# Weather, Macroweather, and the Climate

# Weather, Macroweather, and the Climate

## OUR RANDOM YET PREDICTABLE ATMOSPHERE

Shaun Lovejoy

**OXFORD**

UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

# { CONTENTS }

# { LIST OF BOXES }

# { PREFACE }

In the closing months of the first world war, Lewis Fry Richardson made the first numerical weather forecast, founding the field of numerical weather prediction (NWP). Today, with the help of computers, this brute-force approach has been wildly successful. It is not only ubiquitous in daily weather forecasts, but also has been extended to seasonal predictions through to multidecadal climate projections. It is (almost) the unique tool used to inform policymakers about the climatological consequences of fossil fuel burning and other human impacts.

Yet Richardson was not only the founder of NWP, he also pioneered the development of high-level turbulent laws. In 1926, he proposed the "Richardson 3 law" of turbulent diffusion—a law that wasn't fully vindicated until 2013. Rather than attempting to account for every whirl, cloud, eddy, and structure, the 4/3 law exploits the idea of scaling—a statistical relation between big and small, between fast and slow—to account for and understand the statistical outcome of billions upon billions of structures acting collectively from millimeters up to the size of the planet. Just as the diffusion of milk stirred in a cup of coffee doesn't require tracking every molecule, so too can the atmosphere be *understood* without knowledge of every bump and wiggle on every cloud.

The idea that high-level statistical laws could explain the actions of myriads of vortices, cells, and structures was shared by successive generations of turbulence scientists. Unfortunately, they faced monumental mathematical difficulties largely connected to turbulent intermittency: the fact that most of the activity is inside tiny, violently active regions, themselves buried in a hierarchy of structures within structures. The application of turbulence theory to the atmosphere encounters an additional obstacle: stratification that depends on scale. Although small puffs of smoke seem to be roughly roundish—or even vertically aligned—on a good day, even the naked eye can make out wide horizontal cloud decks that allow us to glimpse chunks of giant strata thousands of kilometers across.

The 1980s marked a turning point when Richardson's deterministic and statistical strands parted company, and when the precarious unity of the atmospheric sciences was broken. On the one hand, computers revolutionized NWP, making the brute-force approach increasingly practical and hence prevalent. On the other hand, the nonlinear revolution—itself a tributary of computers—promised to tame chaos itself, including turbulent chaos with its fractal structures within structures. Throughout the next decades, scientific societies promoted nonlinear science by establishing nonlinear processes divisions and journals. While the nonlinear approaches were advancing understanding, NWPs mushroomed and extended their number crunching to include oceans and the climate.

This book is an insider's attempt to reunite the two strands. It contains some history and a few human touches, but mostly it explains, as simply as possible, how we can *understand* atmospheric variability that occurs over an astonishing range of scales: from millimeters to the size of the planet, from milliseconds to billions of years. The variability is so large that standard ways of dealing with it are utterly inadequate. In 2015, it was found that classic approaches had underestimated the variability by the astronomical factor of a quadrillion (a million billion).

Although familiar treatments focus on a series of "scalebound" mechanisms, each operating over a narrow range of scales ranging from meteorological fronts to convective cells to storm systems—or from El Niño to global warming—in this book I take you by the hand and show you the atmosphere in a new light. Helped by high-level scaling laws operating over enormous ranges of scales from small to large, from fast to slow, I explain this new thing called "macroweather" and describe how it sits in between the weather and climate, finally settling the question: What is climate? I discuss how agriculture—and hence civilization itself—might be a result of freak macroweather.

I answer Richardson's old question: Does the wind have a velocity? And the newer one: How big is a cloud? The answer turns out to explain why the dimension of atmospheric motions is $D = 23/9 = 2.555...$, which is more voluminous than theoreticians' flat value $D = 2$, yet less space filling than the human-scale value $D = 3$. I show that Mars is our statistical twin and why this shouldn't surprise us. I explain how the multifractal butterfly effect gives rise to events that are so extreme they have been called *black swans*. I show how—even accounting for the black swans—we can close the climate debate by statistically testing and rejecting the skeptics' giant natural fluctuation hypothesis. I explain how the emergent scaling laws can make accurate monthly to decadal (macroweather) forecasts by exploiting an unsuspected but huge memory in the atmosphere–ocean system itself. I playfully imagine a 1909 International Committee for Projecting the Consequences of Coal Consumption to show how a good scenario of *economic* development might have led—one hundred years in advance—to accurate projections of our current 1°C of global warming, and I'll show how the same scaling approach can help to reduce significantly the large uncertainties in our current climate projections to 2050 and 2100.

This book is aimed at anyone interested in the weather and climate; it assumes only some basic mathematics: power laws and their inverse, logarithms. However, for those who wish to delve beyond the basic narrative, there are extensive footnotes and endnotes. The footnotes are reserved for supplementary—but nontechnical—information, comments, and occasional anecdotes. The endnotes are more technical in nature, aimed at readers who want to dig deeper. In addition, there are also more than a dozen "boxes" that give even more technical information and explanations. Although they are placed in the text at advantageous locations, they

are designed to be "stand-alone" and can be either skipped or read in any desired order. Overall, there was an attempt to make the book interesting and accessible to readers with a wide range of backgrounds.

The book will have achieved its goal if you achieve a new, unified understanding of the atmosphere and if it convinces you that the atmosphere is not what you thought.

# { ACKNOWLEDGMENTS }

Throughout the years I have had the help and support of many individuals, including students; colleagues; my wife, Hélène; daughter, Vanda; and son, Miro. I especially thank Hélène, who not only faithfully supported me over the decades, but also used her science knowledge and science writing skills to read early drafts of this manuscript critically. I also acknowledge the long-term support of my mother, Margot Lovejoy, an artist who introduced me to fractals, and my father, Derek Lovejoy, a physicist with whom I regularly exchanged on this project but who, unfortunately, did not live to see its completion.

I also thank my long-time friend and colleague Brian Watson, who went through the manuscript painstakingly, pointing out numerous points for improvement. My graduate students Raphael Hébert and Lenin del Rio Amador, helped not only by participating in some of the research (Chapter 7), but also by giving useful feedback on this project. I also thank Joanna Haigh, Director of the Grantham Institute on Climate Change, for hospitality during several months of my 2018 sabbatical leave during which much of this book was written.

Excerpts of a draft of this book were assigned as readings in the course "Environmental Thought," which I taught at McGill University with Darin Barney and Greg Mikkelsen in winter semester 2018. They, and several students, are thanked for their support and critical comments—in particular, Sophie Gagné-Landmann and Chelsea Kingzett.

I can trace the origin of my research into atmospheric scaling to a precise moment: when my mother gave me Mandelbrot's *Fractals: Form, Chance and Dimension* as a Christmas present in 1977; this and Mandelbrot's 1982 book *The Fractal Geometry of Nature* were a source of inspiration throughout. Smitten by fractals, encouraged by my PhD supervisor Geoff Austin, and later by a short collaboration with Mandelbrot and a long one with Daniel Schertzer, this work has now spanned more than four decades. It has involved many people, foremost my collaborators: Geoff Austin, Gunther Bloeschel, Warren Curry, Anthony Davis, Isabel de Lima, Nicholas Desaulniers-Soucy, George Fitton, Philip Gabriel, Jean-Sébastien Gagnon, Hélène Gaonac'h, August Gires, Pierre Hubert, Nikola Jacjay, Philippe Ladoy, Fabrice Lambert, François Landais, Daniel Lavallée, Marc Lilly, Benoit Mandelbrot, David Marsan, Peter Muller, Catherine Naud, Sean Pecknold, Julien Pinel, Alexander Radkevitch, Salvadori Ratti, Gianfausto Salvadori, Daniel Schertzer, Fréderic Schmidt, François Schmitt, Laurent Seuront, John Stix, Jonathan Stolle, Kevin Strawbridge, Ana Tarquis, Ioulia Tchiginskaya, Anastasio Tsonis, Adrian Tuck, and Costas Varotsos.

# Weather, Macroweather, and the Climate

{ 1 }

# Zooming through scales by the billion

## 1.1 What is weather? What is climate?

### 1.1.1 HIGH LEVEL OR LOW LEVEL?

"The climate is what you expect, the weather is what you get"[a]: The climate is a kind of average weather.[1] But is it really? Those of us who have thirty years or more of recall are likely aware of subtle but systematic changes between today's weather and the weather of their youth. I remember Montreal winters with much more snow and with longer spells of extreme cold. Did it really change? If so, was it only Montreal that changed? Or did all of Quebec change? Or did the whole planet warm up? And which is the *real* climate? Todays' experience or that of the past?

The key to answering these questions is the notion of scale, both in time (duration) and in space (size). Spatial variability is probably easier to grasp because structures of different sizes can be visualized readily (Fig. 1.1). In a puff of cigarette smoke, one can casually observe tiny wisps, whirls, and eddies. Looking out the window, we may see fluffy cumulus clouds with bumps and wiggles kilometers across. With a quick browse on the Internet, we can find satellite images of cloud patterns literally the size of the planet. Such visual inspection confirms that structures exist over a range of 10 billion or so: from 10,000 km down to less than 1 mm. At 0.1 mm, the atmosphere is like molasses; friction takes over and any whirls are quickly smoothed out. But even at this scale, matter is still "smooth." To discern its granular, molecular nature, we would have to zoom in 1,000 times more to reach submicron scales. For weather and climate, the millimetric "dissipation scale" is thus a natural place to stop zooming, and the fact that it is still much larger than molecular scales indicates that, at this scale, we can safely discuss atmospheric properties without worrying about its molecular substructure.

---

[a] Often attributed to Mark Twain, but apparently it originally appeared much more recently in a science fiction novel by Heinlein, R. A. *Time Enough for Love.* (G. P. Putnam's Sons, 1973).
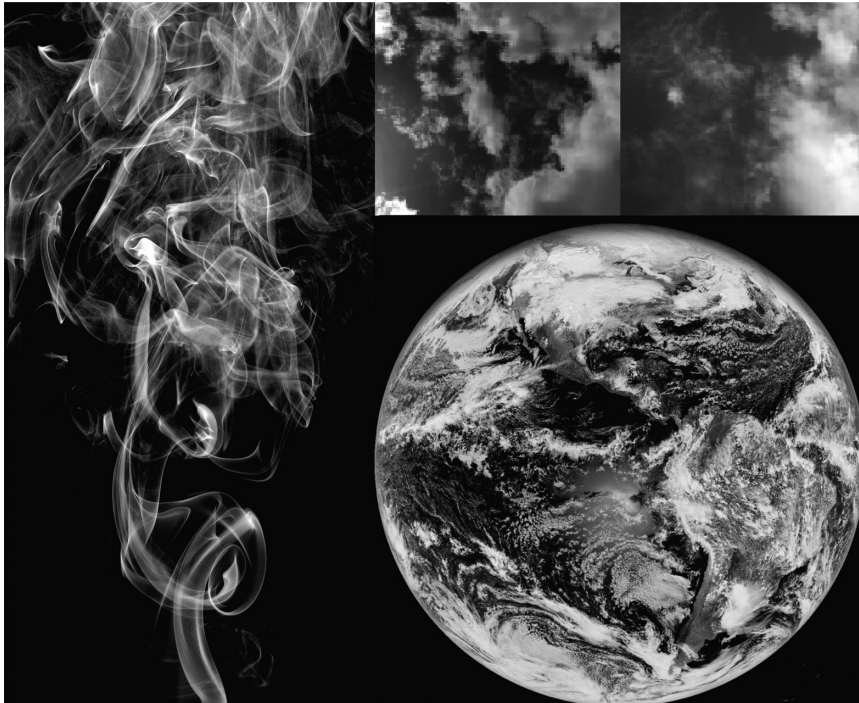
FIGURE 1.1 *(Left) Cigarette smoke showing wisps and filaments smaller than a millimeter up to about a meter in overall size. (Upper middle and right) Two cloud photographs taken from the roof of the McGill University physics building. Each cloud is several kilometers across, with resolutions of a meter or so.[2] (Lower right) The global scale arrangement of clouds taken from an infrared satellite image of Earth, with a resolution of several kilometers.[b] Taken together, the three images of the clouds and Earth span a range of a factor of nearly a billion in spatial scale.*

Clouds are highly complex objects. How should we deal with such apparent chaos? According to Greek mythology, at first there was only chaos; cosmos emerged later. Disorder and order are thus ancient themes of philosophy, physics, and . . . atmospheric science. In the case of clouds, their complex appearances are the result of turbulence. Turbulence arises when fluids are sufficiently stirred. Large-scale structures become unstable and break up into smaller ones. Neighboring structures can interact in complex ways. We can get a feel for this by putting a drop of milk in coffee. If you add the drop gently into a cup of calm black coffee, the milk diffuses very slowly. A brief, wide circular motion rapidly creates a homogeneous milky brew. Although the stirring only directly created a large

---

[b] The satellite picture was taken at infrared wavelengths. The colors are false. The whiteness depends on the coldness (and hence altitude) of the cloud top.

cup-size structure or whirl, it breaks up quickly, creating smaller and smaller ones that disperse the milk rapidly. When these eddies are a bit smaller than a millimeter, molecular diffusion takes over, making the mixture uniform. An important aspect of turbulence is that it typically leads to structures over huge ranges of scale, yet from the point of understanding the system, most of these structures are unimportant details. Rather than acquiring an understanding of each bump and wiggle, what we really need to understand are their *statistics*.

Cloud complexity is analogous to that encountered when considering matter at molecular scales. Figure 1.2A (left) shows a few molecules in a box. Classically, to understand the system at particle scales requires knowledge of the trajectories of each of the molecules.[c] Using a brute-force approach called *molecular dynamics*, today's computers can handle such systems numerically with up to about a million molecules (Fig. 1.2A, middle), yet the systems are still tiny—a billionth of a nanogram or so. But describing and modeling matter at human scales is not just a question of computing power. It's a conceptual problem. Here's a mundane example: I heat my coffee and want to know what happens. Even if I knew the position and velocity of all of its million, billion, billion molecules, it would be mostly useless information. I would be happy with a single number: the temperature. Similarly, even if one could model every single eddy and structure in the atmosphere, every bump and wiggle on every cloud, it would still not answer the questions: Is it a nice day? Will it rain tomorrow? We need not only to eliminate the useless information, but also we must also use the appropriate concepts.

Fortunately, starting in the nineteenth century, techniques have evolved for dealing with huge numbers of molecules: statistical mechanics based on probabilities. If all we want are averages, then we may simplify even further and use thermodynamics and continuum mechanics. The latter theories were developed before the establishment of the atomic theory of matter and make no reference to its particulate nature. Although they treat matter as though it was a homogenous "continuum," this assumption is nevertheless the basis of all theories and models of the weather and the climate. In the jargon of complexity science, they are *emergent* laws and their new concepts such as temperature and pressure are *emergent properties*. Figure 1.2A (right) shows a typical continuum mechanics simulation of an isolated vortex. Even if scientists could make a (low-level) molecular dynamics simulation of this vortex by describing each molecule, there would be no point. To understand it, we need a higher level description that averages out the unimportant molecular details.[3] And although there may be no mathematically rigorous proof, it is widely accepted that there is no contradiction between the higher level continuum description and the lower level molecular one. Both are valid.

---

[c] A quantum mechanical description involves corresponding wave functions.
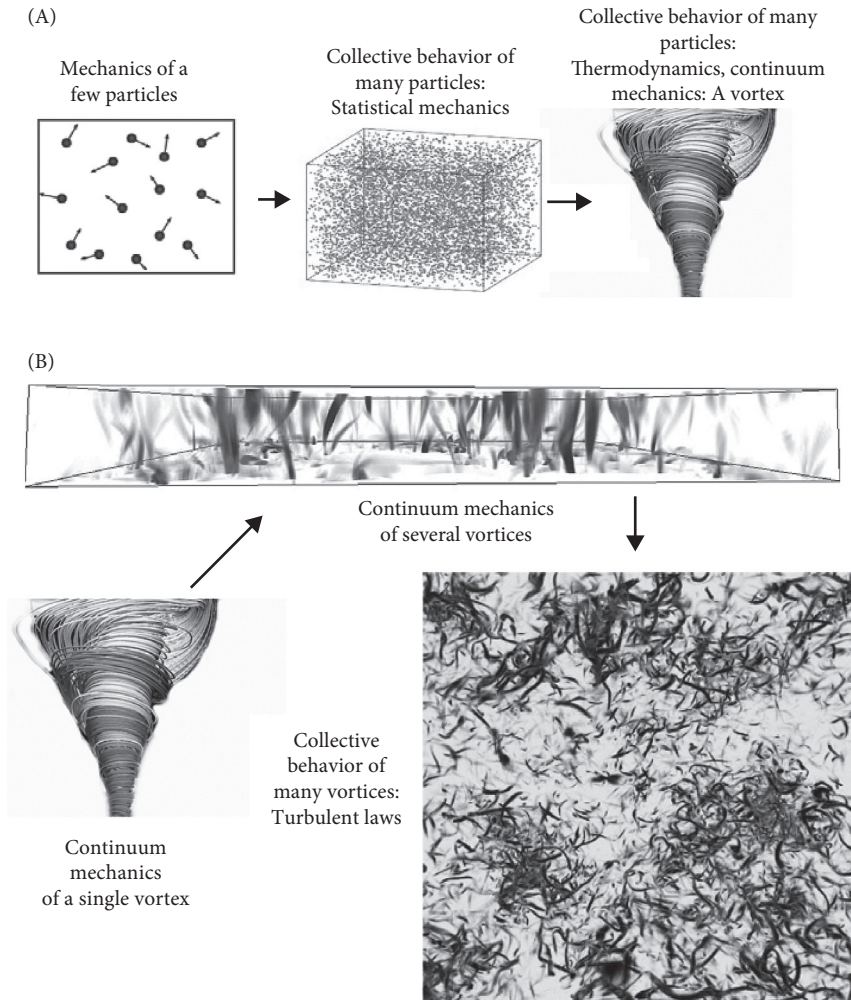
FIGURE 1.2 *(A, left) A particle (molecule) depiction of a gas at molecular scales (nanometers). This is a classic "billiard ball"–type of description. A quantum description involves wave functions that determine the probability of finding the particles in different locations. (A, middle) At somewhat larger scales, we have hundreds of thousands of particles, and we may start to model the system using use statistical mechanics. (A, right) At macroscopic scales, perhaps tens of centimeters or larger, there are so many particles that we can average over the molecular fluctuations and use a continuum description. Here, we see a vortex.[4] This level of description assumes the air is smooth, not granular (i.e., that it ignores its molecular nature). (B, top) Many interacting vortices can still be handled computationally on the basis of continuum mechanics, but the evolution is complex and becomes difficult to understand in a simple mechanistic manner. Each elementary vortex (B, bottom left)[5] is a bit like the molecules in (A, left). (B, bottom right) In this image, we are nearing the strong turbulence limit relevant in the atmosphere. Although this is still a supercomputer simulation, we can already see the problem of huge numbers of interacting vortices. As a result of the seemingly random collection of long, thin vortices, this turbulent view is sometimes called the "spaghetti" picture. This is a direct numerical simulation[6] that models the smallest relevant scales explicitly. Numerical models of the atmosphere in general are only approximations; their smallest scales may be 10 km across. Implicitly, they average over large numbers of structures.*

But the atmosphere is not an isolated single vortex (Fig. 1.2B, bottom left) or even a manageably small collection of vortices (Fig. 1.2B, top). Rather, it is more like the turbulent "spaghetti plate" picture (Fig. 1.2B, bottom right) composed of huge numbers of vortices (eddies, structures). In the atmosphere, typical estimates of the number of interacting components—"degrees of freedom"—are a billion, billion, billion ($10^{27}$), or about the same as the number of molecules in a cubic meter of air. Attempts to understand the corresponding statistics are at the origin of even higher level turbulent laws. Just as statistical mechanics treats large numbers of interacting particles and extracts the important aspects in the form of statistics, so do turbulent laws, which describe the collective behavior of huge numbers of interacting eddies (structures, vortices). Lewis Fry Richardson (1881–1953) proposed the first turbulent law during the 1920s. The saga of how it was finally vindicated in 2013 is told in Chapter 4. When Richardson's law and others of the same type are suitably generalized, they describe the statistical properties of the atmosphere and climate over wide ranges of scales in space and in time. Just as continuum mechanics is a high-level law emergent with respect to the low-level (fundamental) laws of particles and statistical mechanics, the turbulent laws are high-level laws that are emergent with respect to those of continuum mechanics. Once again, the laws at different levels are believed to coexist, to be equally valid.[d] Just as the continuum laws allow new and unprecedented means for understanding, modeling, and forecasting a vortex, the same is true of the turbulent laws for large collections of vortices. Notice the alternation in Figure 1.2 as one moves up the scale from particle laws (deterministic[e]) to statistical mechanics (statistical), to continuum mechanics (deterministic), to turbulent laws (statistical, also called *stochastic*[f]).

In practice, one chooses to use one description (level) or another, one model or another, depending on the application. The graphs in Figure 1.3 illustrate this. At the top, there is an actual wind trace, from measurements taken half a millisecond apart, that displays complex variability. This curve is from real-world data; it is neither random nor deterministic. These are attributes of theories and models. The scientist's job is to find the theory and model that best fits reality, and to do this without a priori prejudices about whether it ought to be deterministic or stochastic[7] (see Box 1.1).

---

[d] I say "believed" because, although there are many arguments and much evidence, there is no mathematical proof of the equivalence.

[e] This means that its evolution follows a rigid rule. This is the classical description. I could have started at the even more fundamental quantum level, which—being statistical—again alternates (see Box 1.1).

[f] Usually, the term *statistical* is used for descriptions of randomness, whereas the term *stochastic* is used for processes and models that incorporate randomness.
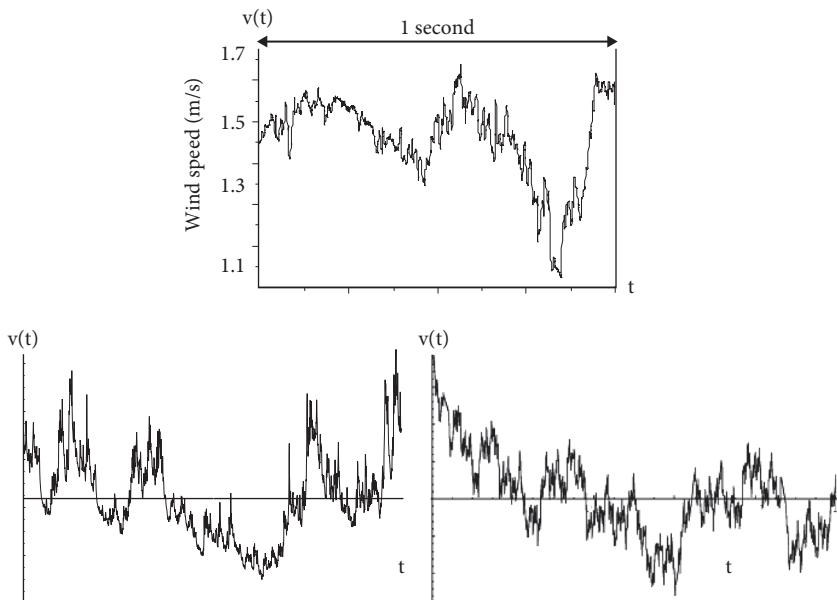
FIGURE 1.3  *(Top) The trace of the wind speed on the roof of the McGill University physics building measured at 2 kHz (2,000 points more than the 1 second of record that is shown). Note the detailed structure. The signal is still not smooth. (Lower left) Here there are 2,048 points of a stochastic (random) multifractal simulation with the same fluctuation exponent (H = 1/3) and measured multifractal parameters (see Box 2.2). (Lower right) Here, there are 2,048 points of a deterministic fractal model, the "Weierstrass function,"[8] with the same fluctuation exponent (Section 2.1).*

BOX 1.1  Which chaos?

**ORDER VERSUS DISORDER AND THE SCIENTIFIC WORLDVIEW**

Although gods were traditionally invoked to explain cosmos and chaos (order and disorder), their role was drastically diminished with the advent of the Newtonian revolution. Sir Isaac Newton himself only required God's assistance for the determination of the initial conditions—for example, starting the planets off in their orbits—the laws of mechanics and gravity would do the rest. By the middle of the nineteenth century, Newtonian laws had become highly abstract, whereas the corresponding scientific worldview had become deterministic. The most extreme views have been attributed to Pierre-Simon Laplace (1749–1827), who went so far as to postulate a purely deterministic universe in which "if a sufficiently vast intelligence exists,"[9] it could solve the equations of motion of all the constituent particles of the universe. In this universe, such a divine calculator could determine the past and future from the present in an abstract high-dimensional "phase space" that defines precisely the position and velocities of every particle and at all times.

Unfortunately, as part of their imperfect knowledge, mortals are saddled with measurement errors that involve notions of chance. The identification of chance with human error led to Voltaire's (1694–1778) contemptuous: "chance is

nothing."[g] By this, he meant that chance is purely subjective. As far as the laws of nature are concerned, it is beneath consideration. It was a full century later that James Clerk Maxwell (1831–1879) introduced probability explicitly into the formulation of physical laws: the distribution of speeds of gas molecules (in 1870). This is the origin of classic statistical mechanics—the idea that the unobserved or unknown degrees of freedom ("details") are the source of random behavior, such as fluctuations about a mean temperature. Although highly partial information is the rule, human-size (macroscopic) objects are typically described by parameters such as temperature, pressure, and density. Most of the details—especially particle positions and velocities—are unimportant and can be reduced to various averages using statistics, hence the dichotomy of objective deterministic interactions of a large number of degrees of freedom coexisting with randomness arising from our subjective ignorance of the details.

The identification of statistics with ignorance evolved—notably thanks to Josiah Willard Gibbs (1839–1903) and Ludwig Boltzmann (1844–1906)—to the more objective identification of statistics with the *irrelevance* of most of the details. For example, Henri Poincaré (1854–1912) rejected the notion of chance as ignorance, and instead favored the view that chance is appropriate when dealing with complex causes. Giving the example of the weather and anticipating Edward Lorenz's (1917–2008) Brazilian butterfly causing a Texas tornado, he stated: "It may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will result in an enormous error in the latter. Prediction becomes impossible and we have the fortuitous [chance] phenomenon."[h] In Chapter 7, we see that this stochastic conclusion was more nuanced than Lorenz's later deterministic one. A little later, Andrey Kolmogorov[10] (1903–1987), perhaps motivated by fluid turbulence, "axiomatized" probability theory, proving that chance could be purely objective. This lent rigor to the work on Brownian motion initiated during the early twentieth century by Louis Bachelier, Albert Einstein, Marian Smoluchowski, and Paul Langevin. It laid the mathematical groundwork for dynamical *stochastic processes* that obeyed objective probabilistic laws.

But even these advances did not alter the deeply held prejudice that statistics and statistical laws were no more than a feeble substitute for determinism. A corollary to this was the hierarchical classification of scientific theories. Fundamental theories were deterministic; only the less fundamental ones involved randomness. Although this poor man's view of randomness might have been appropriate to classic statistical physics, with its billiard ball atoms and molecules, it is incapable of accounting for the fact that the most fundamental physical theory yet devised—quantum mechanics—which is in fact probabilistic,[11] with its fundamental object (the wave function) specifying probabilities.[i] Even today, although numerous developments make it obsolete, determinism is entrenched.

---

[g] He continues: "We have invented this word to express the known effect of unknown causes" (p. 439). Voltaire. *The Ignorant Philosopher*. (1766). (in English: Blue Ribbon Books, 1932.)

[h] Poincaré was a determinist. He was explaining why the weather "seemed" random; he was not arguing for a statistical theory of meteorology. Poincaré, H. *Science and Method*, p. 296. (Thomas Nelson and Sons, 1912).

[i] Not surprisingly, there have been attempts to explain quantum randomness by deterministic chaos acting at a subquantum level.

### THE DETERMINISTIC CHAOS REVOLUTION: THE BUTTERFLY EFFECT

Just as biases toward determinism were starting to unravel, the deterministic chaos revolution of the 1970s and '80s gave them a new lease on life. The trouble started with attempts to solve the equations of motion. Poincaré had already noted that three particles are already too many to allow for nice, regular "analytical" solutions. However, it was Lorenz, in 1963, who fully recognized the general propensity of nonlinear systems to amplify small perturbations ("sensitive dependence on initial conditions")—a fact that only became widely known during the 1970s as the "butterfly effect" (see Chapter 7). Even if Laplace's calculator had *nearly* perfect initial information and had infinitely precise numerics, the future would *not* be predictable. Randomlike "chaotic" behavior would result instead. During the 1950s, Lev Landau had proposed that fluid turbulence arose through an infinite series of instabilities. Now, thanks to Lorenz, it became clear—as the title of a famous article proclaimed[12]—that "three implies chaos." However—and this crucial point is often overlooked—even if only three instabilities are *necessary*, the state of "fully developed fluid turbulence"[13] that approximates the atmosphere still depends on a huge number of components.

To become a fully fledged revolution, in addition to the butterfly effect, chaos theory required two more developments. The first was the reduction of the scope of study to systems with a small number of degrees of freedom. The second was the discovery that, under quite general circumstances, qualitatively different dynamical systems could give rise to quantitatively identical behavior: the celebrated "Feigenbaum constant."[14] This "universality"[15] finally allowed for quantitative empirical tests of the theory. By the early 1980s, these developments had led to what could properly be called the *chaos revolution*.

### LATER DEVELOPMENTS AND PROBLEMS

The basic outlook provoked by the developments in chaos—that randomlike behavior is "normal" and not pathological—is valid regardless of the number of degrees of freedom of the system in question. The success of chaos theory on systems with a small number of degrees of freedom led to some bold prognostications, such as "junk your old equations and look for guidance in clouds' repeating patterns."[16] This fervor was unfortunately accompanied by a drastic restriction of the scope of chaos itself to meaning precisely deterministic systems with few degrees of freedom. The restriction, coupled with the development of new empirical techniques, led to a major focus on applications and a number of curious, if not absurd, claims.

A particularly striking example comes from the climate, for which it had always been assumed that a large (practically infinite) number of interacting components were involved. However, when new chaos tools were applied to paleoclimate data, a famous paper even claimed that only four degrees of freedom were required to specify the state of the climate.[17] It was later pointed out that the conclusion was based on empirical analysis involving only 180 data points, and that even within the deterministic paradigm, this was far too few to justify the conclusions. It was barely noticed that, in any case, stochastic processes with an infinite number of degrees of freedom could have easily given the same result.[18]

During the same period, in geoscience and elsewhere, numerous attempts were made purely from data analysis to prove that, despite appearances, randomlike signals were in fact deterministic in origin.

These attempts were flawed at several levels, the most important of which is philosophical: the supposition that nature *is* (ontologically) either deterministic or random. In reality, the best that any empirical analysis could hope to demonstrate was that specific deterministic models fit the data better (or worse) than specific stochastic ones.

### THE ALTERNATIVE FOR LARGE NUMBERS OF DEGREES OF FREEDOM: STOCHASTIC CHAOS

By the mid 1980s, the ancient idea of chaos had thus taken on a narrow, restrictive meaning that essentially characterized deterministic systems with small numbers of interacting components. The philosophy underlying its use as a model for complex geophysical, astrophysical, ecological, or sociological systems—each involving nonlinearly interacting spatial structures or fields—has two related aspects, each of which is untenable. The first is the illogical inference that because deterministic systems can have randomlike behavior, randomlike systems are best modeled as deterministic. The second is that spatial structures that apparently involve huge variability and many degrees of freedom spanning wide ranges of scale can in fact be effectively reduced to a small number. At a philosophical level, deterministic chaos is a rearguard attempt to resurrect Newtonian determinism.[j]

To overcome the limitations of deterministic modeling, I propose to make systematic use of models with objective randomness: "stochastic chaos."[19] The fundamental characteristic of stochastic theories and models that distinguishes them from their deterministic counterparts is that, mathematically, they are defined on probability spaces, which implicitly involve huge numbers of components.[20] In comparison, their deterministic counterparts—at least when strongly nonlinear—are only *manageable* with small numbers of interacting components (degrees of freedom). The stochastic chaos alternative is now easy to state: Contrary to Einstein's injunction that "God does not play dice," we seek to determine *how* God plays dice with large numbers of interacting components (Fig. 1.4). Deterministic and stochastic models of the same physical system can coexist at different levels in a hierarchy of scientific theories. One can choose the level that is most convenient for the intended application.[21]

---

[j] In a popular book, the father of deterministic chaos, Edward Lorenz, even flirts with teleology—the idea that freewill is an illusion: "We must then wholeheartedly believe in free will. If free will is a reality, we shall have made the correct choice. If it is not, we shall still not have made an incorrect choice, because we shall not have made any choice at all, not having a free will to do so" (p. 160). Lorenz, E. *The Essence of Chaos.* (UCL Press, 1993).

FIGURE 1.4 *God playing dice in geophysics. During the 1930s—the early days of quantum mechanics—Einstein famously proclaimed, "God does not play dice." By this, he meant that, because of its stochastic nature, quantum mechanics could only be a provisional theory, and that it would be replaced eventually by a more fundamental, deterministic one. Today, quantum mechanics is still the most fundamental theory that we have, and the objective nature of such stochastic theories is much clearer than it was in Einstein's day. Stochastic theories can be just as fundamental as deterministic ones. The modern question is: How does God play dice?[22]*

**OBJECTIONS TO STOCHASTIC CHAOS**

*Causality requires determinism*

A common misconception is that causality and determinism are essentially identical or, equivalently, that indeterminism implies a degree of acausality. As emphasized by Bunge,[23] causality is nothing more than a specific type of objective determination or necessity. It, by no means, exhausts the category of physical determination that includes other kinds of lawful production/interconnection, including statistical determination: stochastic causality.

*Structures are evidence of determinism*

Interesting phenomenologically identified large-scale structures—storms, for example—are frequently modeled with mechanistic (and deterministic) models, whereas the presence of variability without noteworthy structures is identified with noise. The inadequacy of this view of randomness is brought home by the still-little-known fact that stochastic models can, in principle, explain the same phenomena. The key is a special kind of "stochastic chaos" involving a scale-invariant symmetry principle in which a basic (stochastic) cascade mechanism repeats scale after scale after scale, from large scales to small scales, eventually building up enormous

variability: multifractals (Box 2.2). Unlike classic stochastic processes, multifractals specifically have extreme events, called *singularities*, that are strong enough to create structures, so they produce potentially large structures that have little in common with featureless white noise.

### Physical arguments for stochastic chaos: Turbulence

We have already seen that because stochastic processes are defined on infinite dimensional probability spaces, stochastic models are a priori the simplest whenever the number of degrees of freedom is large. In geoscience, stochastic chaos is particularly advantageous when—as in fluid turbulence—a nonclassic symmetry is present: scale invariance. This is the thesis of this book.

Consider the bottom graphs in Figure 1.3. On the left, we see a stochastic multifractal model with a basic fluctuation exponent (*H*; discussed in Chapter 2) that is close to the value found in the experimental trace at the top. In this case, it is close to a well-known theoretical turbulence law.[24] On the right, we show a purely deterministic model with the same parameter, but it is based on a simple rule that repeats from large scales to small scales (a fractal). Although the left graph obeys high-level turbulent laws, the right might have come[k] from a (lower level, deterministic) Numerical Weather Model (NWP) or General Circulation Model (GCM[l]). Yet both could be realistic inasmuch as they agree on important statistical properties of the data. In this case, all three graphs in Figure 1.3 turn out to have the same basic relationship between large and small structures: the rule that determines how "fluctuations" change with size. In the data series (Fig. 1.3, top), the relationship can be determined by analysis; in the bottom (Fig. 1.3), the relationship is specified theoretically by the model. The idea is that the detailed wiggles (e.g., the curves in Fig. 1.3) are not of interest, but the way the wiggles typically differ when we move from small to large *is* of interest. Even in these model series, one may notice that their exact characters seem to be a little different. It turns out that the relation between small and large can itself be nontrivial, and a single number (in this case, the fluctuation exponent that we revisit in Chapter 2) is only part of the story.

Another way of viewing this is to ask: Do all the wiggles in Figure 1.3 (top) require detailed explanations? Or, when we consider enough wiggles over a wide range of scales, is there a simplifying principle that comes into play? If either of the bottom curves in Figure 1.3 were realistic wind models, the apparent wind

---

[k] Although, in this case, the finest resolution would be hours, not milliseconds!

[l] The acronym GCM refers to numerical models of the atmosphere that are adapted to timescales longer than the ten-day weather scale, notably by including a model of the ocean. Increasingly, the acronym GCM is alternatively decoded as Global Climate Model. Although GCMs are often distinguished from NWPs, their atmospheric components are fundamentally the same, being run on larger grids and with lower temporal resolutions. For the purposes of this book, the distinction is not always important.

complexity would become much simpler. The emergent laws thus potentially transform the complex into the simple—or, perhaps more accurately—from an unmanageable low-level complexity to a simpler, more manageable higher level *irreducible* complexity (more on this later).

### 1.1.2 SPACE AND TIME

We can use the visual spatial information in Figure 1.1 to learn something about how the atmosphere varies in time. This is because wind—although being far from constant, as Figure 1.3 shows—typically blows at several meters per second. As eddies of all sizes blow past, we feel fluctuations (i.e., variations) in the temperature, humidity, pressure, and, indeed, in the wind itself.[m] At 1 m/s, eddies with millimeter-size temperature variations would cause our thermometer to vary at millisecond scales. Assuming that we had a thermometer with a rapid enough response time, we could measure this, but milliseconds are not enough (Fig. 1.3, top). To see a smooth dissipation scale response, we need a thermometer that measures 10,000 times a second.[25] Pursuing the idea of blowing structures, a cloud outside our window might cross our line of sight in tens of minutes. Structures 10 km in size readily account for fluctuations of this duration. Taking this to its extreme—the size of a planet—we obtain a typical estimate of about ten days for the fluctuations that it implies (corresponding to a mean speed of about 10 m/s). This simple argument already accounts for the temporal variability between a thousandth and a million seconds: another factor of 1 billion.

We have just described the atmosphere's weather regime. Ten days is indeed its limit, and it corresponds roughly to our subjective experience of changing weather and weather patterns, and it is equal to the theoretical limits to which weather can be forecast in a deterministic sense (the discussion of the butterfly effect in Chapter 7). If this weather variability was all there was, then understanding the atmosphere would be much simpler because we could always identify, at least on average, fluctuations of a given duration with those of a corresponding size. But the variability doesn't stop at ten days. Even if we ignore the annual cycle, each month is different from the previous month, each year from the previous year, and this continues through centuries and millennia through to Ice Age scales (tens to hundreds of thousands of years), through to geological scales, right up to the age of the planet—indeed, another factor of 1 billion, an overall range of a billion billion. Incredibly, it wasn't until 2015 that this variability was quantified properly, and it was discovered that, over the range of one hour to the age of Earth, it had been underestimated by a factor of a quadrillion: a one followed by fifteen zeroes[26] (see Chapter 2).

---

[m] This is actually the nonlinear origin of the turbulence: the wind transporting itself.

So at what scale does the climate fit in? It is obviously longer than ten days. Indeed, for most of human history, the climate was considered an essentially constant atmospheric state that characterized a given location at a given time of year. As Mark Twain famously quipped, "Climate lasts all the time and weather only a few days."[27]

The idea of climate constancy—except for geological periods—was still common in 1934, when the newly formed Climate Commission of the International Meteorological Organization[n] adopted 1901 to 1930 as the "climate-normal period" as a quantitative definition of the climate. Ironically, the idea of climate *change* was already dawning, and—following a postwar cooling—climate change was increasingly recognized, leading, for example, to the official American Meteorological Society (AMS) glossary[28] to define the climate as "the synthesis of the weather" and then "the climate of a specified area is represented by the statistical collective of its weather conditions during a specified interval of time (usually several decades)." Not surprisingly, the "climate normal" was shifted shortly thereafter to 1930 to 1960. Now, it is 1980 to 2010. Looking back from today's vantage point, with climate change threatening our future, the view of climate as constant seems quite archaic.[o] Indeed, the main modern update is the proposal "to expand the definition of climate to encompass the oceanic and terrestrial spheres as well as chemical components of the atmosphere."[29]

If the climate is changing, then it does *not* "last all the time," nor can it be "what we expect." So what is it really? And if, as suggested by the AMS glossary—or by fiat from the official duration of climate normals—it only changes on thirty-year scales or longer, then what is the regime in between the weather (up to ten days) and the climate? Incredibly, for an age of supercomputers drowning in petabytes of "big data," the answer has only recently appeared: *macroweather*, a term that is still unfamiliar.[30] We see later that macroweather is defined precisely by the property that had previously been attributed to the climate: that successive fluctuations tend to cancel each other out; that averaging over longer and longer periods reduces the fluctuations, making them *appear* to converge. However; at roughly twenty- to thirty-year periods,[31] the point of apparent convergence itself starts to vary. The successive thirty-year climate normals are, on the contrary, found (thanks to proxy or "paleo" temperature data) to fluctuate almost like the weather all the way to Ice Age scales (100,000 years, Box 1.2). In these pages, I thus explain why we should *not* expect the climate. Expect *macroweather*. But the variability doesn't stop at Ice Age scales, and in Chapter 5 we examine the even longer scale macro- and megaclimate regimes.

---

[n] The precursor of today's World Meteorological Organization.

[o] It persists only as part of a religious-based strand of climate change skepticism. The other strand of skepticism is the exact opposite: that the climate is so variable the warming is claimed to be giant natural fluctuation (GNF)—a claim that is debunked statistically in Chapter 6.

BOX 1.2  Zooming through deep time in the Phanerozoic

Every since Earth was formed 4.5 billion years ago, its temperature has varied. Yet, over most of this time, our knowledge of its temperature is very sketchy. This changed 550 million years ago thanks to the appearance of hard-shelled marine animals at the beginning of the current, Phanerozoic, eon, so we'll concentrate on this.

Figure 1.5A shows reconstructed time sequences of oxygen 18 ($^{18}$O) isotope ratios from calcium carbonate ($CaCO_3$) from paleotemperature ocean floor "stacks." A stack is a collection of cores from different locations in which specific, identifiable common layers (such as tree rings) can be used to produce a common sequence that can be dated. The series in Figure 1.5A is from collections of ocean cores with enough geographical distribution that they are believed to be representative of global conditions. To see how these paleoindicators work, recall that the most abundant oxygen isotope is $^{16}$O. This lighter isotope is evaporated preferentially from the ocean surface, leaving behind water that is enriched in $^{18}$O. The oxygen from this water gets absorbed into the skeletons of tiny creatures (foraminifera, or plankton) and when they die, they fall to the ocean floor, where they accumulate as layered benthic sediments that are later "cored" by drilling ships.

The top series in Figure 1.5A is an update of a global assemblage[32] that covers the entire Phanerozoic. Because lifeforms with $CaCO_3$ exoskeletons did not exist before this, the series goes as far back as this technique will allow. Although 2,980 values were used, they are far from uniformly distributed. Figure 1.5A shows a linear interpolation so that the exactly straight segments are artifacts; they are space-fillers for presumably highly variable missing data. A warning: As is usual in representations of paleodata, the time axis goes from the present at the left to the past at the right. When we approach the modern period—especially when using real (instrumental) data—the axis is reversed so that time increases from left to right.

The vertical axis in Figure 1.5A is 1,000 times the difference in the observed $^{18}$O concentration with respect to a convenient reference concentration, denoted $\delta^{18}$O (the "$\delta$" is thus a tenth of a percent). The basic relationship between $\delta^{18}$O and temperature is an inverse one. Increasing $\delta^{18}$O is associated with decreasing temperatures, although this relationship is complicated by the change in the $\delta^{18}$O composition of seawater as a result of the preferential sequestering of light seawater in ice caps. There are thus both direct and indirect links to temperature via the ice sheets. The temperature range is indicated by the double-headed arrows in the figure. It is based on the "canonical" calibration of $-4.5°C/\delta^{18}$O, and may be too large by as much as a factor of 3.[33] Note the negative sign in the calibration. Large $\delta^{18}$O values correspond to low temperatures and vice versa. (This may be confusing because, in ice cores, the relationship is opposite, as discussed later.) We can see there has been a slow cooling by as much as 40°C during this period. Because this corresponds to roughly one tenth of a degree per million years— assuming that it continues today—it will not do much to alleviate global warming, which is currently occurring at a rate of about 0.1°C per decade (i.e., 100,000 times faster!).

The series show no evidence of fluctuating about fixed values. On the contrary, they appear to "wander" up and down, frequently displaying numerous consecutive fluctuations with the same sign. This is a visual signature of an unstable process. Its

meaning is quantified in Section 2.3, and its implications for the history of Earth are discussed in Chapter 5.[p]

The middle series in Figure 1.5A[34] is from an assemblage covering the Cenozoic era that started 65 million years ago. It is based on deep-sea isotope records from data compiled from more than forty Deep-Sea Drilling Project and Ocean Drilling Project sites. It has 14,828 values and thus gives much more detail than can even be represented visually on the graph. As with the top series in Figure 1.5A, the values are not distributed uniformly in time and they are considered to be globally representative. Note that there are several rather "spiky" excursions. One of them is particularly famous—the Paleocene–Eocene Thermal Maximum (PETM), which is a downward spike (visible on this isotope plot) that took place 55.8 million years ago. Up until the Industrial Revolution, it was the most rapid warming event on record. The warming was estimated to be between 5°C and 8°C, and to be caused by a massive release of carbon into the atmosphere. Indeed, stratigraphy indicates that carbon dioxide ($CO_2$) was released in several geologically short (millennial-scale) pulses. The PETM is thus analogous to today's industrial epoch release, although during the PETM, the planet was nearly ice free and today's release is faster still.

Although Figure 1.5A (middle) indicates symbolically that it is a factor 8.3 blowup of the most recent part of the top series, this is not exactly true because the data are from different cores. It does, however, indicate quite well the type of variation that would be expected in the top series had higher resolutions been available.

The bottom series[35] in Figure 1.5A is a further blowup of a factor of 5.8. It uses 2,560 data points from twelve benthic and five planktic $\delta^{18}O$ records over the Quaternary, mostly from high northern latitudes.[36] The Quaternary is the recent geological period during which (as a result of continental drift) the distribution of landmasses near the poles allowed for the buildup of ice sheets. A particularity of this current geological era is that it alternates between glacial and interglacial (cold and warm, large and small ice caps) periods.

The series has been somewhat smoothed at its smallest scales (this is obvious when it is compared to Fig. 1.5B at an even higher resolution), and this makes its rough oscillations a little more visible. Indeed, this and the blowup at the top of Figure 1.5B are the only series for which some obvious regularity exists. It is associated with astronomical forcings—the "Milanković cycles," named after Serbian geophysicist and astronomer Milutin Milanković (1879–1958), who first proposed them in the 1930s. Without sophisticated analyses, we can see with the naked eye that the most recent 800,000 years or so (in the ellipse in Fig. 1.5, top) have roughly 100,000-year oscillations, whereas the older part of the series has close to 41,000-year oscillations. The latter period is the same as that of small variations in the obliquity (tilt) of Earth's rotational axis, which varies from 22.1° to 24.5° of tilt every 41,000 years. The former is close to a small-amplitude 95,000-year periodic variation in the eccentricity (ellipticity) of Earth's orbit around the sun. It is still not clear why the timing of the recent eight Ice Ages seem to be the result of this small variation, rather than either of the much stronger 41,000-year obliquity or the 413,000-year eccentricity cycles (see Box 5.4).

Zooming in further by factor of 3.2 (Fig. 1.5B, top), we see the longest ice core proxy temperature: the EPICA (European Project for Ice Coring in Antarctica)

---

[p] Notably for James Lovelock's famous "Gaia hypothesis."

Antarctic core determined using a deuterium-based paleotemperature.[37] Deuterium is a heavy isotope of hydrogen (with a proton and neutron in the atomic nucleus); deuterated water is heavier than the much more abundant light water and, as a consequence, it evaporates less easily. Not only is snow enriched in normal water, the amount of enrichment increases with temperature, allowing us to infer roughly the temperature from deuterium measurements.

From the series in Figure 1.5B, we note the loss in resolution (the apparent increase in smoothness) of the curve as we move into the past (to the right); it is an artifact of the compression of the bottom (older) ice in the column and subsequent diffusion of the signal. Clearly, the neat classification of the series into eight glacial and interglacial epochs (indicated by the braces; Fig. 1.5B, top) is a somewhat subjective simplification of the true variability. Up to at least periods of ~100,000 years, we see that the temperature seems to "wander" (e.g., as we consider the change in temperature for increasingly long time periods, the temperature changes more and more). At the same time, we can see that the overall maximum amplitude of the temperature changes (at least at this polar latitude) is about 10 to 12°C, although "typical" 100,000-year variations are closer to ±3°C (see Chapter 5).

A further blowup of a factor of 8.8 takes us to the middle series in Figure 1.5B, from a high-resolution GRIP (Greenland Ice Core Project; from Summit, Greenland) core. The series has more than 17,000 points (roughly every five years) and clearly shows that the variability during the past 91,000 years was far from smooth. One can also note a number of narrow spikes, some of which are associated with "Dansgaard–Oeschger" events in which temperatures can vary by nearly as much as the glacial–interglacial difference, but over time periods as short as fifty years. The mere existence of these still poorly understood events underscores the extreme variability of the atmosphere, even at a fixed timescale. In Chapter 5, we consider whether these are examples of "tipping points" or, rather, "black swans."[q]

The current interglacial period is at the far left of the GRIP series, and an ellipse indicates the most recent 12,000-year period. This is the Holocene, the current warm interglacial period in which civilization arose. A blowup by a factor of 91 of the GRIP series is displayed in the bottom series of Figure 1.5B, showing the most recent millennium. From now on, the direction of time is in the usual left-to-right direction. This Last Millennium series is an example of a multiproxy temperature "reconstruction"[38] (Box 5.3). The series is at annual resolution and it uses data from a wide northern hemisphere geographical distribution, but also from a variety of different paleoindicators (hence the prefix "multi"). Paleoindicators include data from varves (lake sediments), ice cores, dendrochronology (tree rings), and pollen, among others. At around the year 1600, we notice a cooler period—the "little Ice Age." To the right (the circle), we notice the industrial epoch and the phenomenon of global warming.

The zoom sequence just discussed is for temperature proxies, and we have seen that there are many complex issues of interpretation and calibration. The different series are from different locations and often have highly uneven resolutions and sampling

---

[q] Naseem Taleb introduced the term "black swan events." Black swans are metaphors for the impossible. Black swans are discussed in Box 3.1.See Taleb, N. N. *The Black Swan: The Impact of the Highly Improbable.* (Random House, 2010).

intervals. A basic problem faced by all temperature proxies is that, over time, the basic isotopic signal diffuses—gets smeared out—so that the resolution necessarily decreases. The EPICA temperature proxy series, for example, has 3,300 points over the last 100,000 years, but only 137 from 700,000 to 800,000 before present.

To make the zoom simpler to interpret, we can take advantage of a different proxy: estimates of the mass of dust particles deposited per area per unit time, also called the "dust flux." Dust particles are big enough that they don't diffuse significantly, so their resolution can be high even in highly compressed ancient layers of ice. In Figure 1.5C, we note dust data from the same EPICA core but at 1-cm resolution (every centimeter for 3 km: a total of 316,000 data points). To my knowledge, this is the longest single climate data series analyzed to date. We are able to zoom out by a factor of more than 1,000 (upper left to lower right in Fig. 1.5C) with few missing data points. The drawback of the dust flux is that, although it is well measured, its interpretation in terms of the climate state is currently not very clear. It depends, for example, on prevalent wind direction and ~~volcanic eruptions~~ Although the dust flux shows a fairly strong periodicity of about 100,000 years (corresponding to the glacial–interglacial transitions), we clearly see that the strong variability continues at all observed scales.
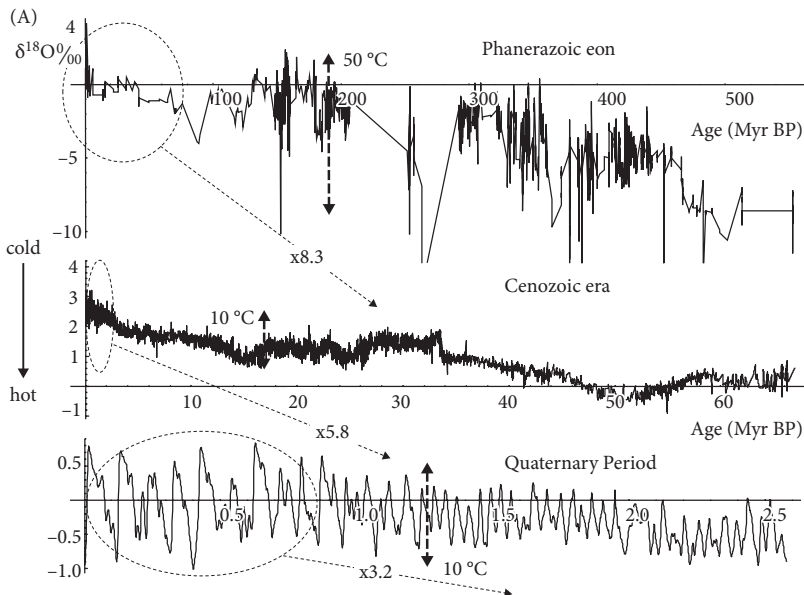


FIGURE 1.5  *(A) δ^18O from assemblies of cores from ocean sediments of benthic organisms, as discussed in the text. The resolutions (here, the time intervals between successive measurements) are variable, but on average are about 200,000 years (top), 4,000 years (middle), and 1,000 years (bottom). Large values correspond to colder temperatures and vice versa. The ellipses, arrows, and numbers indicate the parts of the time axis as well as the zoom factor needed to go from one series to the next. This figure continues in (B). Myr BP, million years before present.[39]*
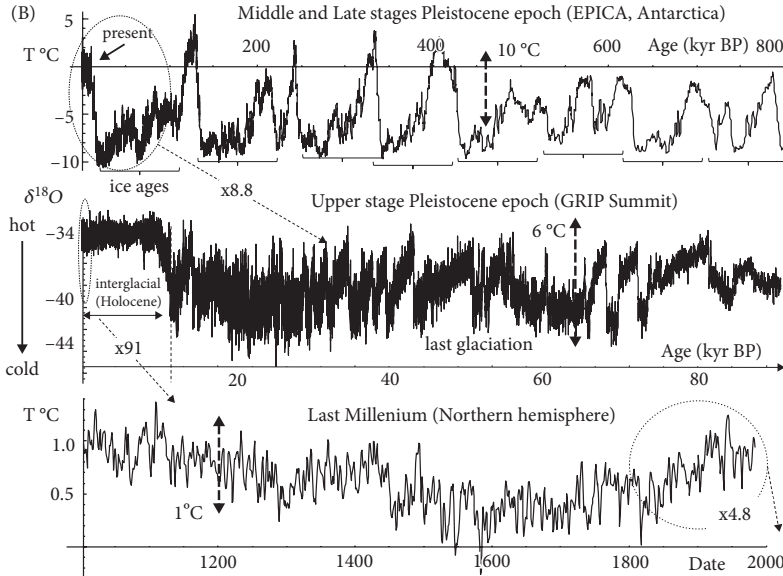
(B) *The top series is the temperature anomaly from the EPICA (European Project for Ice Coring in Antarctica) Antarctic core using deuterium-based paleotemperatures. The anomalies are measured in degrees Celsius. The resolution of the top varies from less than one hundred years (recent data) to about 1,000 years (oldest). In the middle series, it is (a nearly constant) five years; in the bottom series, it is one year. This figure is continued in (D).[40] (C) Zooming out of the Holocene (upper left) with the longest (nearly) complete atmospheric series: 3.2 km of dust fluxes at 1-cm resolution from the Antarctica (EPICA) ice core—the same core that was used for the paleotemperatures in (B, top). The paleotemperature series rapidly loses its resolution with depth; in contrast, the dust flux maintains its high resolution throughout the entire 800,000-year-long core). I zoom out by factors of 2 in depth (the first few are shown by ellipses and arrows).[41] After each zoom, the resolution is degraded by the same factor (2) to maintain 290 points at all times. We then normalize the result by the average value.[42]. All the series start at 373 years before present (BP), and the numbers indicate the number of years that the series covers. For example, the upper left curve covers the period from 373 to 442 BP (moving from left to right).[43] (D) The top series represents the longest available, instrumentally based global temperature estimates (monthly, land only, 3,129 values) for 1753 to 2013. The white line in the middle of the zig-zagging black line (top) is the annual average temperature. All temperatures are measured in degrees Celsius. (Middle) The temperature over Montreal at an altitude of about 3 km as determined by the Twentieth-Century Reanalysis (20CR; a data–model hybrid); the resolution is every six hours from 1871 to 2008. We mainly see the annual cycle. (Bottom) The same as the middle series, but with the annual cycle removed, displaying the larger residual variability. Also shown for reference is the average anthropogenic warming trend over the period.[44] (E) The upper left is the same as the lower series in (D). We successively take the left half of the series and blow it up by a factor of approximately 2, retaining 720 points at each step until we get to the six-hour-resolution series (bottom left). The total length of each series is indicated in red. The bottom right series is from Montreal, but from a millimeter-size thermistor on the roof of the McGill University physics building at 0.067-second resolution. The temperature scale is the same for all series except for the lower right, where it has been reduced by a factor of 10. If higher resolution data were available, the variability would continue for at least another two orders of magnitude to millisecond scales (see Fig. 1.3 for the wind).[45]*

AQ: Please confirm the layout of Figure 1.9

AQ: Please clarify reference to color. This is a B&W figure.

(C)



(D)
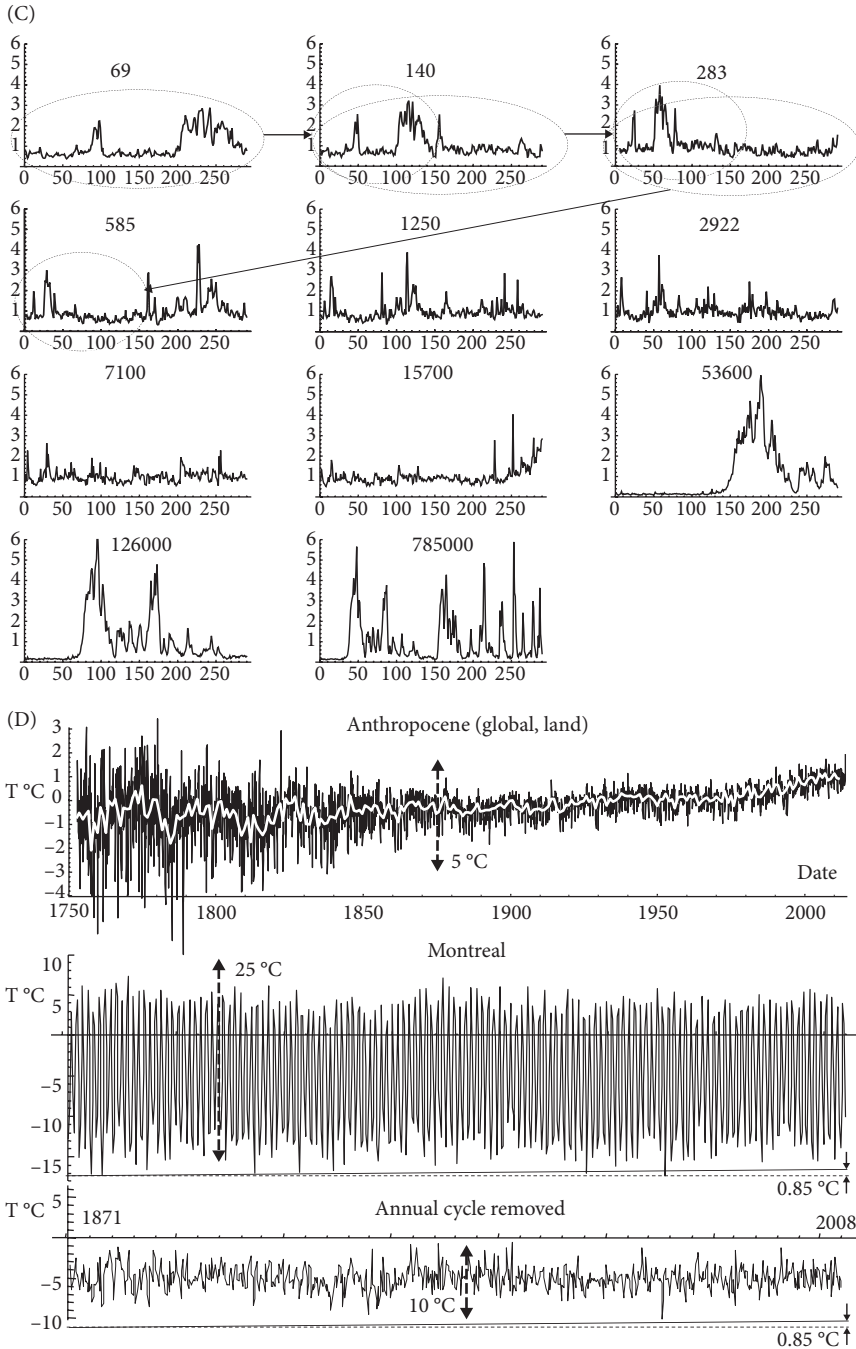


FIGURE 1.5  *Continued*
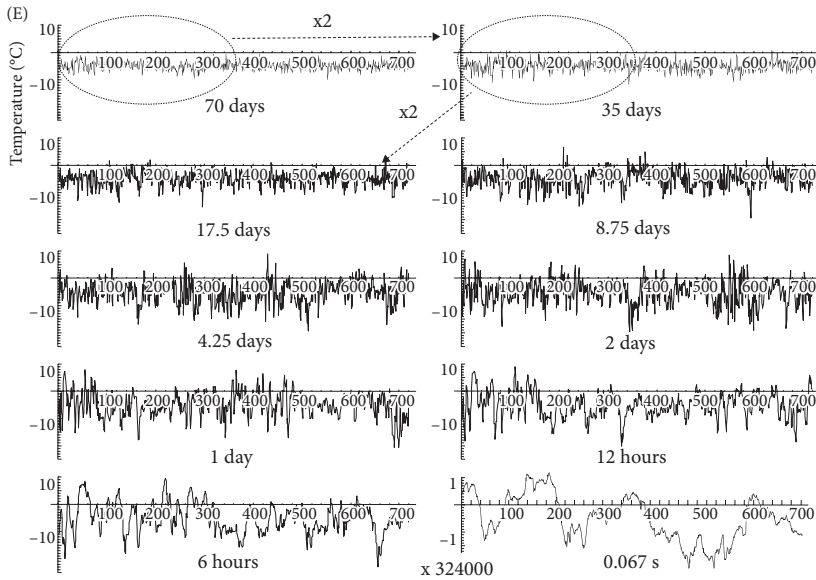
FIGURE 1.5 *Continued*

## 1.2 Zooming in time: The age of Earth to milliseconds
### 1.2.1 PALEOINDICATORS

Variability ranging from milliseconds to billions of years and from millimeters to the size of Earth is difficult even to perceive, let alone to analyze and understand. Before returning to sizes, let us take a voyage through scales in time. For the voyage, we are fortunate that, at least since the beginning of the Phanerozoic, we have useful "paleo" (i.e., proxy, surrogate) temperature data. The Phanerozoic eon is the current geological period in which abundant animal life has existed. It goes back 550 million years to the time when diverse hard-shelled marine animals first appeared.

Box 1.2 details a series of temperature and paleotemperature data sets that collectively span the Phanerozoic, with resolutions varying from about 200,000 years to (Fig. 1.5A, top) down to one year (Fig. 1.5B, bottom). Also shown (Fig. 1.5C) is what is probably the longest single atmospheric series—with 300,000 points—from dust accumulated in Antarctic ice cores. Figure 1.5C shows the series at eleven different resolutions, each differing by a factor of 2, clearly demonstrating the high variability at each octave in resolution. The first set (Fig. 1.5A) is benthic, derived from ocean sediments; the second set (Fig. 1.5B) is from ice cores, with the bottom one (at one-year resolution) being a "multiproxy" combination of ice cores, tree rings, lake sediments, pollen, and other paleoindicators. As explained in Box 1.2, the interpretation of the benthic series is somewhat indirect because

the signal (oxygen isotope ratios) depends on both the temperature of the water and the amount of water stored in ice sheets (which itself depends on the temperature, but also on geology, especially continental drift). Nevertheless, it is clear there is high variability at all observed scales, and the only hint of cyclic regularity is in the bottom series of Figure 1.5A, corresponding to an astronomical cycle (the obliquity, which goes through a cycle every 41,000 years; Box 5.4). The interpretation of the ice core isotope records in terms of past temperatures (Fig. 1.5B, top two series) is more straightforward, although note the apparent smoothing of the temperature variations in the more ancient parts of the EPICA series (top). This is an artifact resulting from molecular diffusion in the highly compacted bottom of the 3.2-km ice column from which the data were taken. As far as we can tell, the real temperature signal has high variability over the entire range. We return to examine quantitative statistical characterizations of the variability later in this chapter. Also discussed in Box 1.2 is a remarkable paleodata set spanning the past 800,000 years at near-annual resolution. Unfortunately, it is for the flux of dust, which is not a temperature proxy, but a more complex-to-interpret indicator of the climate state that—among other factors—is affected by humidity and the direction of the dominant winds. Nevertheless, it clearly shows the high variability at all observed scales.

### 1.2.2 INSTRUMENTAL TEMPERATURES

We finally come to real instrumentally based series. Figure 1.5D (top) shows 3,129 monthly values (from 1753–2013), but only over land.[46] Because the thermometer was only invented in 1638, even by 1753 (when this series starts), there were few recorded thermometer readings, and they were concentrated in Europe. This relative lack of data explains the amplitude of the temperature swings to the left of the graph that are much wider than those at the more modern right part of the series, for which many more measurements are available, giving a better estimate of the global-scale average. To minimize this effect, a running average at an annual resolution curve is also shown. This series covers the period since the Industrial Revolution, which is increasingly known as the "Anthropocene,"[r] the geological epoch strongly influenced by humans. Following the curve, we see the systematic rise since roughly the end of the nineteenth century[s] (we return to this in Chapter 6). Because the land has

---

[r] The term "Anthropocene" denotes the geological era that has been strongly marked by our species. The term includes and transcends the duration of anthropogenic climate change. Although it is now commonly in use, technically it is a geological term. In a million years, our ancestors will find the Anthropocene clearly delineated by a discontinuity in geological strata. The term is currently waiting for official approval by the International Union of Geological Sciences. Although humans started changing their environment significantly at least since the development of agriculture, by far the most important changes have been since the Industrial Revolution, so that—in this book—the two terms are used interchangeably.

[s] As far as the human impact on the atmosphere is concerned, the beginning of the Anthropocene can be dated with some precision. It is about the year 1750 (see Fig. 6.4A).

warmed more than the ocean, in this land-only series, the magnitude of the warming is a little greater than the global average.

The middle series in Figure 1.5D, starting in 1871, is from a localized region around Montreal, Quebec, at 2° spatial resolution (about 200 km), averaged over six hours and at a pressure level of 700 mbars (~70% of surface pressure, at about 3 km in altitude). Because there were no thermometer records from this altitude, the temperatures were inferred by using surface pressure data, monthly sea surface temperature (SST) data, and a meteorological model. This type of data–model hybrid is called a "reanalysis." This series is part of the Twentieth-Century Reanalysis (20CR).[47] For the purposes of this illustration, it has the advantage of being at a well-defined resolution in space and time, and it has no missing data.

In this single series (Fig. 1.5D, middle and bottom), there were more than 200,000 values, so we cannot display them usefully all at once; therefore, we average them and display 720 points only, so that the resolution displayed in the middle and bottom series of Figure 1.5D is about three months. We return to this momentarily. The raw data (Fig. 1.5D, middle series) are totally dominated by the annual cycle, which has an amplitude of about ±10°C. Because the annual cycle is very regular, to see the underlying natural variability, we remove it, leaving the "anomalies"[48]; the bottom series shows the result. We see that the amplitude of the three-month averaged anomalies is typically about ±2°C. Also shown for comparison is the estimated total change in the globally averaged temperature since 1880: 0.85°C.[t] We can see that, even after removing the large annual cycle, it is very hard to detect such a small change in a local temperature series. To determine the anthropogenic signal properly, we need to reduce the local variations (noise) by averaging over large areas—preferably over the entire globe.

Having removed the annual cycle to leave the anomalies, we can follow Figure 1.5C and make a series of blowups at regular intervals to be able to discern any changes in appearance systematically. Figure 1.5E shows the result of successive zooms by taking the left half of the series and blowing them up by a factor of about 2, retaining 720 points at each step, until we get to the six-hour resolution series (bottom left).[u] The bottom right series in Figure 1.5E is also from Montreal, but from a millimeter-size thermistor on the roof of the McGill University physics

---

[t] We also show for reference an estimate of the amplitude of the anthropogenic change for the global change since 1880. It is ~0.85°C [close to the IPCC's Assessment Report 4 (AR4) 2013 estimate; today, it is estimated to be 1°C; see Box 5.1 and Chapter 6]. See Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014). For the land only (top series), the estimate is about 1.5K. ~~See Moll, R., Cameron, R. H., & Schussler, M. Vortices in simulations of solar surface convection. *A&A* **533**, A126–A140 (2011). Moll, R., Cameron, R. H., & Schussler, M. Vortices, shocks and heating in the solar atmosphere: Effect of a magnetic field. *A&A* **541**, A68–A80 (2012). Twain, M. *English as She Is Taught: Genuine Answers to Examination Questions in Our Public Schools.* Collected by Caroline B. Le Row (Cassell & Co., Ltd. 1887). Lovejoy, S. & Lambert, F. High resolution EPICA ice core dust fluxes: Intermittency, extremes and Holocene stability. *Climate Past* (submitted, August 2018).~~ Rohde, R., Muller, R. A., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., Wurtele, J., Groom, D., & Wickham, C. A new estimate of the average earth surface land temperature spanning 1753 to 2011. *Geoinfor. Geostat. An Overview* **1** (1), 1–7 (2013).

[u] The diurnal cycle was removed.

building. The data were taken at 0.067-second resolution (i.e., fifteen times per second, at a resolution of a factor of 324,000 times higher). The temperature scale is the same for all the series except for the lower right. If higher resolution data were available, the variability would continue for at least another factor of 100 (as in Fig. 1.3, for the wind). To see the signature of the dissipation scale caused by molecular friction—"viscosity"—one would have to take data about 10,000 times a second (requiring extremely small, delicate devices!).

Starting at the lower right in Figure 1.5E, we see that—as for the EPICA series (Fig. 1.5A, top row)—the temperature appears to wander randomly, with temperature differences tending to grow with longer time intervals. This characteristic is still apparent at the next six-hour resolution (lower left)—at least for intervals as long as 10% to 20% of the series length (i.e., up to ten to twenty days). As we move upward to longer and longer resolutions to the series-indicated sixteen-day resolution, which is 34.5 years long, we may notice that the overall variation of the series doesn't change much (i.e., the rough range between the maximum and minimum is nearly independent of the resolution). Also notice that the "wandering" characteristic is gradually lost and that, by the time we reach a sixteen-day resolution, fluctuations are tending to cancel in a fairly systematic manner. A consequence is that as we move to even lower resolutions (the upper two series in Fig. 1.5E), the amplitude of the fluctuations starts to decrease systematically; we appear to be converging slowly to a well-defined climate state.

## 1.3  Zooming in space: From the size of the planet and from the top to the bottom
### 1.3.1  HORIZONTAL

Although Figure 1.1 covered visually pretty much the full range of spatial scales, it compared smoke (0.1 mm–1 m), visible cloud imagery (1 m–10 km), and infrared emissions (4 km–20,000 km). To allow for quantitative comparisons across spatial scales, Figure 1.6 (top) shows an aircraft transect (at about 12 km in altitude) of the temperature at 280-m resolution with 8,192 points, over nearly 3,000 km. The transect already displays variability over a range of scale of a factor of nearly 10,000, but its variability is much more extreme than appears at first sight. To bring out this surprising[v] but hidden feature, Figure 1.6 (bottom) also shows a "spike plot" of the absolute changes in temperature from one 280-m segment to the next (for more spike plots, see Fig. 5.2). The series of changes appear to be a hierarchy of spikes, with some being quite extreme; this is an example of turbulent "intermittency." To quantify this, we can normalize the changes by their "typical" value: one standard deviation.[49] When this is done, Figure 1.6 shows that the extreme temperature change is nearly seventeen times greater. If the changes were

---

[v] It isn't the spikes themselves that are surprising; rather, it is their occasional, unexpectedly huge amplitudes.

from the familiar "bell curve[w]" distribution of errors, then the probability of such an event would be one in $10^{86}$ (a one followed by eighty-six zeroes)—so infinitesimally small as to be all but impossible. These spikes, with their occasional extreme values, are thus so far outside the realm of the expected (here, the bell curve), that they are examples of black swan events (discussed in Box 3.1). Further zooms into aircraft trajectories (including wind and aircraft altitude) are discussed in Box 4.1.
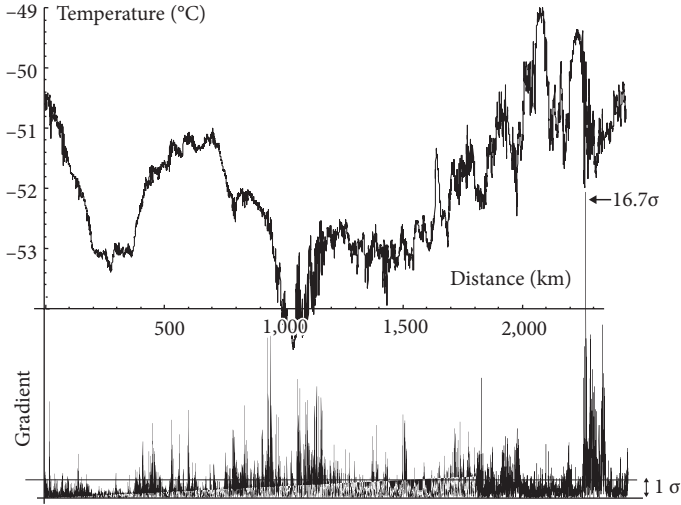


FIGURE 1.6  *(Top) The first 8,192 points of the temperature series measured by a Gulfstream 4 flight over the Pacific Ocean at 196 mb and at 1-second resolution, corresponding to a spatial resolution of 280 m. Because the aircraft speed is much greater than the wind, this can be considered a spatial transect. (Bottom) The change in temperature (gradient) from one measurement to the next normalized by dividing by the typical change (the standard deviation).[50]*

### 1.3.2  VERTICAL

We have seen evidence that the atmosphere displays extreme variability over wide ranges of scale in time and in the horizontal direction in space, but what about the vertical? It turns out that detailed (high-resolution) data in the vertical are very hard to obtain. For instance, the main source of weather data at different altitudes is from radiosondes, which are balloons that are still launched every twelve hours from several hundred locations globally. They measure wind, temperature, humidity, and pressure. But, they have low resolutions (they average over layers

---

[w] Gaussian distribution.

100–200 m thick) and have interpretation problems resulting, for example, from instruments swinging under the balloon.

Instead of lifting instruments upward with balloons, dramatically improved vertical measurements can be obtained by dropping them from aircraft—otherwise known as dropsondes. Dropsondes have small parachutes and fall at 10 m/s or so. Dropped from 10- to 12-km altitudes, and acquiring data every half a second, they each yield more than 1,000 measurements of the main atmospheric variables— wind, temperature, pressure, and humidity—corresponding to measurements taken roughly every 5 m in the vertical. The only drawback is that they cost several thousand dollars apiece . . . and they might hit people! In practice, they are dropped over oceans and are lost.

Nevertheless, experimental campaigns involving hundreds of dropsondes have revolutionary implications for our understanding of the atmosphere's vertical structure, which is described in more detail in Chapter 3. For the moment, let's focus our attention on the visual appearance of a few profiles. To bring out their unusual features clearly, I present them in the same way as the aircraft transect in Figure 1.6. I took the absolute differences. Figure 1.7 shows eight variables of fundamental importance to the atmosphere's dynamics: the horizontal wind speed ($v$), temperature ($T$), humidity ($q$), pressure ($p$), logarithm of the potential temperature ($\theta$), and of the equivalent potential temperature[x] ($\theta_E$), density of the air ($\rho$), and vertical wind ($w$). Only $v, T, q, p$, and $w$ are measured directly; the other quantities are derived from them using theoretical relationships. Figure 1.7A shows the overall profiles of the absolute changes presented in the usual manner, with the vertical coordinate in the vertical direction on the plot. The actual distance between the raw points where the data were taken varies somewhat because the sonde doesn't fall at a completely constant rate and, in addition, there were numerous data outages. For these reasons, two sondes were dropped 0.3 second apart.[y] Beyond the generally high level of intermittency ("spikiness"), one is immediately struck by a single enormous spike that occurs at near 600 m in altitude in the humidity ($q$) and (derived) equivalent potential temperature ($\theta_E$) series. Indeed, if it weren't for the fact that both sondes agree almost perfectly about the spike, it would be easy to dismiss this as a data glitch. A blowup (Fig. 1.7B) shows this feature quite clearly.

---

[x] These are thermodynamic functions depending on the entropy per unit mass for dry air and wet air, respectively.

[y] Corresponding to a separation of about 100 m.
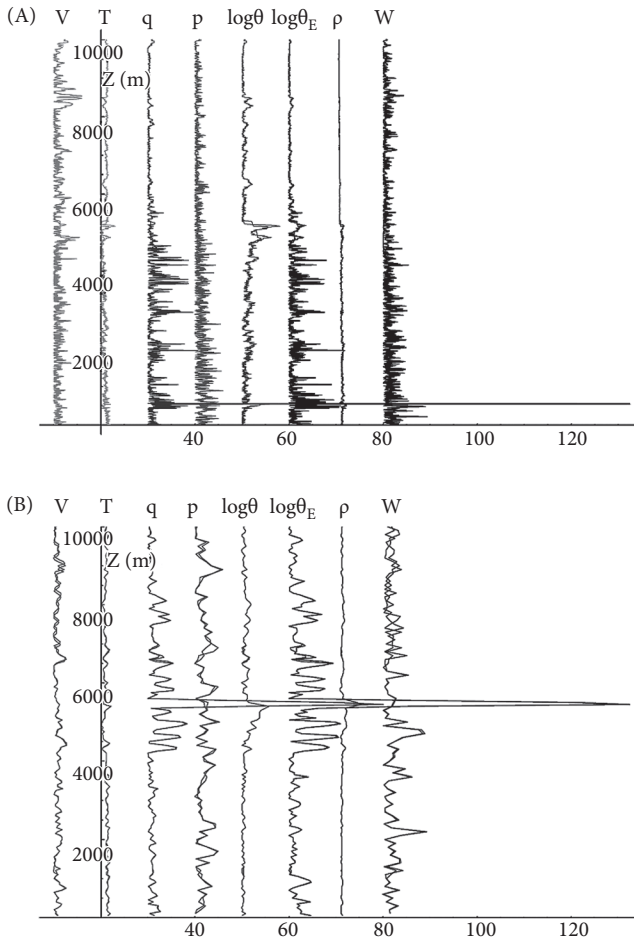
FIGURE 1.7 *(A) Profiles for the lower 10 km of the atmosphere from two sondes launched 0.3 second apart over the Pacific Ocean in the US National Oceanic and Atmospheric Administration's Pacific Storm 2004 experiment. The horizontal axis shows the absolute normalized changes of the measured and derived quantities; the vertical is the altitude (z). (B) A blowup of (A) by a factor of 10, from 100 m to 1 km.*[51]

Figures 1.6 and 1.7 show that the atmosphere is highly variable—indeed, "spiky"—intermittent over huge ranges of scale in both the horizontal and vertical directions. As a result of gravity, the atmosphere is stratified so that one does not expect that transects and profiles should look the same. Indeed, the visual appearances of temperature changes in Figures 1.6 and 1.7 are different. To assess the stratification directly, it is best to study vertical sections, which requires remote sensing. The data with the greatest range of scales covered—more than

1,000 in both the vertical and the horizontal—are from airborne lidars.[z] A lidar is a laser that sends short pulses of light (in Figure. 1.8, these are 3 m in length) and then—like a radar—it measures the distance of the reflection by the time it takes for the pulse to return.
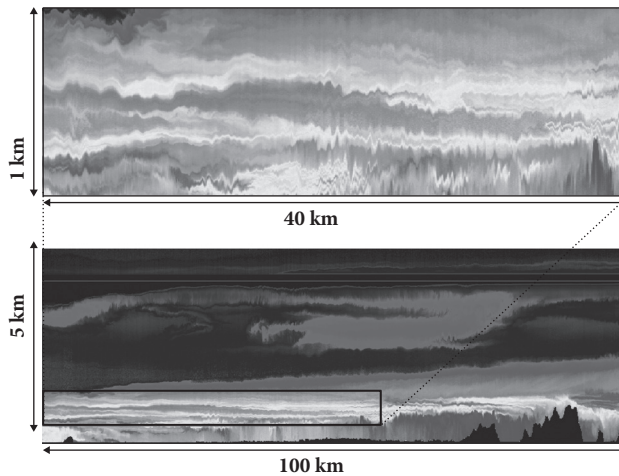


FIGURE 1.8 *(Bottom) A vertical section of laser backscatter from aerosols (smog particles) taken by an airborne lidar (laser) flying at a 4.5-km altitude (indicated as a horizontal line at that level) over British Columbia near Vancouver (the topography is shown in black; the lidar shoots two beams, one up and one down). The resolution is 3 m in the vertical and 96 m in the horizontal. This view is at a fairly coarse resolution, so we see mostly a layered structure. (Top) The black box at the lower left is blown up in the top of the figure, revealing the full resolution. We can now start to discern vertically aligned and roundish structures. The aspect ratio is about 40:1.[52]*

Figure 1.8 shows an~~other~~ example~~, this time~~ from an airborne lidar that detects aerosol particles (essentially from smog, in this case) and quantifies their back-scatter, which is related to their concentration in the air. In the low-resolution image (Fig. 1.8, bottom), we see mostly fairly stratified (flat) layers of smog, but when we zoom into the section in the box (top), we start to see small roundish or even vertically aligned structures, some of which are even wavelike. By comparing the shapes before and after the zoom, we see a systematic change in the morphologies, in this case from flat to roundish to vertically oriented. As quantitative analysis shows (see Chapter 3), this transformation involves no characteristic scale. It turns out that, to understand it, rather than imposing a preconceived

---

[z] Analogous to radar, lidar is an acronym for LIght Detection And Ranging.

notion of size or scale, we need instead to consider the notion of scale itself to be an emergent turbulent property. This is analogous to Einstein's theory of general relativity in which the distribution of energy and mass determines the notion of distance (the space–time "metric"). In the atmosphere, it means we must let the turbulent dynamics determine the appropriate notion of size.[53] A consequence is a revolutionary new theory of atmospheric dynamics: the 23/9 dimensional (D) model discussed in Chapter 3.

## 1.4  The unfinished nonlinear revolution: Junk your old equations and look for guidance in clouds . . .

### 1.4.1  THE DEVELOPMENT OF NUMERICAL MODELS

Although[aa] it can be traced back to the 1920s (or earlier, see Chapter 4), the modern quantitative era of weather science really got underway during the 1950s with the development of computers.[bb] Figure 1.9 shows the historical development of numerical models of the atmosphere as quantified by their size—the number of points they use in the horizontal and vertical directions. Today, typical models have more than 1,000 grid points in each horizontal direction,[54] and about one hundred in the vertical (Fig. 1.9).[cc] Over this time, computing power increased by a factor of a quadrillion, and the efficiency of numerical algorithms increased by a factor of perhaps another million.[55]

---

[aa] The second part of the title is a rhapsodic quote celebrating the nonlinear revolution. See Cvitanovic, P. Introduction. In: *Universality in Chaos* (ed. P. Cvitanovic), pp. 3–34; quote, p. 4. (Adam Hilger, 1984).

[bb] Indeed, in the early days (1948), mathematician John von Neumann (1903–1957) promoted the development of computers specifically to forecast the weather. From the very beginning, computer technology and weather forecasting have been intimately linked, with each stimulating the development of the other.

[cc] Because of atmospheric stratification, the models use systematically fewer points in the vertical. Indeed, as Figure 1.9 shows, the historical development of the ratio of vertical to horizontal grid points even follows the theoretical expectations of the 23/9 D model!

| | Date (approx.) | Name | Horizontal (points, rough) | Vertical (levels) | Comment |
|---|---|---|---|---|---|
| 1 | 1956 | Phillips | 16x17 | 2 | First simulation of long range general circulation |
| 2 | 1958 | Barotropic | 30x34 | 3 | Early Operational model |
| 3 | 1966 | Joint Numerical Weather Prediction Unit | 53x57 | 6 | First operational primitive equation model |
| 4 | 1974 | Global model | 144 | 9 | First global model |
| 5 | 1979 | ECMWF | 192 | 15 | First ECMWF model |
| 6 | 1980 | Global spectral model | 30 | 12 | First spectral model |
| 7 | 1991 | ECMWF | 650 | 31 | High res spectral model |
| 8 | 2006–present | Weather Research and Forecasting | 240 (typical high resolution) | 40 | Public domain |
| 9 | ≈2000 | HadCM3, | 96x73 | 19 | Met office coupled Atmosephere-Ocean model |
| 10 | 2010–present | Unified Model | 1600 | 70 | Met office operational model |
| 11 | current | Global Forecasting System | 1100 (high res mode) | 64 | NWS, operational model: 1988–present |
| 12 | current | Integrated Forecast System | 1600 | 91 | ECMWF operational model |
| 13 | current | Global Environmental Multiscale model | ≈2500 | 58 | Canadian Meteorological service operational model: 1997–present |
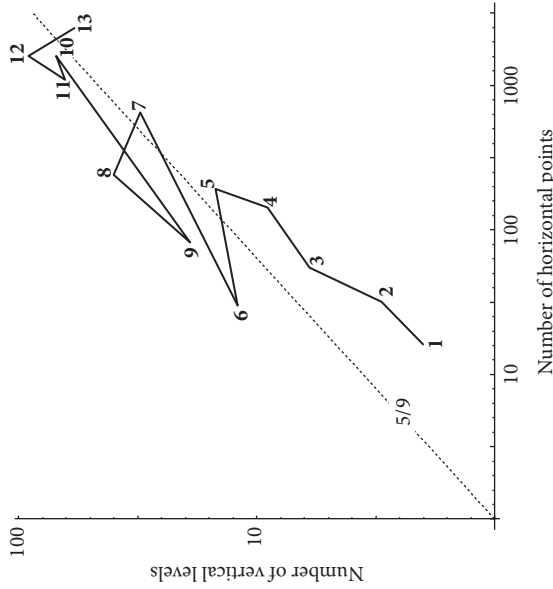


FIGURE 1.9  *The development of numerical weather models as quantified by their size—the number of points they use in calculation grids. The straight reference line has a slope of 5/9, which is required theoretically for an adequate treatment of the vertical stratification (the 2 + 5/9 = 23/9 dimensional model; see Chapter 3). Each model is labeled by a number with some details in the right-hand table.*

Simultaneous with the revolution in computation was a revolution in observations, especially from satellites and radar. Today, weather forecasting has skill up to several days in advance. It is starting to approach the theoretical limit to predictability—the time it takes for tiny initial errors to be amplified catastrophically, destroying the forecast.[dd]

The quantitative era of climate science began at around the same time, during the 1960s and '70s, with the extraction and analysis of the first long ice and ocean cores, with the first satellite imagery, and with the development of GCMs—which, when run far enough into the future (and when coupled with ocean models, and later cryosphere and carbon cycle models), can be averaged to characterize existing or future climates (see Chapters 5, 6, and 7). At first, the notion of climate change itself was poorly discerned, so that, notably, throughout the 1960s, there was widespread scientific skepticism of anthropogenic warming.[ee] However, the development of GCMs changed everything. One of the first large-scale numerical experiments was to investigate the effects of greenhouse gases.[56] By 1979, the results were sufficiently convincing that, in a report to the US National Academy of Science,[57] the effect of a $CO_2$ doubling was estimated as 1.5 to 4.5°C—a range that has remained unchanged ever since[58] (as discussed later). By the time the evidence for the warming was unequivocal (the 1980s), most climate scientists were already convinced of its reality.[ff]

The large-scale development of numerical weather and climate models represents an essentially brute-force approach to atmospheric dynamics analogous to the molecular dynamics approach to handling large numbers of molecules. At first, the models were so small they were not expected to be realistic; they were viewed primarily as tools for understanding the dynamics. However, over time, as they became bigger and more realistic, the modelers increasingly abdicated attempts at understanding, often becoming satisfied with representing the atmosphere's past and current state numerically or with forecasting its future development.

### 1.4.2  THE NONLINEAR BACKLASH

In parallel with the development of numerical models, theoretical attempts to understand the atmosphere as a complex dynamical system had also been developing. By the 1970s and '80s, these attempts led to a kind of backlash—a revolution in nonlinear science with new nonlinear paradigms developed solely for the

---

[dd] The deterministic limit, resulting from the butterfly effect (see Chapter 7).

[ee] At the time, the theory was seen as simplistic, and many found it repugnant to think that humans could change something as immutable as the climate. For a historical context, see Sherwood, S. Science controversies past and present. *Physics Today* **October**, 39–44 (2011). For a history, see Weart, S. R. *The Discovery of Global Warming*. (Harvard University Press, 2008).

[ff] There is an urban legend to the effect that, during the 1970s, the scientific community believed there was a danger of global cooling. This was never more than a fringe viewpoint, which was amplified by the media.

purpose of understanding turbulence, the atmosphere, and other strongly non-linear systems. A nonlinear process is one in which a small cause can be amplified by feedback to yield a potentially enormous effect. It is not hard to believe that the atmosphere is highly nonlinear and that this property is essential to understanding its variability. Indeed, a standard way of quantifying nonlinear systems is to take the ratio of the nonlinear part of the system to the linear part. In the atmosphere, this nonlinear-to-linear ratio is called the *Reynolds number*, after nineteenth-century fluid mechanist Osborne Reynolds (1842–1912). It is typically about a trillion, so that standard (linear) mathematical techniques are nearly useless.[59]

It was hardly an accident that nonlinear science was born thanks to a meteorologist: Ed Lorenz (1917–2008). In 1963, Lorenz published a seminal article with the innocuous-sounding title "Deterministic Nonperiodic Flow,"[60] in which he showed that a system with only three interacting components (his simple model of the atmosphere, the "flow") could follow randomlike ("nonperiodic") behavior, even though the rule governing its evolution was defined strictly with no room for chance (it was "deterministic"). The random appearance of his modeled flow was a symptom of the phenomenon of sensitive dependence on initial conditions in which tiny changes in the starting conditions are amplified and can lead to huge changes later on: the famous "butterfly effect" (see Chapter 7).

Until Lorenz's discovery of deterministic chaos,[gg] it was generally believed that typical real-world systems with few interacting components—such as the solar system, with the sun and each planet treated as a point—behaved in a calm, smooth, regular manner: Newton's "clockwork" universe. The discovery of deterministic chaos showed that, on the contrary, complex randomlike behavior was the norm, not the exception (Box 1.1).[hh]

While these ideas were generating excitement throughout the sciences, Benoit Mandelbrot (1924–2010) was investigating another aspect of the variability: fractals. The basic idea of fractals is scale invariance: that a small part of a system could in some way be similar to the whole. Indeed, by 1982, the fractal principle had already been applied to structures spanning the range from one to several thousand kilometers in scale[ii] and, by the mid 1980s, my colleague Daniel Schertzer and I had already proposed that scale invariance be used as the fundamental way of dividing atmospheric dynamics into qualitatively different regimes.[61] Both of these avenues are explored extensively in this book. The resulting high-level turbulent laws are illustrations of random "stochastic chaos"[62] (Box 1.1). What the intervening decades have taught us is that both deterministic chaos such as NWPs and GCMs

---

[gg] As discussed in Box 1.1, chaos is a much more general concept than this—a fact that is often forgotten!

[hh] Ironically, it has since been shown that, in reality—as a result of interactions among planets—the solar system itself is chaotic, although it isn't noticeable at periods less than tens of millions of years. See Laskar, J. Large-scale chaos in the solar system. *A&A* **287**, L9–L12 (1994).

[ii] See Section 2.8 in this book. See also Lovejoy, S. Area perimeter relations for rain and cloud areas. *Science* **187**, 1035–1037 (1982). The basic idea actually goes back to Richardson, L. F. Atmospheric diffusion shown on a distance-neighbour graph. *Proc. Roy. Soc.* **A110**, 709–737 (1926).

can coexist with stochastic chaos without contradiction (Fig. 1.3, bottom row), and that both can be used to describe, model, and forecast the atmosphere and climate.

So far, the scientific community has only sparsely assimilated the consequences of the nonlinear revolution. Although numerical models continue to grow in size and their skill for forecasting up to several days is uncontested, their brute-force approach has begun to reach a point of diminishing returns. On the one hand, during the past few years, NWP forecasts have shown only marginal improvements.[63] On the other hand, although climate models have developed rapidly (notably by increasing their spatial resolutions and the number of physical processes they attempt to represent), they too have been reaching a limit. A single number illustrates this perhaps better than anything else: the estimated sensitivity of GCMs to a doubling of $CO_2$ in the atmosphere. In 1979, based on early numerical experiments, the US National Academy of Sciences concluded that such doubling would lead to an increase in global temperature between 1.5°C and 4.5°C—a very large range.[jj] In its latest (fifth) assessment report (AR5, 2013), the IPCC found exactly the same range. Indeed, with the minor variant in the AR4,[64] every one of its reports iterated an identical range. The AR5 value was based primarily on thirty-two different GCMs, and the range reflected a fundamental limitation of the GCM approach: Each GCM has its own climate and each is different from the real world[kk] (see Chapters 5 and 7).

Although this numerical approach seems to be nearing its limit, the turbulent, stochastic chaos approach that has emerged from the past few decades is finally mature enough that it can make important contributions. It is needed to understand wide-range atmospheric variability properly, to reinstate the missing quadrillion, and to handle the new macroweather regime properly. It leads to improved applications in the form of better monthly, seasonal, annual, and decadal forecasts; and to longer term climate projections to 2050 and beyond (Chapter 7). The thesis of this book is that the nonlinear revolution is unfinished and, in the atmosphere—especially for weather and climate—it must be taken to its logical conclusion by recognizing and exploiting the high-level turbulent laws. By doing so, it promises to reunite the statistical and the deterministic strands of atmospheric science that have been divorced since the 1970s.[ll]

---

[jj] The range of 1.5 to 4.5°C for $CO_2$ doubling goes back to 1979 to the National Academy of Sciences Committee on Anthropogenic Global Warming. The story of how this wide range was arrived at was later recounted by a famous participant, Syukuro Manabe. The committee had the results of two models: one that gave a sensitivity of 2°C and the other, 4°C. According to Manabe's recollections, the committee chair, Jule Charney, "chose 0.5 °C as a not-unreasonable margin of error, subtracted it from Manabe's number, and added it to Hansen's. Thus was born the 1.5 °C-to-4.5 °C range of likely climate sensitivity that has appeared in every greenhouse assessment" (p. 933). See Kerr, R. A. Three degrees of consensus. *Science* **305** (5686), 932–934 (2004).

[kk] The scientific difficulty in evaluating the consequences of increasingly greenhouse gases—and the main source of disagreement between the GCMs—is how to estimate the amplifying effects of water vapor and clouds (see Box 6.1).

[ll] In March 2018, in preparation for the centennial of the American Geophysical Union (AGU), section leaders were asked to predict the major changes they expected in their field over the coming century. My answer:

### 1.4.3 COMPLEXITY SCIENCE

Before moving on, a word about complexity and complexity science. The atmosphere and climate system may be considered examples of "geocomplexity."[65] Complexity science emerged as one strand of the nonlinear revolution, as an attempt to generate a series of general complexity mechanisms, models, and paradigms that might apply widely, including to biosystems (where many complexity paradigms originated). An ambiguity arises because complexity itself may be somewhat subjective: Is the wind trace (Fig. 1.3, top) simple or complex?[66] Although it certainly looks *complicated*, both the stochastic and deterministic models at the bottom of Figure 1.3 are based on simple rules so that, if these models are realistic, then Figure 1.3 is only *apparently* complex.[67] Indeed, the history of science is replete with examples of how apparently complex systems turn out to be simple after they are theorized with the appropriate concepts. The turbulence approach to the atmosphere developed in these pages is indeed an attempt to make the apparently complex atmosphere as simple as possible.

But complexity science has more to it than just trying to find simple models for the apparently complex. It has also given rise to overarching theories of how our real, ordered cosmos actually evolved from the Big Bang through to the solar system to Earth, with its atmosphere and its biota, and even to the brain human. Eric Chiason, in his book *Cosmic Complexity*,[68] has gone further in this direction than most. Interestingly, he chose a precise complexity metric, the energy rate density: ε. ε is the typical power (energy rate) per unit of mass necessary to maintain a complex system in a quasi-steady state far from thermodynamic equilibrium. It turns out that ε is also the fundamental notion in classic turbulence theory, and we find that it that governs the horizontal weather and ocean dynamics up to planetary scales (Chapter 4).[69] Placing the atmosphere's average ε value into Chiason's scheme (Fig. 1.10) shows that it fits in reasonably well.

---

"I expect that in one hundred years, geophysical modeling (including the ones that are currently fluid based) will have fully developed stochastic approaches that are complementary to our current deterministic ones.

Rather than attempting to account deterministically for all the "details," the stochastic approaches will be based on the collective behavior of vast numbers of interacting components. They will exploit scale symmetries and, for fluid systems, they will be the descendants of classic approaches to turbulence.

This will be like molecular dynamics versus statistical mechanics: Both may be valid, and one uses one approach or the other depending on the application.

In particular, I expect that weather and climate predictions will be primarily stochastic."
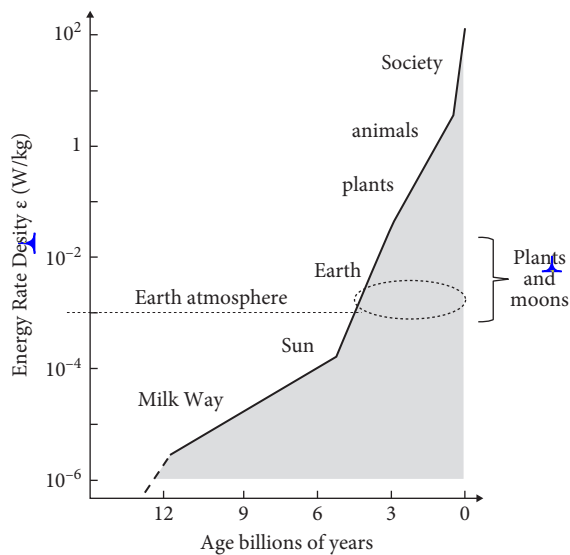
FIGURE 1.10 *The evolution of complex systems since the Big Bang as quantified by the energy rate density (ε, measured in watts per kilogram).[70]*

## 1.5  An overview of this book

This book is about scale in both space and in time, and how its systematic study is revolutionizing the way we model and understand the atmosphere—from small to big and fast to slow—all the way from weather to climate to megaclimate. It's about how we *think* about the atmosphere. It is an attempt to give you the fundamental information needed to understand what weather and climate really *are*, so you can fully understand the meaning of weather, of climate, of climate *change*. It focuses on developments during the past three decades, many of which I have been involved in personally, often under the rubric of nonlinear geophysics. Nonlinear geophysics is the geoscience application and development of ideas spawned by the nonlinear revolution. It can be traced back to a series of workshops on nonlinear variability in geophysics (NVAG), the first of which was held at McGill University in 1986,[71] and then on to the establishment of the nonlinear processes division at the European Geophysical Society[mm] (1989), the *Nonlinear Processes in Geophysics* journal (1994), and the nonlinear geophysics focus group at the AGU (1998). Nonlinear geoscience is highly interdisciplinary and is distinguished from other scientific specializations by its strong focus on the strongly nonlinear aspects of geosystems.

---

[mm] Since 2002, the European Geosciences Union.

A few years ago, Daniel Schertzer and I wrote a lengthy technical review titled *The Weather and Climate: Emergent Laws and Multifractal Cascades*.[72] Although the book you now hold in your hands is (hopefully) a widely accessible nontechnical distillation, it goes well beyond that review, extending its scope—especially with regard to macroweather and the climate. The basic text is intended to be readable without special technical knowledge. However, for those who wish to learn more, there are numerous "boxes" that can either be read in sequence as part of the narrative, skipped, or read independently on their own. There are also numerous footnotes and endnotes that give more information, with the latter being more technical.

The book is structured as follows: In Chapter 2, I discuss the two opposing ways of conceptualizing systems with variability over wide ranges: scalebound and scaling. In Chapter 3, we examine a third, intermediate view based on a generalization of the notion of scale to anisotropic systems, which is necessary for answering the question: How big is a cloud? In Chapter 4, we look at the long saga of vindicating Richardson's wide-range scaling view of the atmosphere, and consequences for the temporal scaling up to the end of the weather regime. In Chapter 5, I discuss macroweather, climate, and macro- and megaclimate. In Chapter 6, we see how humans have altered the atmosphere, by how much, and how we know what we know. In Chapter 7, I discuss predictability and its limits— both the (nearly familiar) deterministic butterfly effect as well as the lesser known stochastic predictability limits. I go on to describe how high-level turbulent laws can be used to make predictions at monthly to decadal scales, and how scaling laws can be used to make predictions at monthly to decadal scales, and can help project the state of the atmosphere to the year 2050 or 2100.

## Notes

1. This is essentially the same as the International Panel on Climate Change (IPCC) that states: "Climate in a narrow sense is usually defined as the 'average weather,' or more rigorously, as the statistical description in terms of the mean and variability of relevant quantities over a period of time ranging from months to thousands or millions of years." See IPCC. Glossary. In: The Fourth Assessment Report of the IPCC. (IPCC, 2007).

2. See Sachs, D., Lovejoy, S., & Schertzer, D. The multifractal scaling of cloud radiances from 1m to 1km. *Fractals* **10**, 253–265 (2002).

3. There exist somewhat artificial particle-level "lattice–gas cellular automaton" models that are used for the simulation of fluid flows, and these can reproduce vortices. See, for example, Wolf-Gladrow, D. A. *Lattice–Gas Cellular Automata and Lattice Boltzmann Models*. (Springer, 2000).

4. Wedemeyer-Böhm, S., Scullion, E., Steiner, O., Rouppe van der Voort, L., & de la Cruz Rodriguez, J. Magnetic tornadoes as energy channels into the solar corona. *Nature* **486**, 7404–7408 (2012).

5. Moll, R., Cameron, R. H., & Schussler, M. Vortices in simulations of solar surface convection. *Astron. Astrophysics* **533**, A126–A140 (2011).

6. Thanks to M. Wilcek.

7. To make this point very clear, we could note that deterministic theories and models can be thought of as simply limiting special cases of stochastic ones in which all the probabilities reduce to certainties.

8. See Weierstrass, K. Presented to the Prussian Academy of Sciences, 1872.

9. Laplace, P. S. *Théorie Analytique des Probabilités.* Vol. 7. (Gauthier-Villars, 1886). (originally published 1812).

10. Kolmogorov, A. N. *Grundebegrisse der Wahrscheinlichkeitrechnung.* (Springer, 1933). An English translation by N. Morrison appeared under the title *Foundations of the Theory of Probability* (Chelsea, New York) in 1950, with a second edition in 1956. For a history of the work, see Shafer, G., & Vovk, V. The sources of Kolmogorov's Grundbegriffe. *Stat. Sci.* **21** (1), 70–98 (2006).

11. The probabilistic nature of quantum mechanics is a consequence of the "Born rule" for interpreting the modulus squared of the wave function as proportional to the probability (after Max Born, circa 1926).

12. York, J. & Li, T. Y. Period three implies chaos. *Amer. Math. Monthly* **82**, 985–992 (1975).

13. Such a state is usually considered to be the strongly nonlinear limit of *incompressible* hydrodynamic turbulence so that the atmosphere would be a stratified and compressible variant. It has been suggested that fully developed turbulence is a new state of matter akin to the familiar solid, liquid, and gaseous states. See Manneville P. *Instabilities, Chaos and Turbulence*, p. 642. 2nd ed. (Imperial College Press, 2010).

14. Feigenbaum, M. J. Quantitative universality for a class of nonlinear transformations. *J. Stat. Phys.* **19**, 25 (1978).

15. The development of nonlinear theory had been hampered by the fact that each distinct nonlinear system had its own quirks, properties, and behavior, each requiring specific—often lengthy—analysis. *Universality* meant that quite different detailed models shared the same basic "universal" behavior; therefore, it was only necessary to study broad "universality classes." Because no model is a perfect fit to reality, without some kind of universality, nonlinear models could not be expected to be able to explain the real world or to be experimentally testable. Universality in itself was not a new idea. In probability theory, it explains the ubiquity of the *bell curve* (Gaussian probability distributions). The term *universality* gained currency in the 1960s and '70s in the study of phase transitions.

16. Cvitanovic, P. Introduction In: *Universality in Chaos* (ed. P. Cvitanovic), pp. 3–34; quote, p. 4. (Adam Hilger, 1984).

17. Nicolis, C. & Nicolis, G. Is there a climate attractor? *Nature* **311**, 529 (1984).

18. Osborne, A. R., & A. Provenzale. Finite correlation dimension for stochastic systems with power-law spectra. *Physica D* **35**, 357–381 (1989).

19. Lovejoy, S. & Schertzer, D. In: *ECO-TEC: Architecture of the In-between* (ed. A. Marras), pp. 80–99. (Princeton Architectural Press, 1999).

20. Mathematically, these probability spaces are usually infinite dimensional.

21. Fortunately, atmospheric scientists are rarely concerned with philosophical purity. During the 1990s, despite the ambient deterministic worldview, they improved weather forecasts by introducing random numbers at the model grid scale to represent the dynamics of structures that were too small to be modeled explicitly. This was the origin of today's hybrid deterministic–stochastic ensemble forecasting systems. I predict that this is

only the first step; ultimately, these models will be replaced by pure stochastic forecasting systems.

22. Illustration by I. Tchigirinskaya; reproduced from Schertzer, D., & S. Lovejoy.Lecture Notes: Resolution Dependence and Multifractals in Remote Sensing and Geographical Information Systems. McGill University, Montreal, June 10, 1996.

23. Bunge, M. *Causality*. (Dover, 1956).

24. The Kolmogorov law (see Chapter 4).

25. This is a typical value estimated with average energy fluxes (see Chapter 4).

26. Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

27. Twain, M. *English as She Is Taught: Genuine Answers to Examination Questions in Our Public Schools*, p. 29. Collected by Caroline B. Le Row. (Cassell & Co., Ltd., 1887).

28. Huschke, R. E. *Glossary of Meteorology*, p. 581. (American Meteorological Society, 1959).

29. Committee on Radiative Forcing Effects on Climate, National Research Council. *Radiative Forcing of Climate Change: Expanding the Concept and Addressing Uncertainties*, p. 12. (National Academies Press, 2005).

30. For a compact and nontechnical exposition, see Lovejoy, S. What is climate? *EOS* **94** (1), 1–2 (2013). The term *macroweather* replaces the slightly earlier cumbersome expression low-frequency weather [Lovejoy, S. & Schertzer, D. Low frequency weather and the emergence of the climate. In: *Extreme Events and Natural Hazards: The Complexity Perspective* (ed. A. S. Sharma, A. Bunde, D. N. Baker, & V. P. Dimri), pp. 231–254. AGU monographs. (AGU, 2012).] Macroweather is the subject of Chapter 10 in Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

31. In the industrial epoch. In the preindustrial epoch, the timescale is longer and still poorly understood (see Chapter 5).

32. Veizer, J., Ala, D., Azmy, K., Bruckschen, P., Bruhn, F., Buhl, D., Carden, G., Diener, A., Ebneth, S., Goddris, Y., Jasper, T., Korte, C., Pawellek, F., Podlaha, O., & Strauss, H. 87Sr/86Sr, δ $^{18}$O and δ $^{13}$C evolution of Phanerozoic seawater. *Chem. Geol.* **161**, 59–88 (1999).

33. Shaviv, N. J. & Veizer, J. Celestial driver of Phanerozoic climate? *GSA Today* **July**, 4–10 (2003).

34. Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686–693 (2001).

35. Huybers, P. Glacial variability over the last two million years: An extended depth-derived age model, continuous obliquity pacing, and the Pleistocene progression. *Q. Sci. Rev.* **26**, 37–55 (2007).

36. For this and the middle series in Figure 1.5A, a roughly 50% larger calibration constant of –6.5°C/δ$^{18}$O was used to take into account the greater high-latitude variations.

37. Using either deuterium or $^{18}$O gives very similar results.

38. Moberg, A., Sonnechkin, D. M., Holmgren, K., Datsenko, N. M., & Karlén, W. Highly variable northern hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature* **433**, 613–617 (2005).

39. Reproduced with permission from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dyn.* **44**, 3187–3210 (2015).

40. Reproduced with permission from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dyn.* **44**, 3187–3210 (2015).

41. The resolutions in years are thus not constant as a result of due to the compression of the ice with depth. They are given by the durations divided by 290: respectively, 0.24, 0.48, 0.98, 2.02, 4.32, 10.0, 24.5, 54.1, 185, 434, and 2,710 years per point (upper left to lower right).

42. The means of the first eight are about the same; the means of the last three are about seven times greater larger than those of the last three (these are the ones that include large fractions—or all—of glacial-–interglacial cycles).

43. Reproduced from Lovejoy, S., & Lambert, F. High resolution EPICA ice core dust fluxes: Intermittency, extremes and Holocene stability. *Climate Past* (submitted, August 2018).

44. Reproduced with permission from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

45. Reproduced with permission from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

46. Rohde, R., Muller, R. A., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., Wurtele, J., Groom, D., & Wickham, C.A new estimate of the average Earth surface land temperature spanning 1753 to 2011. *Geoinfo. Geostat.* **1**, 1–7 (2013).

47. Compo, G. P., et al. The twentieth century reanalysis project. *Q. J. Roy. Meteorol. Soc.* **137**, 1–28, (2011).

48. Here we remove the annual cycle based on the entire record since 1871. The more usual practice is to remove it over the most recent thirty-year climate-normal period.

49. The square root of the average of the square of the difference between the value and the mean.

50. In Chapter 5, we see similar spike plots but we normalize instead by the mean absolute change rather than by the standard deviation, which makes only a minor difference. Reproduced from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades* (Cambridge University Press, 2013).

51. Reproduced from Lovejoy, S., Tuck, A. F., Hovde, S. J., & Schertzer, D. The vertical cascade structure of the atmosphere and multifractal drop sonde outages. *J. Geophys. Res.* **114**, D07111 (2009), doi: 07110.01029/02008JD010651.

52. Reproduced from Lilley, M., Lovejoy, S., Strawbridge, K., & Schertzer, D. 23/9 Dimensional anisotropic scaling of passive admixtures using lidar aerosol data. *Phys. Rev. E* **70**, 036307-036301-036307 (2004).

53. With generalized scale invariance (discussed in Chapter 3), size is not a distance. In general, we define it without a metric.

54. Because Earth is spherical, usually they use special "Gaussian grids," so things are not quite as simple. Nevertheless, typically each of the hundred or so vertical layers of a global model has a million or so points.

55. Lynch, P. *The Emergence of Numerical Weather Prediction: Richardson's Dream*. (Cambridge University Press, 2006).

56. Manabe, S. & Wetherald, R. T. The effects of doubling the $CO_2$ concentration on the climate of a general circulation model. *J. Atmos. Sci.* **32**, 3 (1975).

57. Climate Research Board. *Carbon Dioxide and Climate: A Scientific Assessment.* (US National Academy of Science, 1979).

58. Identical to the IPCC AR3, 2002, range and the IPCC AR5, 2013, range. The latter qualifies this range as "high confidence." The AR4 (2007) claimed the slightly reduced range of 2 to 4.5°C.

59. This has not prevented linear techniques from being used extensively for attempts at understanding atmospheric stability and waves. For a criticism and a scaling alternative, see Pinel, J. & Lovejoy, S. Atmospheric waves as scaling, turbulent phenomena. *Atmos. Chem. Phys.* **14**, 3195–3210 (2014).

60. Lorenz, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).

61. Lovejoy, S. & Schertzer, D. Scale invariance in climatological temperatures and the local spectral plateau. *Ann. Geophys.* **4B**, 401–410 (1986).

62. Lovejoy, S. & Schertzer, D. Stochastic chaos and multifractal geophysics. In: *Chaos, Fractals and Models* (eds. F. M. Guindani & G. Salvadori), pp. 38–52. (Italian University Press, 1998). Lovejoy, S. & Schertzer, D. Stochastic chaos, scale invariance, multifractals and our turbulent atmosphere. In: *ECO-TEC: Architecture of the In-between* (ed. Amerigo Marras), pp. 80–99. (Storefront Book series, copublished with Princeton Architectural Press, 1999).

63. Stern, H. & Davidson, N. E. Trends in the skill of weather prediction at lead times of 1–14 days. *Q. J. Roy. Meteorol. Soc.* **141**, 2726–2736 (2015).

64. In 2007, the AR4 instead gave a range of 2 to 4.5°C.

65. The term was apparently first used as a title in a workshop in 2009. See Lovejoy, S., Agterberg, F., Carsteanu, A., Cheng, Q., Davidsen, J., Gaonac'h, H., Gupta, V., L'Heureux, I., Liu, W., Morris, S. W., Sharma, S., Shcherbakov, R., Tarquis, A., Turcotte, D., & Uritsky, V. Nonlinear geophysics: Why we need it. *EOS* **90** (48), 456–457 (2009).

66. Lovejoy, S., Schertzer, D., Allaire, V., Bourgeois, T., King, S., Pinel, J., & Stolle, J. Atmospheric complexity or scale by scale simplicity? *Geophys. Res. Lett.* **36**, L01801 (2009), doi: 10.1029/2008GL035863.

67. One can define complexity mathematically by the number of bits of information needed to specify a process in a computer algorithm. On this basis, the bottom models are both very simple.

68. Chiason, E. J. Energy rate density as a complexity metric and evolutionary driver. *Complexity* **16**, 27–40 (2010).

69. In turbulence, it $\varepsilon$ is viewed as the rate that energy flows from large structures to small ones, and it is usually given in the equivalent units of meters squared per cubic second. It also governs the oceans up to about one to two years, as well as the Martian atmosphere.

70. Adapted from Chiason, E. J. Energy rate density as a complexity metric and evolutionary driver. *Complexity* **16**, 27–40 (2010).

71. The second was in Paris in 1988; the third in Corsica, France, in 1994; and the fourth in Roscoff, France, 1998. See Schertzer, D. & Lovejoy, S. EGS Richardson AGU Chapman NVAG3 conference: Nonlinear VAriability in Geophysics: Scaling and multifractal processes. *Non. Proc. Geophysics* **1**, 77–79 (1994). See also, Schertzer, D. & Lovejoy, S. *Non-linear variability in geophysics: Scaling and Fractals.* (Kluwer, 1991).

72. Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades.* (Cambridge University Press, 2013).

{ 2 }

# New worlds versus scaling: From van Leeuwenhoek to Mandelbrot

## 2.1 Scalebound thinking and the missing quadrillion

### 2.1.1 NEW WORLDS AND SPECTRAL BUMPS

We just took a voyage through scales, noticing structures in cloud photographs and wiggles on graphs. Collectively, they spanned ranges of scale over factors of billions in space and billions of billions in time. We are immediately confronted with the question: How can we conceptualize and model such fantastic variation?

Two extreme approaches have developed. For the moment, I call the dominant one the *new worlds* view, after Antoni van Leeuwenhoek (1632–1723), who developed a powerful early microscope. The other is the self-similar (scaling) view by Benoit Mandelbrot (1924–2010), which I discuss in the next section. My own view—scaling but with the notion of scale itself an emergent property—is discussed in Chapter 3.

When van Leeuwenhoek peered through his microscope,[a] in his amazement he is said to have discovered a "new world in a drop of water": "animalcules," the first microorganisms[1] (Fig. 2.1). Since then, the idea that zooming reveals something completely new has become second nature. In the twenty-first century, atom-imaging microscopes are developed precisely because of the promise of such new worlds.

---

[a] The inventor of the first microscope is not known, but van Leuwenhoek's was powerful, with up to about three hundred times magnification.

FIGURE 2.1 *Antoni van Leuwenhoek discovering "animalcules" (microorganisms), circa 1675.*

The scale-by-scale "newness" idea was graphically illustrated by K. Boeke's highly influential book *Cosmic View*,[2] which starts with a photograph of a girl holding a cat, first zooming away to show the surrounding vast reaches of outer space, and then zooming in until reaching the nucleus of an atom. The book was incredibly successful. It was included in Hutchins and Adler's *Gateway to the Great Books*,[3] a ten-volume series featuring works by Aristotle, Shakespeare, Einstein, and others. In 1968, two films were based on Boeke's book—*Cosmic Zoom*[4] and *Powers of Ten* (1968[5], re-released in 1977[6]), encouraging the idea that nearly every power of ten in scale hosts different phenomena. More recently (2012), there's even the interactive Cosmic Eye app for the iPad, iPhone, or iPod, not to mention a lavish update: the "Zoomable Universe."[7]

In an article published in 1981, Mandelbrot coined the term *scalebound* for this "new worlds" view—a convenient shorthand[b] that I use frequently later.[c] Often, scaleboundedness is obvious. For example, adult humans have heights in the range of 57 cm to 270 cm (these are records!)—a factor of 5—and give or take a factor of 10, atoms are several tenths of a nanometer across. Stretching the

---

[b] He wrote it as used here, as one word, as a single concept.

[c] He was writing in *Leonardo*, to an audience of architects: "I propose the term scalebound to denote any object, whether in nature or one made by an engineer or an artist, for which characteristic elements of scale, such as length and width, are few in number and each with a clearly distinct size" (p. 44). Mandelbrot, B. Scalebound or scaling shapes: A useful distinction in the visual arts and in the natural sciences. *Leonardo* **14**, 43–47 (1981).

idea only a little, bacteria are scalebound even though they range in size from a tenth of a micron to a tenth of a millimeter. But what about a cloud, a coastline, or a storm?

While *Powers of Ten* was proselytizing the new worlds view to an entire generation, there were other developments that pushed scientific thinking in the same direction. During the 1960s, long ice and ocean cores were revolutionizing climate science by supplying the first quantitative data at centennial, millennial, and longer timescales. This coincided with the development of practical techniques to decompose a signal into oscillating components: "spectral analysis." Although it had been known since Joseph Fourier (1768–1830) that any time series may be written as a sum of sinusoids, applying this idea to real data was computationally challenging and, in atmospheric science, had been largely confined to the study of turbulence.[d] The breakthrough was the development of fast computers combined with the discovery of the "fast Fourier transform" (FFT)[8] algorithm[e] (1965).

The beauty of Fourier decomposition is that each sinusoid has an exact, unambiguous timescale. Its period (the inverse of its frequency) is the length of time it takes to make a full oscillation (Fig. 2.2A, upper left, for example). Fourier analysis thus provides a systematic way of quantifying the contribution of each timescale (inverse frequency) to a time series. The spectrum is obtained from the analysis by averaging[9] the square of the contribution at a given frequency per unit frequency band. Qualitatively, it indicates the variability of the process as a function of timescale (one divided by the frequency). In the following, this qualitative idea is the main thing to keep in mind.[f]

---

[d] Even there, spectra were often estimated by using specialized circuitry involving numerous narrow-band filters.

[e] The speed-up resulting from the invention of the FFT is huge. Even for the relatively short series in Figure 1.3 (2,048 points), it is about a factor of 100. In numerical weather models, it accelerates calculations by factors of millions.

[f] A precise interpretation of the spectrum is technically challenging and, as discussed later, even the professionals missed key implications!
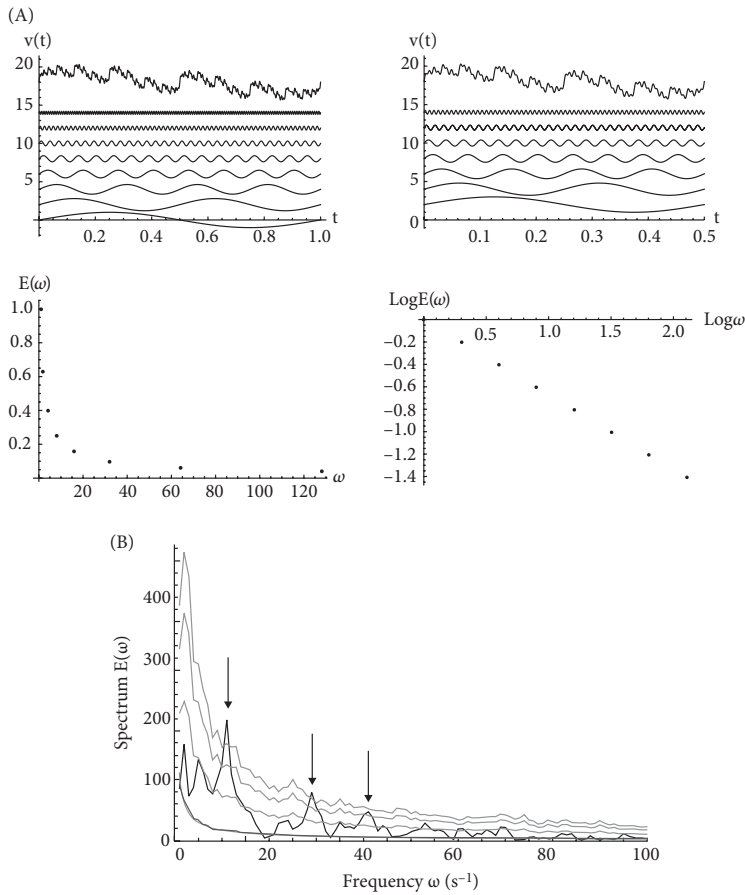
FIGURE 2.2  *(A) (Upper left) The first eight contributions to the Weierstrass function (displaced in the vertical for clarity). Sinusoids with frequencies of 1, 2, 4, 8, 16, 32, 64, and 128 cycles per second (the time t is in seconds). (Upper right) Sinusoids with frequencies of 2, 4, 8, 16, 32, 64, and 128 cycles per second, stretched by a factor of $2^H = 1.25$ in the vertical and a factor of 2 in the horizontal. The sum (top) is the same as that on the left, but it is missing the highest frequency detail (the difference is barely perceptible; see the discussion a little later). (Lower left) The spectrum on a linear–linear scale, with each point indicating the contribution (squared) and the frequency. (Lower right) The same as lower left, but on a logarithmic plot (it is now linear). (B) The darker line represents the Fourier spectrum[10] [E(ω), proportional to the amplitude squared of the components discussed in (A)] of the changes in wind speed in the one-second-long simulation shown at the bottom left of Figure 1.3, representing the amplitudes of the first one hundred frequencies (ω).[11] The points the farthest to the left are thus at frequencies of one cycle over the length of the simulation (i.e., one cycle per second, a period of one second). The far right shows the variability at one hundred cycles per second, giving the amplitude of the wiggles at 10 ms (higher frequencies are not shown for clarity). At the bottom, there are actually two gray curves that are nearly indistinguishable. One shows the average over 5,000 random series, each identical statistically to that in the bottom left of Figure 1.3. As expected, it is nearly identical to the other superposed theoretical (scaling) power law gray curve. The top three dotted curves show the theoretical 1%, 1 in 1 million and 1 in 10 billion extreme fluctuation limits (bottom to top) determined by assuming that the spectrum has bell curve (Gaussian) probabilities. The arrows show the most extreme spikes, each of which has a probability of less than 1 in 1 million.[12]*

Figure 2.2A illustrates this for the Weierstrass function, which in this example is constructed by summing sinusoids with frequencies increasing by factors of 2 so that the $n$th frequency is $\omega = 2^n$. The amplitudes ($A$) decrease by factors of $2^{-H}$ (here, $H = 1/3$ so that $2^{-H} = 0.79$), so that the $n$th amplitude is $2^{-nH}$. Figure 2.2A (upper left) shows the result for $H = 1/3$, with all the terms up until 128 cycles per second (upper row). Eliminating $n$, we find the power law relation $A = \omega^{-H}$. More generally for a scaling process, we have

$$\text{Spectrum} = (\text{Frequency})^{-\beta},$$

where $\beta$ is the usual notation for the "spectral exponent."[13] The spectrum is the square of the amplitude, so that in this (discrete) example,[14] we have $\beta = 2H$. The spectrum of the Weierstrass function is shown in Figure 2.2A (bottom row left) as a discrete series of dots, one for each of the eight sinusoids in the upper left construction. In the bottom row right, we see the same spectrum, but on a logarithmic plot on which power laws are straight lines. Of course, in the real world—unlike this academic example—there is nothing special about powers of two, so that all frequencies—a continuum—are present.

The Weierstrass function was created by adding sinusoids: Fourier *composition*. Now let's take a complicated-looking series—for example, the multifractal simulation of the data series (lower left in Fig. 1.3). It has small, medium, and large wiggles. To analyze it, we need the inverse of composition, and this is where the FFT is handy. In this case, by construction, we know that the wiggles are unimportant[g]; they are generated randomly by the process. However, if we had no knowledge—or only a speculation—about the mechanism that produced it, we would wonder: Do the wiggles hide signatures of important processes of interest? Or are they simply uninteresting details that should be averaging out and ignored?

Figure 2.2B shows the spectrum of the multifractal simulation (Fig. 1.3, lower left) for all periods longer than 10 ms. How do we interpret the plot? One sees three strong spikes, at frequencies of 12, 28, and 41 cycles per second (corresponding to periods of 1/12, 1/28, and 1/41 of a second—about 83, 35, and 24 ms). Are they signals of some important fundamental process or are they just noise?

Naturally, this question can only be answered if we have a mental model of how the process might be generated, and this is where it gets interesting. First of all, consider the case in which we have only a single series. If we know the signal is turbulent (as it is for the data series at the top), then turbulence theory tells us we could expect all the frequencies in a wide continuum of scales to be important and, furthermore, that at least on average, their amplitudes should decay in a power law manner (as with the Weierstrass function). But, classic theory

---

[g] I mean they don't imply any special origin or mechanism. However, various applications might only be sensitive to a narrow range of frequencies—for example, wind blowing against a swing. In this case, the wiggles, fluctuations in the wind that happens to occur at the natural frequency of oscillation of the swing would be important for pushing the swing even if—from the point of view of the underlying turbulent wind—they had no special role.

tells us only the spectrum we would expect to find if we averaged over a large number of identical experiments[h] (each one with different "bumps" and wiggles, but from the same overall conditions). In Figure 2.2B, this average is the smooth blue curve.[i]

But in Figure 2.2B, we see there are apparently large departures from this average. Are these departures really exceptional or are they just "normal" variations expected from randomly chosen pieces of turbulence? Before the development of cascade models and the discovery of multifractals in the 1970s and '80s, turbulence theory would have led us to expect that the up-and-down variations about a smooth line through the spectrum should roughly follow the bell curve. If this was the case, then the spectrum should not exceed the bottom curve more than 1% of the time and the top curve more than 1 in 10 billion times. Yet, we see that even this 1/10,000,000,000 curve is exceeded twice in this single but relatively unexceptional simulation[15] (indicated by the two leftmost arrows in Fig. 2.2B).

The spectrum turns out to be very sensitive to "jumps" and "spikes" that are hiding in the signal, as illustrated in Figure 1.6 (but see also Plates 4.1 and 4.2, and Figs. 4.5 and 5.2). This turns out to be an example of extreme black swan variability discussed in Box 3.1. Had we encountered this series in an experiment, turbulence theory itself would probably have been questioned—as indeed it repeatedly was (and still is). Failure to appreciate fully the huge variability that is *expected* in turbulent processes, and the continued embrace of inappropriate bell curve–type paradigms has spuriously shed discredit on many attempts at establishing turbulent laws and has been a major obstacle to their understanding.

In conclusion, until the development of multifractals in the 1980s, even if we knew the series came from an apparently routine turbulent wind trace on the roof of the physics building, we would still have concluded that the bumps were indeed significant.

But what would be our interpretation if, instead, Figure 2.2B was the spectrum of a climate series?[j] We would have no good theory of the variability and we would typically only have a single trace.

Let's take the example of an ice core record. The series itself was likely the product of a near-heroic scientific effort, possibly involving months in freezing conditions near the South Pole. The sample would first be cored and then transported to the lab. This would have been followed up by a painstaking sampling, analysis of the isotopic composition using a mass spectrometer, and finally a digitization of the result. Careful comparison with other cores or with ice flow models would eventually establish a chronology.

At this point, the researcher would be eager for a quantitative look at what she had found. If the darker curve in Figure 2.2B was the spectrum of such a core,

---

[h] An "ensemble" or "statistical" average (see Chapter 4).

[i] It is the ensemble average spectrum. In this example, it lies mostly below the spectrum of the single (somewhat exceptional) realization shown by the darker line.

[j] Obviously at totally different timescales!

how would she react to the bumps in the spectrum? Unlike the turbulence situation, in which there was some theory, an early core would have had little with which to compare. This is the point where the new worlds view could easily influence the researcher's results. She would be greatly tempted to conclude that the spikes were so strong, so far from the bell curve theory, that they represented real physical oscillations occurring over a narrow range of timescales. Alternatively, there might be incorrect theories that could be spuriously supported by fortuitously placed random spectral bumps, and much time would be wasted chasing blind alleys.

Further support for hypothesizing the action of special physical processes would come from the fact that the two extreme bumps in the spectrum involve several successive frequencies and, according to usual statistical assumptions, "background noise" should not be correlated in this way. This wide bump would strengthen the spurious interpretation that there was a hidden oscillatory process at work.[16] Armed with the series of bumps, she might start to speculate about possible physical mechanisms to explain them.[17]

We should thus not be surprised to learn that the 1970s witnessed a rash of articles based on spectra resembling that seen in Figure 2.2B; oscillators were suddenly ubiquitous.[18] It was in this context that Murray Mitchell[19] (1928–1990) famously made the first explicit attempt to conceptualize temporal atmospheric variability (Fig 2.3A). Mitchell's ambitious composite spectrum ranged from hours to the age of Earth ($\approx 4.5 \times 10^9$ to $10^{-4}$ *years*; bottom, Fig. 2.3A). Despite his candid admission[k] that this was mostly an "educated guess," and notwithstanding the subsequent revolution in climate and paleoclimate data, more than forty years later it has achieved an iconic status and is still cited and reproduced regularly in climate articles and textbooks.[20] Its continuing influence is demonstrated by the slightly updated version shown in Figure 2.3B, which (until 2015) adorned the National Oceanic and Atmospheric Administration's (NOAA's) National Climate Data Center paleoclimate website.[l] The site was surprisingly forthright about the figure's ideological character. While admitting that "in some respects it overgeneralizes and over-simplifies climate processes," it continued: "the figure is intended as a mental model to provide a general 'powers of ten' overview of climate variability, and to convey the basic complexities of climate dynamics for a general science savvy audience"[m] Notice the explicit reference to the "powers of ten" mindset more than fifty years after Boeke's book was published.

---

[k] I should make it clear that the problem was not Mitchell—who made an important pioneering contribution—but rather the later elevation of Mitchell's provisional, tentative interpretation into an unexamined ideology.

[l] The site explicitly acknowledged Mitchell's influence.

[m] If this were not enough, the site adds a further gratuitous interpretation. To assure skeptics, it continues "[just] because a particular phenomenon is called an oscillation, it does not necessarily mean there is a particular oscillator causing the pattern. Some prefer to refer to such processes as variability" (p. 3190). Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam*. **44**, 3187–3210 (2015). Because any time series—whether produced by turbulence, the stock market, or a pendulum—can be decomposed into sinusoids, the decomposition has no physical content per se, yet we are told that variability and oscillations are synonymous.
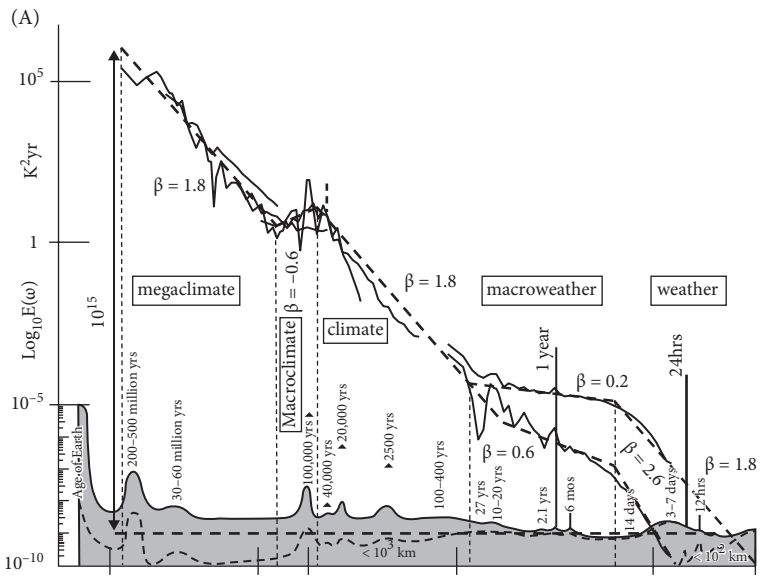
FIGURE 2.3  *(A) A comparison of Mitchell's relative scale, "educated guess" of the spectrum (gray, bottom),[21] with modern evidence from spectra of a selection of the series displayed in Figure 1.5 (the plot is logarithmic in both axes). There are three sets of thick black lines. On the far right, the spectra from the 1871 to 2008 Twentieth-Century Reanalysis (20CR; at daily resolution) quantifies the difference between the globally averaged temperature (bottom right) and local averages (2° × 2°, top right).[22] The spectra were averaged over frequency intervals (ten per factor of 10 in frequency), thus "smearing out" the daily and annual spectral "spikes." These spikes have been reintroduced without this averaging, and are indicated by vertical lines at daily and annual resolution curves. Using the daily resolution data, the annual cycle is a factor of ≈1,000 above the continuum, whereas using hourly resolution data, the daily spike is a factor of ≈3,000 above the background. Also shown is the other striking narrow spectral spike at 41,000 years[−1] (obliquity; approximately a factor of 10 above the continuum, shown as a short vertical dashed line), which is shown in a dashed gray line because it is only apparent over the period 0.8 to 2.56 million years before present. The thick sloping dashed lines have slopes indicating the scaling behaviors. The thin dashed lines show the transition periods that separate out the regimes discussed in detail in Chapters 4 and 5. These are at 20 days, 50 years, 80,000 years, and 500,000 years. Mitchell's original figure has been reproduced faithfully many times. It is not actually very important to be able to read the lettering near the spikes. If needed, most of them can seen in (B), which was inspired by the original [see also (C) and Box 2.1].[23] (B) The updated version of Mitchell's spectrum is reproduced from the National Oceanic and Atmospheric Administration's National Climate Data Center paleoclimate website.[n] The "background" on this spectrum is perfectly flat; hence, in comparison with the empirical spectrum in (A), it is in error by an overall factor of ≈10[16].[24] (C) The spectrum in (B), replotted in terms of fluctuations (gray), where the flat baseline of (B) now has a slope of –1/2, corresponding to an uncorrelated Gaussian "white noise" background. Because the amplitudes in (B) were not specified, the amplitude of the transformed gray "bumps" is only notional.[25] At the top is superposed the typical fluctuation at timescale Δt, as estimated from various instrumental and paleoseries, from those displayed in Figure 1.5. More details can be found in Section 2.3 and Figure 2.4A.[26]*

---

[n] The page has since been taken down.
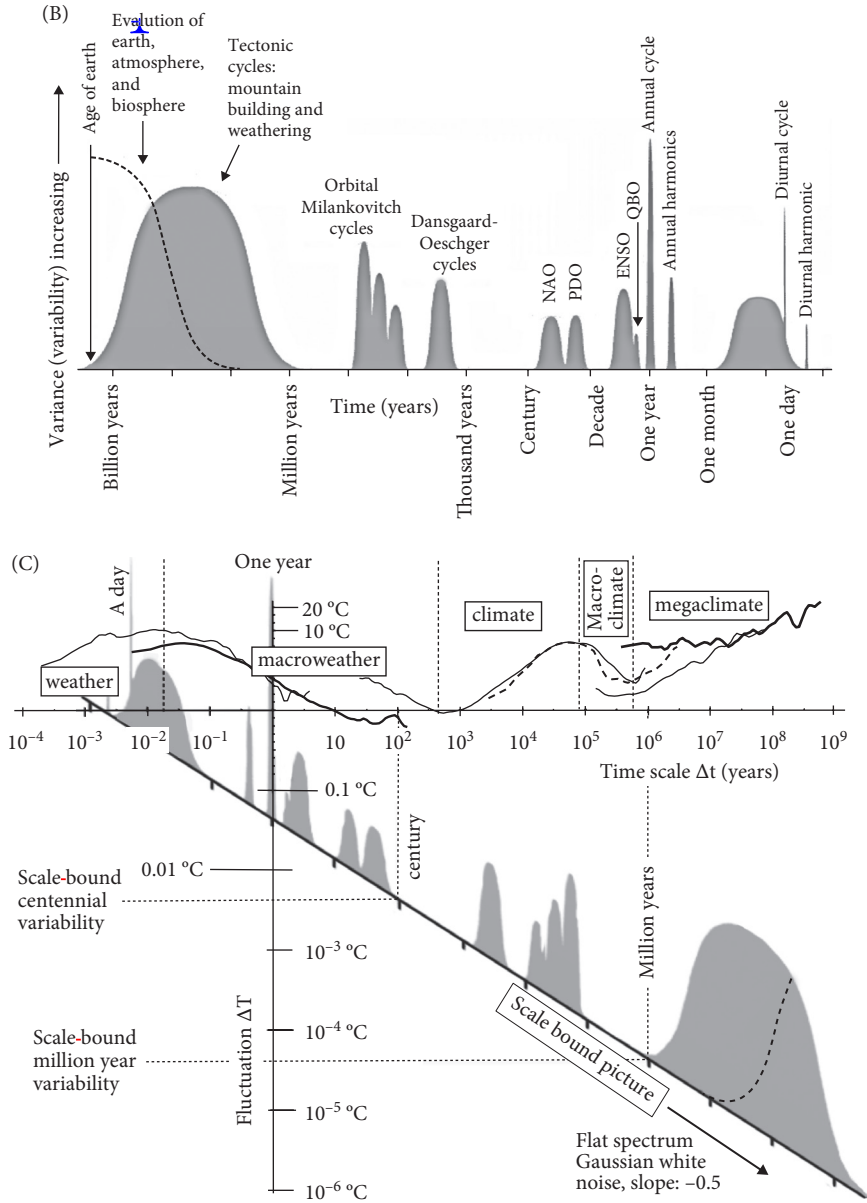
FIGURE 2.3 *Continued*

Certainly the continuing influence of Mitchell's figure has nothing to do with its accuracy. Within fifteen years of its publication, two scaling composites (one close to that shown in Fig. 2.3A[o] and the fluctuation analysis shown in Fig. 2.18)[27] already showed colossal discrepancies. In Figure 2.3A, the spectra of several of the series analyzed in Chapter 1 have been superposed; the difference with Mitchell's original is literally astronomical. Although over the range of one hour to 1,000,000,000 years, Mitchell's background varies by a factor of ≈150 (bottom, gray), the spectra from real data imply instead that the true range is a factor of a quadrillion[28] ($10^{15}$). With its completely flat background, NOAA (Fig. 2.3B) extends this error by a further factor of 10.[29]

Rather than plotting the data in the difficult-to-interpret terms of amplitudes of sinusoids, we can plot it in a much simpler-to-interpret way using typical amplitudes of fluctuations. Fluctuations are discussed in detail in Section 2.3, but even with only an intuitive idea that they quantify the variation of a quantity over an interval of time $\Delta t$, if they are scaling, they will follow a power law with fluctuation exponent $H$:

$$\text{Fluctuations} \approx \ (\text{Interval}\,\Delta t)^{H} \approx (\text{Scale})^{H}.$$

Figure 2.3C (top) shows the analysis of many of the same series in Figure 2.3A, but using the fluctuations described in detail in Section 2.3 (see Fig. 2.4A, B). We can see that Figure 2.3C makes intuitive sense. By reading the numbers off the top curve of the graph, we see that typical temperature fluctuations vary between a little less than a degree to as much as 15°C or 20°C at hundreds of millions of years (full details and a discussion are included in Section 2.3). In comparison, Figure 2.3C (bottom, the line and the gray areas) shows that the quantitative implications of Mitchell's spectrum are quite implausible. Although the amplitude of the fluctuations in Figure 2.3B was not specified, the baseline background in Figure 2.3B is a constant level of white noise that corresponds to fluctuations decreasing with a slope of –1/2, as indicated. We see, for example, that it implies that successive century-long average temperatures would typically differ by less than one hundredth of a degree and that consecutive 1,000,000-year average temperatures would differ by millionths of a degree centigrade! Figure 2.3A and C (top) shows how these analyses can be used to establish a basic typography to categorize the dynamical regimes. We return to this in Section 2.3 and Chapter 5. For a critical discussion, see Section 2.3 and especially Box 2.1.

---

[o] A composite spectrum over the range of scales from 1,000 to 100,000,000 years was published by Shackleton, N. J. & Imbrie, J. The δ^{18}O spectrum of oceanic deep water over a five-decade band. *Climat. Change* **16**, 217–230 (1990).
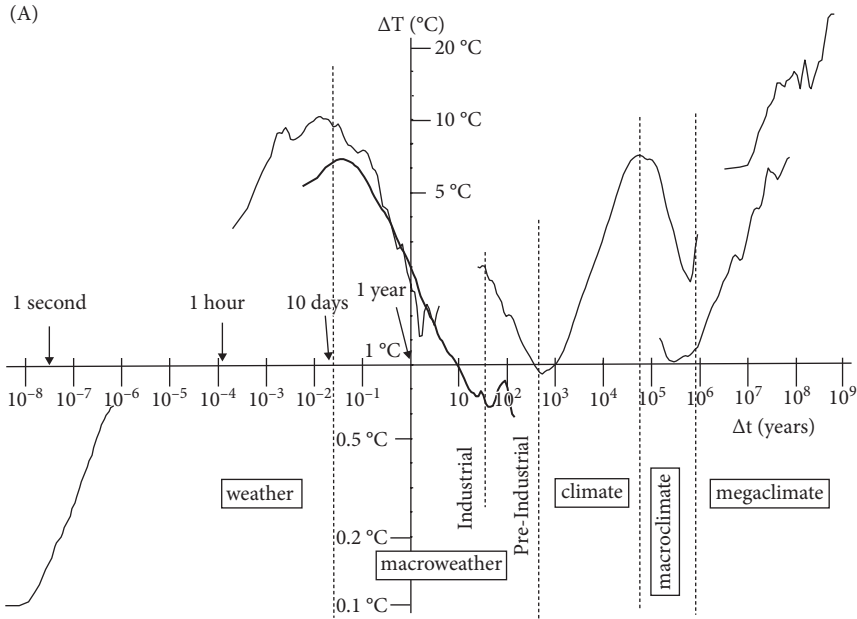
FIGURE 2.4  *(A) A composite of typical Haar fluctuations.[30] This composite is mathematically equivalent,[31] but is easier to understand than the spectrum in Figure 2.3A. Using largely the same data, it extends the range of timescales by a factor of 100,000 from an hour down to 2 × 0.017 = 0.034 second. From left to right, the curves are from thermistors at 0.017-second resolution (the same data as the lower right of Fig. 1.5E) from (daily and annually detrended) hourly temperatures (second from the left, from a station in Lander, Wyoming), Twentieth-Century Reanalysis temperatures (thick, middle; the same data as Fig. 1.5D, but at 75°N), and paleotemperatures from EPICA ice cores (the S-shaped curve at the right; from the data shown in Fig. 1.5B and in the spectrum in Fig. 2.3A) from the past 800,000 years. The two far right curves are from the upper benthic curves in Figure 1.5A and Figure 2.3A. The different dynamical regimes are indicated by the dashed lines, which roughly separate regions with different linear scale dependencies. The slopes are estimates of H.[32] (B) The same as (A), but shows spatial fluctuations. The curves display analyses from data averaged over weather (W), macroweather (M), and climate (C) timescales. The straight reference line has the slope H = 0.4. The lower left is from the aircraft data used in Figure 1.6 (280-m resolution), the upper right is from daily temperatures in January 2006, and the fluctuations in the longitudinal direction are every 1° in longitude (the same data as used in Fig. 5.2; see the description there). The middle (M) and lower (C) right curves are from monthly and 140-year averaged data in the longitudinal direction (see Fig. 5.2 for more details).[33] (C) Representative series from each of the five scaling regimes taken from Figure 1.5 and (A), with the addition of the hourly surface temperatures from Lander, Wyoming (bottom, detrended daily and annually), and a thermistor series in Montreal. To contrast their appearances fairly, each series had the same number of points (180) and was normalized by its overall range (the maximum minus the minimum), and each series was offset by 1°C in the vertical for clarity.[34] The series resolutions were 1 hour, 1 month, 400 years, 14,000 years, 370,000 years, and 1,230,000 years, bottom to top respectively. The fluctuation exponent H is indicated at the left, H < 0.[35]*

(B)

(C)
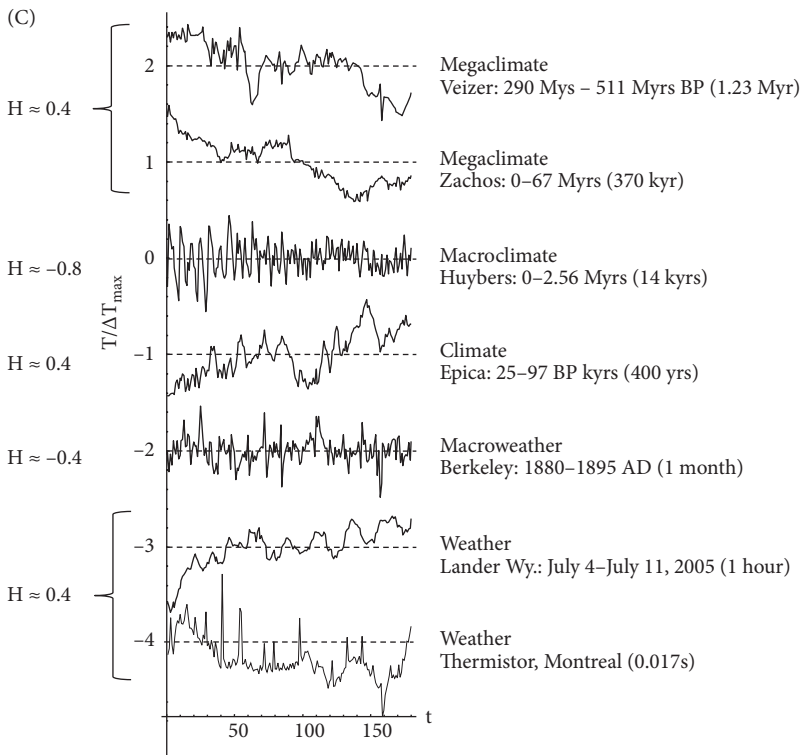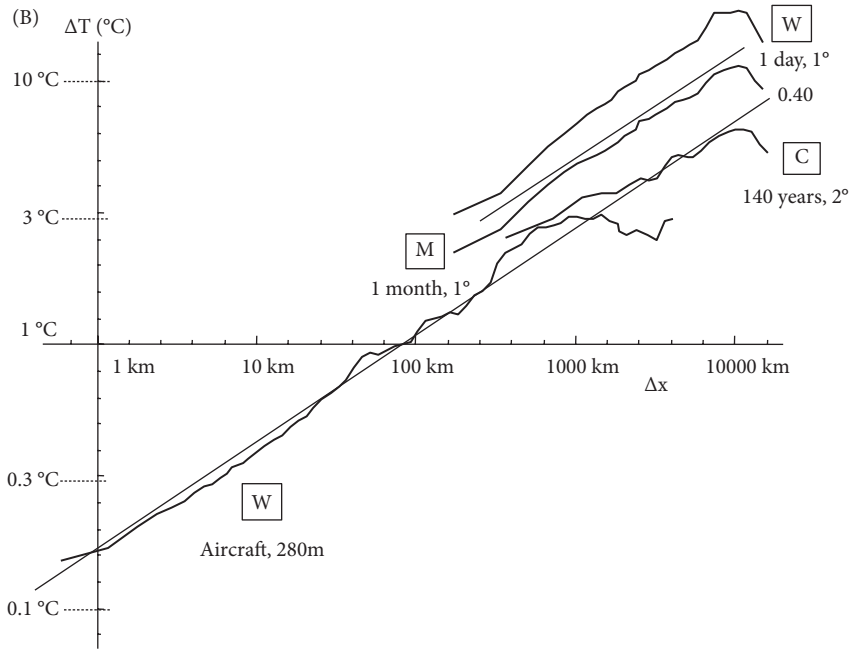
FIGURE 2.4 *Continued*

BOX 2.1  Lucy's fluctuations

The fluctuation analyses (Fig. 2.4A) and the (statistically equivalent) corresponding spectra (Fig. 2.3A) are based on real data, with the behavior of fluctuations being particularly straightforward to interpret. It points to a seductively simple picture involving four or—with macroclimate—possibly five scaling regimes.

But is this classification into different dynamical regimes really justified? So far, the main criticism of the composites is not that they are wrong, but that they are *misleading*. In particular, Nielsen et al.[36] pointed out that the composites will be somewhat different depending on the epoch at which they are compiled. This is because the shorter timescale series making up the composite come from the more recent past, which may not be representative of the behavior further back in time.[37] For example, the dashed lines in Figure 2.4A already indicate that the weather–macroweather transition scale depends on whether we include the industrial epoch. Anthropogenic warming increases the multidecadal variability, yielding a shorter transition scale (see Chapter 6).

Indeed, as discussed in Section 5.6 on freak macroweather, the preindustrial macroweather–climate transition time is still not clear, with Greenland ice cores showing an exceptionally long (multimillennial) transition scale in our current post-Ice Age Holocene epoch (Fig. 5.13).[p] Because a Holocene ocean core near Greenland shows a much shorter (centennial-scale) transition, both the epoch-to-epoch (Holocene, pre-Holocene) as well as the geographical variations of the transition scale need clarification. Currently there is a PAGES (Past Climate Change) working group on climate variability across scales (CVAS) that is investigating this issue. In Figure 2.4A, this exceptional Holocene feature is hidden, because the middle (EPICA ice core) curve is an average over the past 800,000 years (the middle and late Pleistocene epoch). However, throughout this nearly 1,000,000 -year-long epoch, our interpretation of Figure 2.4A would be valid as long as it was nuanced by these qualifiers.

Broadening our view to the past 550 million years, it might nevertheless be misleading to think of Figure 2.4A as representative of the entire (Phanerozoic) eon. To explore this further, let's consider what the composites would have looked like had they been produced during the Pliocene epoch, 3.2 million years ago, by Lucy, our famous brainy hominid ancestor.[q] Lucy lived before the Quaternary (see Box 1.2, Fig. 1.5A, bottom) at a time when—thanks to continental drift—the landmasses and ocean currents were sufficiently different so that there were no ice sheets, no glacial–interglacial transitions.

Let's imagine that Lucy was a climatologist who was trying to make sense out of extensive archives of climate data from a wide variety of sources. She got the idea of analyzing the fluctuations systematically and making composites that show the broad sweep of atmospheric variability. Her goal was to clarify and classify an otherwise complex mass of information. If her paleodata were good enough, she might have been able to detect the coming and going of ice sheets in the Late Eocene to Early Oligocene, 28 to 31 million years earlier[38] (Fig. 1.5A, middle), but the temporal resolution of that ancient data would have been poor and her centennial-scale data

---

[p] As discussed in Section 5.6, there may be a more mundane explanation: The ice core data are not representative of global conditions.

[q] Lucy was an *Australopithecus afarensis*, discovered in Ethiopia by Donald Johanson and Tom Gray on November 24, 1974.

would have come from her own ice sheet-free epoch. Over the range of scales from centuries to a million years, Lucy's composite would probably not reproduce the middle bump in Figure 2.4A[r] at 100,000-year scales.

But how different would her composite really have been? First consider the basic weather and macroweather regimes—the nine or ten orders of magnitude of the graph covering milliseconds to decades. This part of Lucy's composite would have been almost identical to ours. We can be fairly certain of this for several reasons. First, weather and macroweather variability are fundamental characteristics of turbulence, a fact that is underscored (Section 4.8) by the finding that nearly identical weather and macroweather regimes exist on Mars. Second, very similar regimes are also well reproduced by climate or turbulent models when they are run with fixed external conditions ("control runs," Chapter 5, Fig. 5.10). Although Lucy's orbital parameters were different, this difference would only have been important at tens of thousands of years (Box 5.4); Lucy's GCM would also produce weather and macroweather regimes, and they would have today's statistics.[39]

Let's now turn our attention to the right-hand side of Figure 2.4A, which includes the temperature variability for periods of hundreds of thousands of years and longer—the megaclimate regime. The megaclimate turns out to be quite robust—at least its most recent 65 million years (the Cenozoic era)—and Lucy's composite would also have agreed with ours about this. This was demonstrated[40] using the Zachos stacks (Fig. 1.5A, middle) with fluctuations that are shown in the second curve from the right in Figure 2.4A. This Cenozoic series was analyzed over six successive 10-million-year periods, and each of these segments was found to have very similar variability over the range of 500,000 years on up.[41] This is not surprising because, at these long timescales, (bio)geological processes drive the system, and they are also likely to be scaling.[s] Lucy's composite would therefore be quite similar to our own, with the main exception being the scales from centuries to hundreds of thousands of years.

Putting this together, if we restrict the applicability of Figure 2.4A to the past 800,000 years, then there would be relatively minor differences, depending on when the series was compiled. Even if this is extended to 65 million years, with the added caveats discussed earlier, it would also be a realistic picture of atmospheric variability.

But Lucy would have understood that the real importance of her composite was paradigmatic: that atmospheric variability should be understood using a relatively small number of wide-range scaling dynamical processes, with quasi-periodic scalebound processes generally playing only minor roles.[42] Yet, in different epochs, the boundaries between the regimes will shift to shorter or longer timescales, and using high-resolution data unavailable to us, Lucy might have discovered a new scaling "mesoclimate" regime in the place of our climate and macroclimate. Nevertheless, her composite would still lead her to conceptualize the climate system in a scaling way rather than a scalebound way.

---

[r] Interestingly, had she lived 30 millions earlier, she might have again produced a composite similar to our Figure 2.4A.

[s] Geological scaling is likely because the processes of erosion and continental drift don't have preferred scales, and the field for which we have the best data—Earth's topography—is also scaling. There is also evidence for scaling in geogravity, geomagnetism, and seismicity. See a review in Lovejoy, S. & Schertzer, D. Scaling and multifractal fields in the solid earth and topography. *Nonlin. Proc. Geophys.* **14**, 1–38 (2007).

Writing a decade and a half after Mitchell, leading climatologists Nicholas Shackelton and John Imbrie laconically noted that their own spectrum was "much steeper than that visualised by Mitchell," [43] a conclusion already anticipated five years earlier (Fig. 2.11) and subsequently reinforced by several scaling composites.[44] Over at least a significant part of this range, Carl Wunsch[45] further underlined its misleading nature by demonstrating that the contribution to the variability from specific frequencies associated with specific "spikes" (presumed to originate in oscillatory processes) was much smaller than the contribution resulting from the continuum.

But none of this perturbed the dominant scalebound view. Just as van Leeuwenhoek peered through the first microscope and discovered a new world, today we automatically anticipate finding new worlds by zooming in or out of scale. It is a scientific ideology so powerful that even quadrillions do not shake it.

The scalebound view led to a framework for atmospheric dynamics that emphasized the importance of numerous processes occurring at well-defined timescales, the quasi-periodic "foreground" processes illustrated as bumps— the signals—on Mitchell's nearly flat background that was considered to be an unimportant noise.[t] Although in Mitchell's original figure the lettering is difficult to decipher, Figure 2.3B spells them out more clearly with numerous conventional examples. For example, QBO is the "Quasi-Biennial Oscillation," ENSO is the "El Niño Southern Oscillation," PDO is the "Pacific Decadal Oscillation," and NAO is the "North Atlantic Oscillation." The Dansgaard-Oeschger and Milankovitć and tectonic cycles[u] will be discussed in Chapter 5 at longer timescales. The point here is not that these processes, mechanisms, are wrong or nonexistent[v]; rather, it is that—at best—they only explain a small fraction of the overall variability.

Even the nonlinear revolution was affected by scalebound thinking. This included atmospheric applications of low-dimensional deterministic chaos. When chaos techniques were applied to weather and climate, the spectral bumps were associated with specific chaos models, analyzed with the help of the dynamical systems machinery of bifurcations, limit cycles, and the like.[46] Of course—as discussed later—from the alternative scaling, turbulence view, wide-range continuum spectra are generic results of systems with large numbers of interacting components (degrees of freedom)—stochastic chaos[47]—and are incompatible with the usual small number of interacting components (low-dimensional) deterministic chaos (Box 2.2).[w] Similarly, whenever there are no dynamically important characteristic scales or scale breaks, the spectra will be scaling (i.e., power laws[48]; Chapter 3).

---

[t] Mitchell actually assumed that his background was either a white noise or over short ranges, sums (integrals) of a white noise.

[u] Figure 2.3B refers to these as cycles rather than oscillations, perhaps because they are broader.

[v] Although, in some cases, maybe they are!

[w] During the early 1980s, excitement about chaos was so strong that enthusiasm sometimes replaced cool heads. A famous article published in *Nature* that was based on a new time series analysis algorithm even claimed that four interacting components were enough to describe and model the climate! Nicolis, C. & Nicolis, G. Is there a climate attractor? *Nature* **311**, 529 (1984).

### 2.1.2 NEW WORLDS AND THE METEOROLOGICAL ZOO

At weather scales, and at virtually the same time as Mitchell proposed a scalebound framework for temporal variability, Isidoro Orlanski proposed a scalebound spatial classification of atmospheric phenomena by powers of ten.[49] Figure 2.5 shows a reproduction of Orlanski's phenomenological space–time diagram[x] with eight different dynamical regimes indicated on the right according to their spatial scales. The diagram does more than just classify phenomena according to their size; it also relates their sizes to their lifetimes. Along the diagonal, various preexisting conventional phenomena are indicated, including fronts, hurricanes, tornadoes, and thunderstorms. The dashed straight-line embellishment was added by my colleagues and me in 1997[50]; it shows that the figure actually shows scaling, not scalebound behavior! This is because straight lines on logarithmic plots such as this are power laws, and even the slope of the line ($-3/2$) turns out to be predicted theoretically using the energy rate density (Chapter 4).
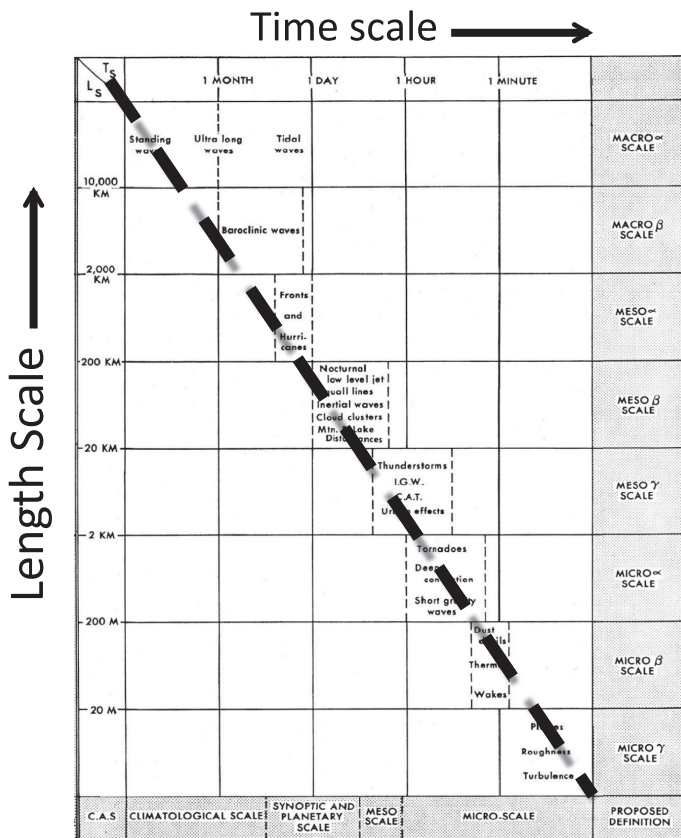


FIGURE 2.5 *Orlanski's space–time diagram with eight different dynamical regimes indicated on the right according to their spatial scales. Notice that he indicates that the climate starts at about two weeks (bottom row). The straight line shows that the figure is actually scaling (straight on this logarithmic plot).[51]*

---

[x] Sometimes called "Stommel diagrams" after Henry Stommel (1920–1992), who produced such diagrams in oceanography.

At the time of Orlanski's classification, meteorological thinking was already largely scalebound. There were several reasons for this. One was, in part, a result of its near-total divorce from turbulence theory, which was primarily scaling.[y] Another resulted from its heritage from the older more qualitative traditions of "synoptic" meteorology with its "zoo"[52] of atmospheric shapes, morphologies, and structures.[z] Finally, I should note the influence of analytical linearized approaches, which were the only ones available in the precomputer era. Therefore, Orlanski's classification rapidly became popular as a systematic rationalization of an already strongly, phenomenologically scalebound approach. It is ironic that just as Orlanski tried to perfect the old scalebound approach, and unbeknown even to the modelers, the computer revolution was ushering in the opposite scaling approach: NWPs and GCMs.[53]

## 2.2  Scaling: Big whirls have little whirls and little whirls have lesser whirls

### 2.2.1  THE FRACTAL REVOLUTION

Scalebound thinking is now so entrenched that it seems obvious that "zooming in" to practically anything will disclose hidden secrets. Indeed, we would likely express more wonder if we zoomed in only to find that nothing had changed, if the system's structure was scaling! Yet during the past thirty years, antiscaling prejudices have started to unravel. Much of this is thanks to Mandelbrot's pathbreaking *Fractals, Form, Chance and Dimension*[54,aa] (1977) and *The Fractal Geometry of Nature*.[55] Thanks to his avant-garde use of stunning computer graphics, his books made an immediate visual impact. They were the first to display the realism of scaling. One was also struck by the word "geometry" in the title. The last time scientists had emphasized geometry was sixty years earlier, when D'Arcy Thompson[56] (1860–1948) used it brilliantly to understand the shapes of diatoms and other biomorphologies. Although Mandelbrot's simulations, imagery, and scaling ideas sparked the fractal strand of the nonlinear revolution—and continue to transform our thinking—his insistence on geometry is now nearly forgotten. The basic reason is that scientists are—in my opinion rightly so—more interested in statistics than in geometry. There is also a less obvious reason: The most interesting thing to come from the scaling revolution was probably not fractals per se, but multifractals, and—despite attempts—they cannot, in general, be reduced to geometry (Box 2.2).[57]

---

[y] All the fundamental turbulence theories were scaling. The question was whether one or two—or, in some cases, three—scaling ranges were required. This is discussed in detail in Chapter 4.

[z] For example, as examined in Chapter 3, clouds had been classified since the early nineteenth century, but these classifications were based on cloud shapes—morphologies—not size.

[aa] This was actually a translation and extension of his earlier French book: Mandelbrot, B. B. *Les objets fractals, forme, hasard et dimension*. (Flamarion, 1975).

In contrast with a "scalebound" object, Mandelbrot counterposed his new scaling, fractal one:

A scaling object . . . includes as its defining characteristic the presence of very many different elements whose scales are of any imaginable size. There are so many different scales, and their harmonics are so interlaced and interact so confusingly that they are not really distinct from each other, but merge into a continuum. For practical purposes, a scaling object does not have a scale that characterizes it. Its scales vary also depending upon the viewing points of beholders. The same scaling object may be considered as being of a human's dimension or of a fly's dimension."[58]

I had the good fortune to begin my own graduate career in 1976, just as the scalebound weather and climate paradigms were ossifying, but before the non-linear revolution really took off. I was thus totally unprepared and can vividly remember the epistemic shock when, shortly after it appeared, I first encountered *Fractals, Form Chance and Dimension*. Revealingly, it was neither my PhD supervisor, Geoff Austin, nor any other scientific colleague who introduced me to the book, but rather my mother[bb]—an artist—who was awed by Mandelbrot's imagery and fascinated by its artistic implications. At the time, my thesis topic was on the measurement of precipitation from satellites[cc] and I had become frustrated because of the enormous space–time variability of rain that was way beyond anything that conventional methods could handle.[dd] The problem was that there were several competing techniques for estimated rain from satellites and each one was different, yet there was essentially no way to validate any of them, to determine which was the best. Scientific progress was blocked. Fortunately, this didn't prevent radar and satellite remote sensing *technology* from continuing to advance.

Not long after reading Mandelbrot's book, I started working on developing fractal models of rain, so that when I finally submitted my thesis in November 1980, about half of it was on conventional remote-sensing topics and the other half was an attempt to understand precipitation variability by using fractal analyses and fractal models of rain.[59] Given that three of the more conventional thesis chapters had already been published in journals—and had thus passed peer review—I was confident of a rubber stamp by the external thesis examiner. Because I had already been awarded a postdoctoral fellowship financed by Canada's National Science

---

[bb] She was a pioneering electronic artist and had been working with the first color Xerox machines to develop electronic imagery years before the development of personal computers and cheap computer graphics. She used various fractal simulations that I supplied—including some with computer glitches—as "objets trouvés" to make abstract images, see Lovejoy, M. *Postmodern Currents: Art and Artists in the Age of Electronic Media*. (Prentice Hall College Division, 1989).

[cc] My thesis was titled *The Remote Sensing of Rain*. McGill University, 1981.

[dd] Conventional methods—both deterministic (of NWP type) and stochastic but scalebound—are still in vogue, but during the past ten years, our understanding of precipitation has been revolutionized by the application of the first satellite-borne weather radar [the Tropical Rainfall Measurement Mission (TRMM)], which has demonstrated unequivocally that—like the other atmospheric variables—precipitation is a global-scale cascade process that is distinctive primarily because its intermittency parameter is much larger than for the other fields. Lovejoy, S., Pinel, J., & Schertzer, D. The global space–time cascade structure of precipitation: Satellites, gridded gauges and reanalyses. *Adv. Water Res.* **45**, 37–50 (2012).

and Engineering Research Council, I happily began preparing for a move to Paris to take it up at the Météorologie Nationale (the French weather service).

But rather than getting a nod and a wink, the unthinkable happened: My thesis was rejected! The external examiner, David Atlas (1924–2015), then at the National Aeronautics and Space Administration (NASA), was a pioneering radar meteorologist who was involved in then-fashionable scalebound mesoscale theorizing (Chapter 4). Atlas was clearly uncomfortable with the fractal material; but, rather than attacking it directly, he instead claimed that although the content of the thesis might be acceptable, its structure was not. To his way of thinking, there were in fact two theses, not one. The first was a conventional one that had already been published, whereas the second was a thesis on fractal precipitation that, according to him, was unrelated to the first. The last point piqued me because it seemed obvious that the fractals were there in an attempt to overcome longstanding problems of untamed space–time variability. On the contrary, they were *very* relevant to a remote-sensing thesis.

At that point, I panicked. According to the McGill thesis regulations, I had only two options: either I accept the referee's strictures, amputate the offending fractal material, and resubmit, or I could refuse to bend. In the latter case, the thesis would be sent without change to *two* external referees, both of whom would have to accept it—a highly risky proposition. Although I was ready to defend the fractal material, I knew full well it had not received serious critical attention. There might be errors that would sink the whole thing. A second rejection would be disastrous because, then, McGill would not permit me to resubmit a thesis on the same topic. Therefore, before making a decision—and with the encouragement of Austin—I contacted Mandelbrot, visiting him at his Yorktown heights IBM office in January 1981.

Mandelbrot was very positive about the material in the draft thesis. Not being very familiar with atmospheric science,[60] and wanting to give me the best possible advice, he contacted his friend and oceanographer Eric Mollo-Christensen (1923–2009) at the Massachusetts Institute of Technology. Mollo-Christensen advised me simply to remove the fractal material and get the thesis out of the way. I could then try to publish it in the usual scientific literature. Beyond that, Mandelbrot advised me to make a short publication out of the analysis part, hinting that we could later start a collaboration to develop an improved fractal model of rain.

With the fractals excised, the thesis was accepted without a hitch,[ee] and at the end of June, a week after defending my thesis, I went off to my Paris postdoc at the

---

[ee] Twenty-five years later, I met up with Atlas, who was by then in his eighties but still occupying an office at NASA. His rejection of my thesis had been a fatherly act, intending to steer me back toward mainstream science. During our discussion, he was mostly intrigued that I was still pursuing the material he had rejected so long ago!

Météorologie Nationale, to work with radar specialist Marc Gilet.[ff] In literally my first week in the French capital, I wrote up the analysis part—the empirical rain and cloud area–perimeter relation (Fig. 2.6)—and submitted it to *Science*.[gg] A few months later, I started a series of month long visits to Mandelbrot in Yorktown Heights. This collaboration eventually spawned the Fractals Sums of Pulses (FSP) model[61] close to the *H* model described in Box 2.3.
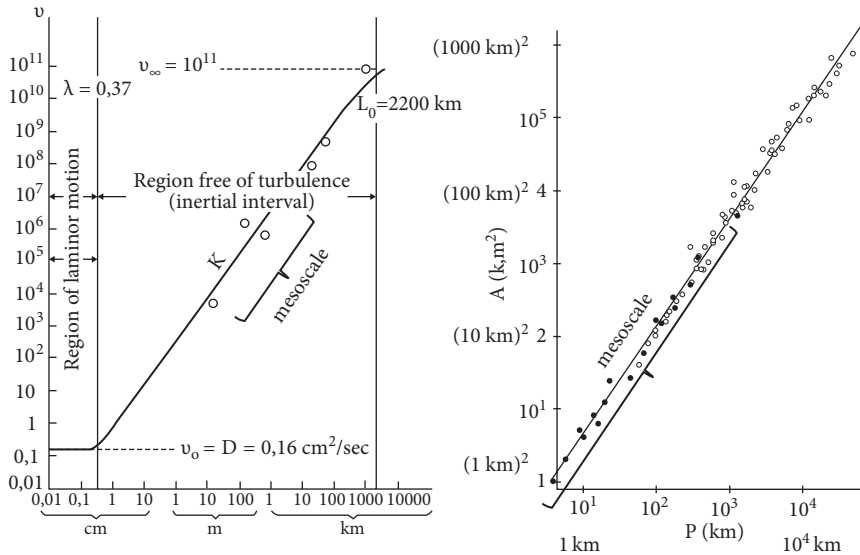


FIGURE 2.6  *(Left) Richardson's proposed 4/3 law of turbulent diffusion,[62] which includes a few estimated data points. (Right) The area–perimeter relation for radar-detected rain areas (black) and infrared satellite cloud images (open circles). The perimeter is the horizontal axis; the area, the vertical axis. The slope corresponds to D = 1.35. The mesoscale (roughly 1–100 km) is shown in the brackets—nothing special.[63]*

### 2.2.2  FRACTAL SETS AND DIMENSIONS

The converse of scalebound objects are scaling, fractal objects.[64] Let's take a look at some examples and see how it's possible for zooming to lead to sameness: to old worlds, not to new ones.

An object is scaling if, when blown up, a small part in some way resembles the larger whole. An unavoidable example is the icon that now bears his name—the

---

Mandelbrot set, the black silhouette in Figure 2.7. It can be seen that, after a series of blowups, we find reduced-scale copies of the original (largest version) of the set.[hh] Although the Mandelbrot set has been termed "the most complex object in mathematics,"[65] it is simultaneously one of the simplest, being generated by simply iterating the algorithm: "I take a number, square it, add a constant, square it, add a constant, . . . ."[66] Precisely because of this algorithmic simplicity, it is now the subject of a small cottage industry of computer geeks who superbly combine numerics, graphics, and music. YouTube is replete with examples; the last time I looked, the recordholder displayed a zoom by a factor of more than $10^{4000}$ (a one followed by 4,000 zeroes)!



FIGURE 2.7 *The scaling approach. Looking through the microscope at the Mandelbrot set (the black in the upper left square), after two magnifications (lower left and right-hand side), Mandelbrot noticed one of an infinite number of reduced-scale versions (slightly deformed, near the lower right-hand corner).*[67]

The Mandelbrot set may be easy to generate, but it is hardly easy to understand. To understand the scaling, fractal idea, consider instead the simplest (and historically the first) fractal: Cantor's "perfect" set (Fig. 2.8A). Start with a segment one unit long (infinitely thin; this is mathematics!)—the "base." Then, remove the middle one third; this is the "motif" (second from the top in Fig. 2.8A). Then iterate by removing the middle third of each of the two one-third-long segments from the previous. Continue iterating so that at every level, you remove all the middle segments before moving to the next level. When this is repeated to infinitely small segments, the result is the Cantor set.[68] From Figure 2.8A, we can see that if either the left or right half of the set is enlarged by a factor of 3, then one obtains the same set.

---

[hh] The small versions are actually slightly deformed versions of the larger ones.
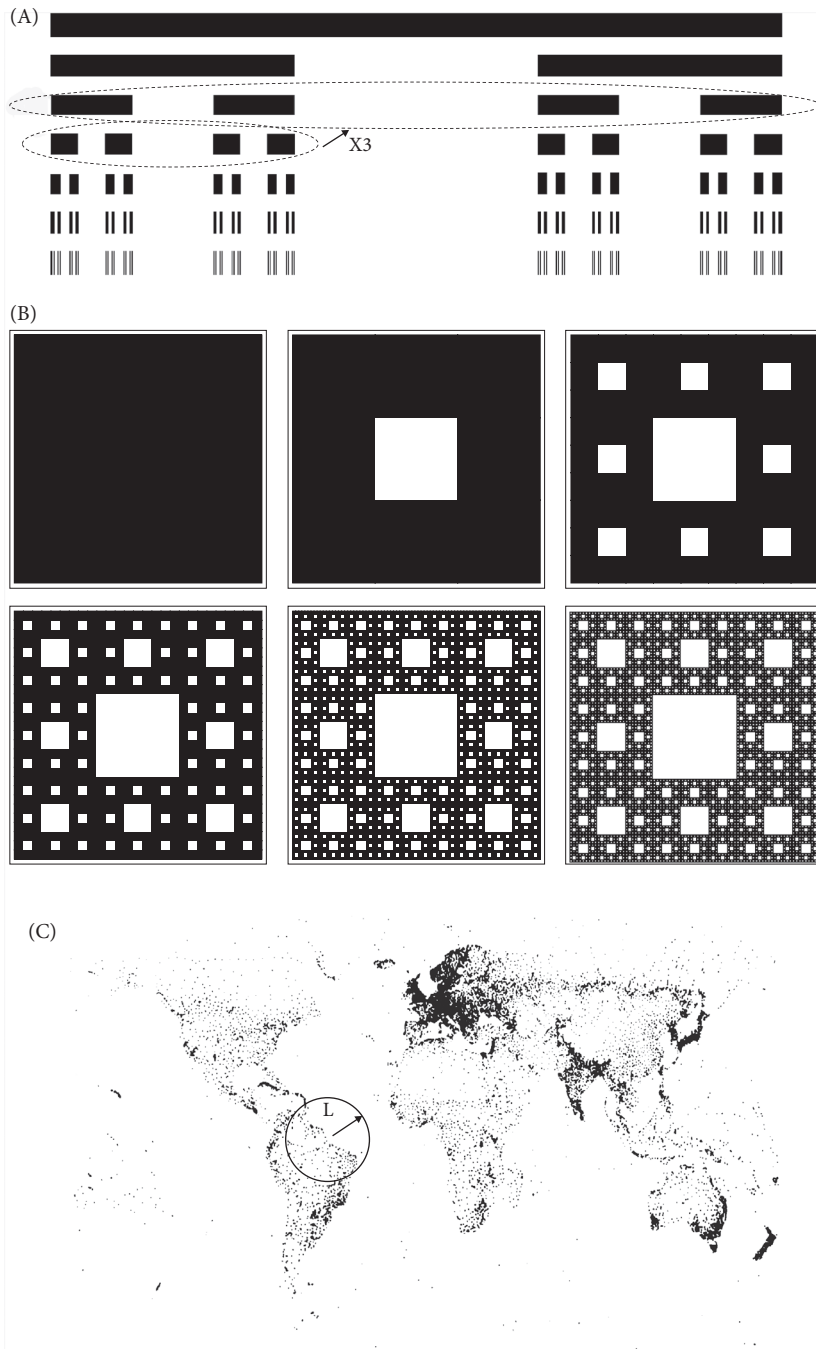
FIGURE 2.8 *(A) The Cantor set. Starting at the top (the "base"), a segment one unit long (the "motif") is obtained by removing the middle third. The operation of removing middle thirds is then iterated to infinitely small scales, the first six iterates are shown. The dashed ellipses show the property of self-similarity: The left-hand half of any level (not only the one explicitly shown), when blown up by a factor of 3, gives the next level up.*
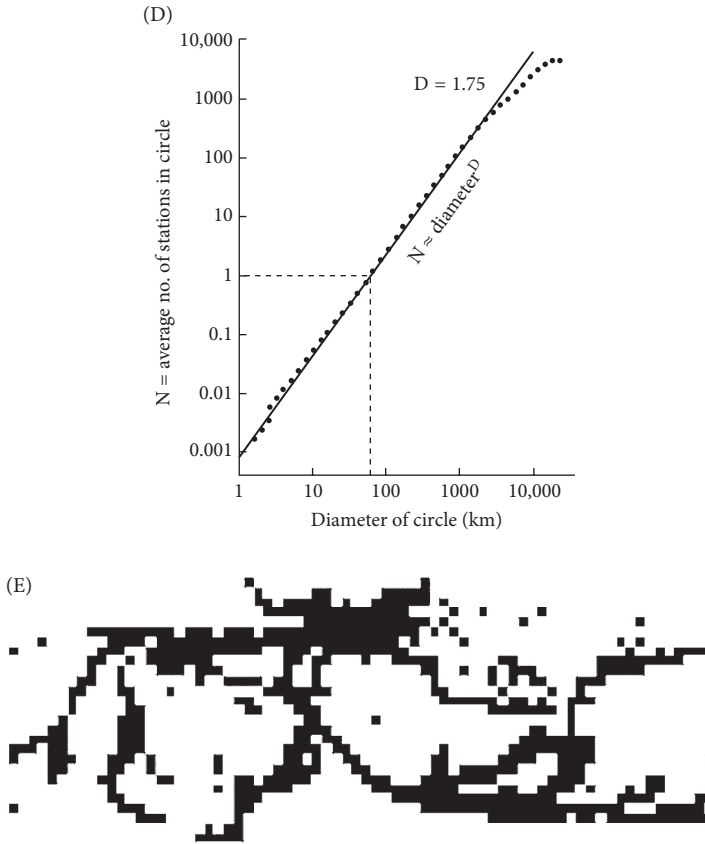
FIGURE 2.8  *Continued*

*(B) The construction of the Sierpinski carpet. The base (upper left) is transformed to the motif (upper middle) by dividing the square into nine subsquares, each one third the size, and then removing the middle square. The construction proceeds left to right/top to bottom to the fifth iteration. (C) The geographical distribution of the 9,962 stations that the World Meteorological Organization listed as giving at least one meteorological measurement every twenty-four hours (in 1986). It can be seen that it closely follows the distribution of landmasses and is concentrated in the rich and populous countries. The main visible artificial feature is the Trans-Siberian Railroad. Also shown is an example of a circle used in the analysis.[69] (D) The average number of stations (vertical axis) within a circle radius L horizontal axis (in kilometers). The slope of the top straight line is D = 1.75.[70] (E) The black indicates the 5° × 5° grid points for which there are some data in the month of January 1878 (20% of the 2,560 grid points were filled).[ii] Although highly deformed by this map projection, we can almost make out the South American continent (lower left) as white surrounded by black data that are from ship measurements, and Europe (the central upper black mass of grid points). The fractal dimension of the data is close to that in (C).*

---

[ii] This is from the Hadley Centre Climate Research Unit Temperatures (HadCRUT) data set using an equal angle projection that greatly distorts high latitudes. See Kennedy, J. J., Rayner, N. A., Smith, R. O., Saunby, M., & Parker, D. E. Reassessing biases and other uncertainties in sea-surface temperature observations measured in situ since 1850. Part 2: Biases and homogenisation. *J. Geophys. Res.* **116**, D14104–D14121 (2011).

This property—that a part is in some way similar to the whole—is for obvious reasons called "self-similarity." In this case, the left or right halves are identical to the whole. In atmospheric applications, the relationship between a part and the whole will generally be statistical; small parts are only the same as the whole *on average*. The Cantor set has many interesting properties. For our purposes the main one is its fractality, a consequence of its self-similarity.

Let's consider it a little more closely. After $n$ construction levels, the number of segments is $N = 2^n$, and the length of each segment is $L = (1/3)^n$. Therefore, $N$ and $L$ themselves are related by a power law. Eliminating the level $n$, we find $N = L^{-D}$, where $D = \log(2)/\log(3) = 0.63 \ldots$. $D$ is the fractal dimension ("log" for logarithm). In this case, it is called the "box-counting" dimension because—if we considered a fully formed Cantor set—the number of segments $L$ (one-dimensional "boxes") that we need to cover the set would be the same[71] $N$. If this fractal dimension seems a bit weird, consider what would happen if we applied the definition to the entire initial (one-dimensional) line segment. We can check that we do indeed recover the usual value for a line: $D = 1$. To see this, consider what would happen if we did not remove the middle third (we kept the original segment), but analyzed it using the same reasoning. In this case, would still divide by three at every iteration so that, as before, $L = (1/3)^n$, but now the number of segments is simply $N = 3^n$ instead of $2^n$. This would lead to the dimension $D = \log 3/\log 3 = 1$, simply confirming that the segment does indeed have the usual dimension of a line.

When a quantity such as $N$ changes in a power law manner with scale $L$, it is called "scaling," and thus $N = L^{-D}$ is a *scaling law*. Contrary to a scalebound process that changes its mechanism (its "laws") every factor of 10 or so, a unique scaling law may hold over a wide range of scales. For the Cantor set and other mathematical fractals, it holds over an infinite range. Of course, real physical fractals can only be scaling over finite ranges of scale. There is always a smallest and largest scale beyond which the scaling will no longer be valid.

Why does a power law imply scaling (and vice versa)? The answer is simply that if $N = L^{-D}$ and we zoom in by a factor $\lambda$ (so that $L \to L/\lambda$), then we see that $N \to N\lambda^D$; the *form* of the law is unchanged. Writing this in words and referring to the dashed ellipses in Figure 2.8A, we could write

Number of segments inside the large ellipse = (Length of large ellipse)$^D$
Number of segments inside the small ellipse = (Length of small ellipse)$^D$,

where, in this case, $\lambda = 3$ (the ratio of the large ellipse to the small one) and $D = 0.63$, as explained earlier.

In contrast, for a scalebound process, changing scales by zooming would give us something quite different. Whenever there is a scaling law, there is something that doesn't change with scale, something that is *scale invariant*. In the previous example, it is the fractal dimension $D$. No matter how far we zoom into the Cantor set, we will always recover the same value $D$. Self-similarity is a special case of *scale invariance* and it occurs when—as its name suggests—some aspect of the system is unchanged under a usual blowup. In physics, quantities (such as energy) that are invariant (conserved) under various transformations are of fundamental

importance, hence the significance of exponents such as fractal dimensions that are invariant under scale transformations.

More generally, a system can be invariant under more generalized "zooms" (i.e., blowups combined with stretchings, rotations, or other transformations). As an example, let's return to the graph of the Weierstrass function, which is scale invariant but *not* self-similar. To show that it is indeed scale invariant, we must combine a blowup with a squashing—or, alternatively, a blowup by different factors in each of the coordinate directions. This property is shown in Figure 2.2A by comparing the full Weierstrass function on the interval between zero and one, with the upper right that shows the left half[72] stretched in the horizontal direction by a factor of 2 and stretched in the vertical direction by a factor of $2^H = 1.26$. Objects that are scale invariant only after being blown up by different factors in perpendicular directions are called "self-affine." The graph of the Weierstrass function is thus self-affine. As we study at length in Chapter 3, scale invariance is still much more general than this.

On the other hand, in the infinitely small limit, the Cantor set is simply a collection of disconnected points (Mandelbrot calls such sets "dusts"),[jj] and a mathematical point has a dimension zero.[73] The Cantor set is thus an example of set with a fractal dimension of 0.63... that is between zero and one, and $D$ quantifies the extent to which it fills the line. Sets of points with such in-between dimensions (they are usually noninteger) are fractals.[kk] More generally, for the purposes of this book, a fractal is a geometric set of points that is scale invariant.[74]

As another mathematical example, consider Figure 2.8B, the "Sierpinski carpet."[75],[ll] The figure shows the base (upper left) and motif (upper middle) obtained by dividing an initial square into squares one third the size, and then removing the middle one. The bottom right shows the result after five iterations. Using the same approach as that described earlier, after $n$ construction steps (levels), the number of squares is $N = 8^n$, and the size of each is $L = (1/3)^n$. Thus, $N = L^{-D}$, with $D = \log 8 / \log 3 = 1.89...$ . Indeed, the Cantor set, the Sierpinski carpet, and the unit segment illustrate the general result:

$$\text{Probability} \approx (\text{Scale})^C$$

where the scale is identified with the length of the side of a square.

---

[jj] At some stage in the construction, any connected segment would have been cut by the removal of a middle third.

[kk] As a result of nontrivial mathematical issues, there are numerous mathematical definitions of dimension. A full discussion of them would take us too far afield.

[ll] This construction and the analogous one based on removing middle triangles is credited (in 1916) to Waclaw Sierpinski (1882–1969). The name "carpet" was added by Mandelbrot. Mobile phone and Wi-Fi antennae have been produced using a few iterations of the Sierpinski carpet, exploiting their scale invariance to accommodate a wide range of frequencies. The Sierpinski triangle goes back to at least the thirteenth century; it is seen in churches as decorative motifs.

Just as the Cantor set has a fractal dimension $D = 0.63...$ between zero and one—between a point and a line—the value of $D$ for the Sierpinski carpet is between one and two (i.e., between a line and a plane), and it quantifies the extent that the Sierpinski carpet exceeds a line while partially filling the plane. These examples show a basic feature of fractal sets. As a result of their hierarchical clustering of points, they are "sparse." Their fractal dimension quantifies their sparseness.

Although the number of boxes gives us information about the absolute frequency of occurrence of parts of the set of size $L$, it is often more useful to characterize the density of the boxes of size $L$, which is obtained by dividing the number of boxes needed to cover the set by the total number of possible boxes—for example, the Cantor set by $L^{-1}$ and the Sierpinski carpet by $L^{-2}$ because they are sets on the line ($d = 1$) and plane ($d = 2$), respectively.[76] This ratio is their relative frequency. In other words, it is the probability that a randomly placed segment ($d = 1$) or square ($d = 2$) will happen to land on part of the set. The ratio is $L^{-D}/L^{-d} = L^{d-D} = L^{C}$, where $C = d - D > 0$ is the *codimension* of the set. Although $D$ measures *absolute* sparseness and frequencies of occurrence, $C$ measures *relative* sparseness and probabilities of occurrence. For the Cantor set, $C = 1 - \log2/\log3 = 0.36...$ , and for the Sierpinski carpet, $C = 2 - \log8/\log3 = 0.11...$ , so that their relative sparsenesses are not so different. If I put a circle (or square) of size $L$ at random on the Sierpinski carpet (iterated to infinitely small scales), the probability of it landing on part of the carpet is $L^{0.11}$, whereas for the Cantor set, putting a random segment of length $L$ would have almost the same probability of landing on the set: $L^{0.36}$. Notice that in both cases the probability gets smaller as the scale $L$ is reduced. This is because small segments/boxes are more likely to land in "holes" than larger ones.

This example illustrates the general result

$$\text{Probability} \approx (\text{Scale})^{C}.$$

In science, we're usually interested in probabilities, so that fractal codimensions are generally more useful than fractal dimensions.

An example of fractal sets that is relevant to atmospheric science includes the locations where meteorological measurements are taken, represented as points in Figure 2.8C. In this case, the set is sparse because the measurement stations are concentrated on continents and in richer nations. To estimate its fractal dimension, one can place circles of radius $L$ on each station[77] (one is shown in the figure) and determine the average number of other stations within distance $L$. If one repeats this operation for each radius $L$, averaging over all the stations, one finds that, on average,[78] there are $L^{D}$ stations in radius $L$ and that this behavior continues down to a scale of 1 km.[79] For the measuring network, we found $D = 1.75$ (Fig. 2.8D).

Even today, much of our knowledge of the atmosphere comes from meteorological stations. For climate purposes, such as estimating the evolution of the atmosphere over the past century, we must also consider ship measurements, place the data on a grid (typically $5° \times 5°$ in size), and average them over a month (e.g., Fig. 2.8E). It turns out that for any given month, the set of grid points having some

temperature data is similarly sparse,[80] so that both in situ weather and climate data are literally taken on fractal sets. An immediate consequence of a fractal network is that it will not detect sparse fractal phenomena—for example, the violent centers of storms that are so sparse that their fractal dimensions are less than $C$ (in this example, less than about 0.25).

This analysis shows that as we use larger and larger circles, they typically encompass larger and larger voids, so that the number of stations per square kilometer decreases systematically. The measuring network effectively has holes at all scales. This means that the usual way of handling missing data must be revised (Box 5.1). Currently, one thinks of the measuring network as a two-dimensional spatial array or grid of data (three dimensional if we include time), although with some grid points empty because of missing data. According to this way of thinking, because Earth has a surface area of about 500 million $km^2$, each of the 10,000 stations represents about 50,000 $km^2$. This corresponds to a box about 220 km on a side so that atmospheric structures (e.g., storms) smaller than this will typically not be detected. Although it is admitted to be imperfect,[81] the grid is therefore supposed to have a *spatial resolution* of 220 km. Our analysis shows that, on the contrary, the problem is one of inadequate *dimensional resolution*.[82]

### 2.2.3 FRACTAL LINES AND WIGGLINESS

The preceding examples of fractal sets were deliberately chosen so that we could get an intuitive feel for the sparseness that the dimension quantifies. In many cases, we instead deal with sets made up of "wiggly" lines, such as the Koch curve,[83] shown in Figure 2.9A.[mm] The fractal dimension can often quantify wiggliness. The construction proceeds from top to bottom by replacing each straight segment by segments in the shape of the second curve from the top (i.e., made of pieces, each of the original size). Again, after $n$ iterations, we have $N = 4^n$ and $L = (1/3)^n$, hence the fractal dimension is $D = \log 4/\log 3 = 1.26...$ . In this curve, the "wiggles" have a dimension between one and two; the wiggliness is quantified by $D$. Notice that as we proceed to more and more iterations, the length of the curve increases. Indeed, after each iteration, because each segment is replaced with four segments each one third the previous length, the length increases by a factor of 4/3. Therefore, after $n$ iterations, the length is $(4/3)^n$, which becomes infinite as $n$ grows. If a completed Koch curve is measured with a ruler of length $L$ (by definition, such a ruler will be insensitive to wiggles smaller than this[nn]), then in terms of the fractal dimension, the length of the Koch curve is $L^{1-D}$. As the ruler gets shorter and shorter, it can measure more and more details, and the length increases accordingly. Since $D$ (= 1.26 > 1), the length grows as $L^{-0.26}$ and becomes infinite for rulers with a small enough $L$.

---

[mm] If three Koch curves are joined at the vertices of a triangle, one obtains the Koch "snowflake," which is probably more familiar.

[nn] This method is sometimes called the "Richardson dividers method" after Lewis F. Richardson, who first used it to estimate the length of coastlines and other geographical features.
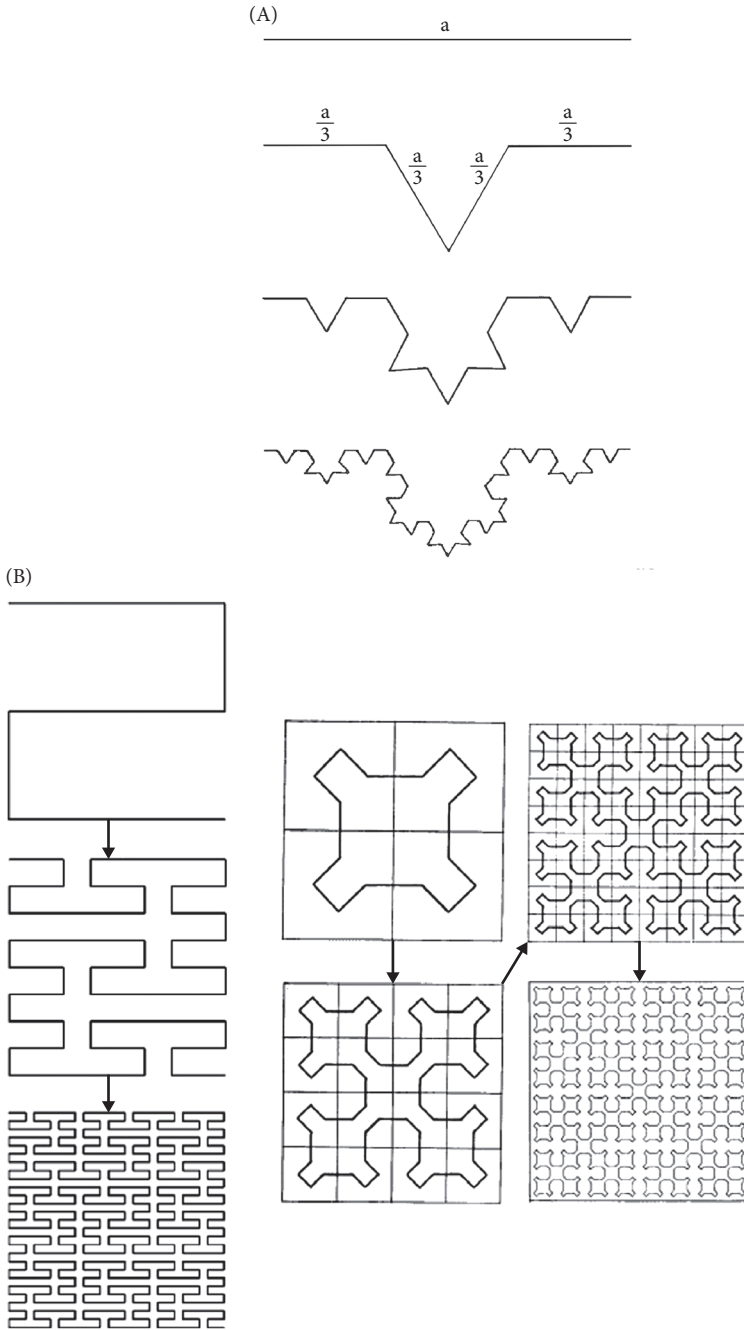
FIGURE 2.9 *(A) (Left) A fractal Koch curve used as a model of the interface between two parts of a turbulent fluid by Weylander.[84] (B) (Left) The first three steps of the original Peano curve, showing how a line (dimension 1) can literally fill the plane (dimension 2). (Right) A variant reproduced from Steinhaus, who used it as a model for a hydrographic network, illustrating how networks of streams can fill a surface.[85]*

How far can we take wiggliness? In 1890, Giuseppe Peano (1858–1932) proposed the fractal construction that bears his name (Fig. 2.9B). The Peano curve is made from a line so wiggly that—by successive iterations—it passes through every single point in a square so that it ends up literally filling part of the plane! At the time, this stunned the mathematical community because it was believed that a square was two dimensional as it required two numbers (coordinates) to specify a point in it. Peano's curve allows the position of a point in a square to be indicated instead by a single coordinate specifying its position on an (infinite) line wiggling its way around the square.[86]

We have already seen another example of wiggliness, the Weierstrass function (Figs. 1.3 and 2.2A), which is constructed by adding sinusoids with geometrically increasing frequencies and geometrically decreasing amplitudes. The Weierstrass function was originally proposed as the first example of a function with a value that is well-defined everywhere (it is continuous), but does not have a tangent anywhere (it is "nowhere differentiable"). A visual inspection (Fig. 2.2A) shows why this is so: To determine the tangent, we must zoom in to find a smooth-enough part over which to estimate the slope. Because the Weierstrass function is a fractal, we in zoom forever without finding anything smooth.

An atmospheric example of a wiggly curve is the perimeter of a cloud, which can be defined from a cloud photograph as the wiggly line separating regions that are brighter or darker than a fixed brightness threshold.[87] To estimate the fractal dimension of a single cloud perimeter, we could try to measure it with shorter and shorter rulers and use the fact that as $L$ gets smaller, the perimeter length increases as $L^{1-D}$ (since $D > 1$). It turns out that it is more convenient to use fixed-resolution satellite or radar images (i.e., $L$ is constant) and to instead use many clouds of different sizes. If we ignore any holes in the clouds,[88] and if their perimeters all have the same fractal dimensions, then their areas ($A$) turn out to be related to the perimeter[89] as $P = A^{D/2}$. Figure 2.6 (right) shows an example when this technique is applied to real cloud and rain areas. Although various theories were later developed to explain the empirical dimension[oo] ($D = 1.35$), the most important implication of this figure is that it gave the first modern evidence of the complete failure of Orlanski's scalebound classification. Had Orlanski's classification been based on real physical phenomena, with each different and acting over narrow ranges of scales, then we would expect a series of different slopes—different fractal dimensions—one for each of his scale ranges.

The expectation that the behavior would be radically different over different scale ranges was especially strong concerning the mesoscale, the horizontal range from about 1 to 100 km, where it was believed that atmospheric thickness would play a key role in changing the behavior: the "mesoscale gap" (Chapter 4). Before this, the only other quantitative evidence for wide-range atmospheric scaling was from various empirical tests of Richardson's 4/3 law of turbulent diffusion.[90] Figure 2.6 (left)

---

[oo] These results were in my original PhD thesis. In the final version they were excised to satisfy the external referee. They were subsequently published in *Science*.

shows his original verification using, notably, data from pilot balloons and volcanic ash.[pp] Throughout the 1950s and 1960s, Richardson's atmospheric 4/3 power law was confirmed repeatedly, with theorists invariably complaining that it extended beyond the range for "which it can be justified theoretically." [qq] However the story didn't end there. During the 1970s, in the wake of the two-dimensional isotropic turbulence theory, a large-scale balloon version, the EOLE experiment was undertaken and claimed to invalidate his law. In Chapter 4, I describe the saga of how this conclusion was based on an erroneous analysis, and how (partially) in 2004 and (fully) in 2013, Richardson was vindicated.

---

BOX 2.2  Intermittency, multifractals, and the α model

In Section 1.3, we showed that lurking inside an apparently ordinary-looking aircraft temperature transect is a highly intermittent (spiky) signal (the absolute differences between consecutive values). Figure 5.2 gives more examples. It turns out that this spikiness cannot be reproduced by additive fractal constructions such as the Wierstrasse function or the *H* model (Box 2.3). Instead, it requires a multiplicative construction and it yields multifractals. This is the subject of this box.

In Figure 1.6, I pointed out that the extreme spike in the figure was so much above the mean that the probability of it being generated by a process with a standard bell curve–type probability was microscopic. How can we understand and model such a process? Because the temperature field is scaling, the natural thing to do is to see what happens when we degrade the resolution of the spikes systematically. Figure 2.10 shows how averaging the spikes in Figure 1.6 by successive factors of 4 smooths and reduces the resolution until, at the bottom, there are only two halves left. We could imagine a final step in which the two halves are averaged to yield a completely flat (constant) function. Notice that the spiky shape persists at all the lower resolutions. This is because the spikes are clustered and the clustering is hierarchical. For example, the largest spike (just above the letter m in "km" on the axis of the top row) is so strong and so highly correlated with the neighboring spikes, that it continues to stand out as we degrade the resolution.

---

[pp] The ocean is also an example of a stratified turbulent system, and the 4/3 law holds fairly accurately. Richardson tested his law creatively in the ocean using—among other things—bags of parsnips, which he watched diffuse from a bridge. Richardson, L. F. & Stommel, H. Note on eddy diffusivity in the sea. *J. Meteorol.* **5**, 238–240 (1948). Still later, it was verified in the ocean over a range from 10 m to 10,000 km. See Okubo, A. & Ozmidov, R. V. Empirical dependence of the horizontal eddy diffusivity in the ocean on the length scale of the cloud. *Izv. Akad. Nauk SSSR, Fiz. Atmosf. Okeana* **6** (5), 534–536 (1970).

[qq] Meaning, it cannot be accounted for by the dominant three-dimensional isotropic homogeneous turbulence theory, cited from p. 557 of Monin, A. S. & Yaglom, A. M. *Statistical Fluid Mechanics.* (MIT Press, 1975).

Temperature turbulent flux φ at 280m resolution

High to low Resolution: degrading by factors of 4
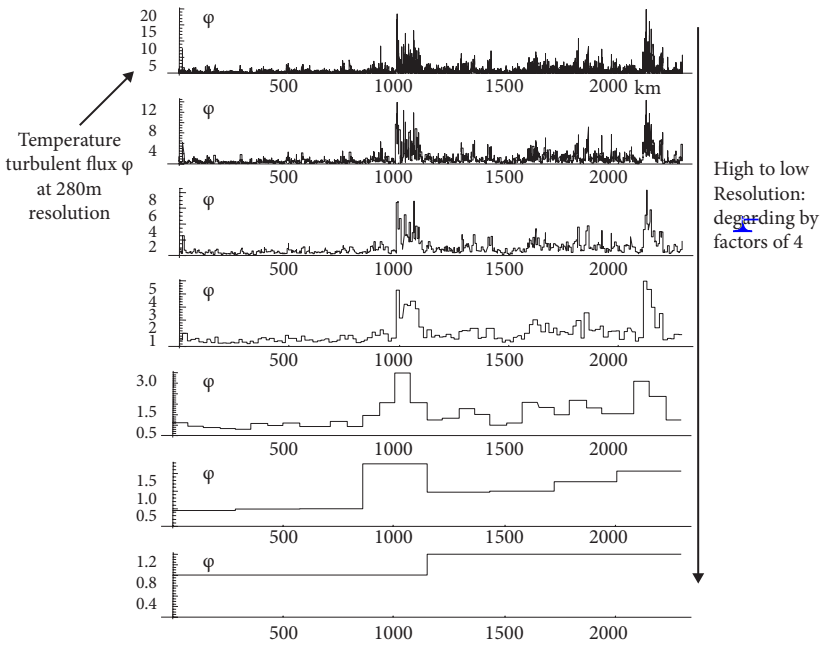
FIGURE 2.10 *The top row is a reproduction of the intermittent spikes taken from the gradients in the aircraft data at the bottom of Figure 1.6. Moving from top to bottom, each row degrades (by averaging) the resolution of the previous by a factor of 4. Note the scale on the left is constantly changing.*[91]

Our objective now is to find a simple model that can reproduce this type of behavior. Let's consider a simple process, the first two steps of which are shown in Figure 2.11: the "α model."[92] The idea is to inverse the order of the left column that increasingly averaged the data; to go from the bottom to the top in Figure 2.10. In other words, start with a completely uniform interval and make it more and more spiky. This time we start with the unit interval—the black line on the left half of the figure, one unit above the axis. We divide it in two, but this time we multiply each half by one of two possible numbers, choosing them randomly by separate "coin tosses" (although, in general, the probabilities will not be equal). In the example shown, the left multiplier was $2^{\gamma_-}$ and the right one was $2^{\gamma_+}$ where $\gamma_-$, $\gamma_+$ are the two basic "singularities" of the model, with $\gamma_- < 0$ and $\gamma_+ > 0$. Because $\gamma_- < 0$, $2^{\gamma_-}$ and it decreases the level of the left-hand side. And because $\gamma_+ > 0$, $2^{\gamma_+}$, so that multiplication increases the level on the right-hand side. Note that we could have any of the four combinations $(\gamma_-, \gamma_-)$, $(\gamma_-, \gamma_+)$, $(\gamma_+, \gamma_-)$, $(\gamma_+, \gamma_+)$—in other words, (decrease, decrease), (decrease, increase), (increase, decrease), (increase, increase)—although only the probabilities of the middle two would generally be equal.
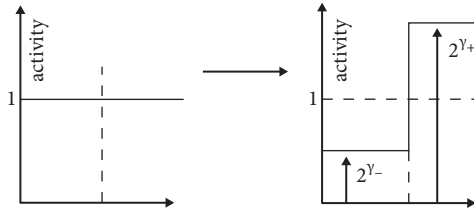
FIGURE 2.11 *The first two steps in the construction of an α model in one dimension. The unit interval (the segment at vertical level 1 on the left) is broken into halves and each is multiplied by a randomly chosen factor, as explained in the text.*

Figure 2.12 shows the result for two different sets of $\gamma_-$, $\gamma_+$. As we iterate the process by dividing and multiplying to smaller and smaller scales, the process becomes more and more spiky (note the vertical axes are constantly changing to accommodate the higher and higher extreme values that are generated). The simulation on the right is particularly extreme. This is an example of the "α model," so named because of the exponent α that it introduces[93].
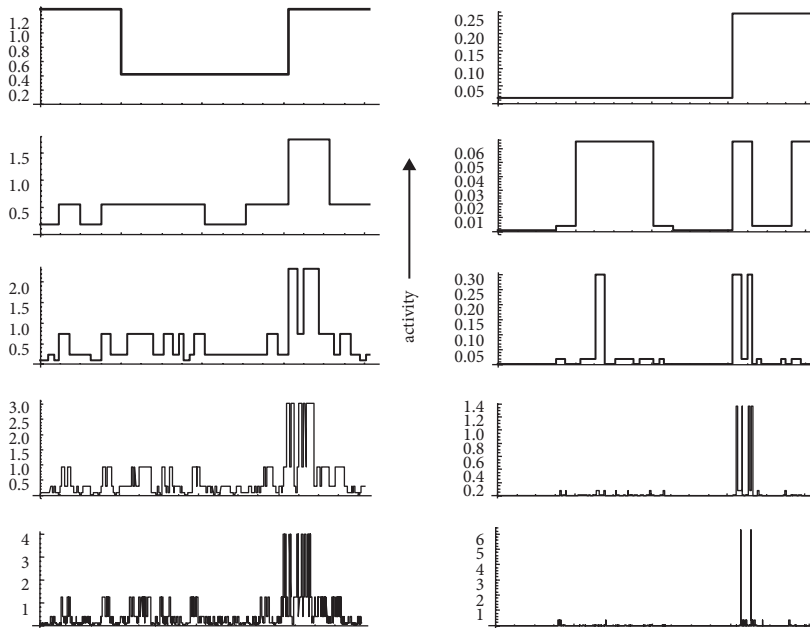


FIGURE 2.12 *Each column (top to bottom) shows every other stage in the construction of an α model over the unit interval (i.e., with resolutions 1/4, 1/16, 1/64, 1/256, and 1/1,024). Each column is a realization of an α model with different parameters ($\gamma_+$, $\gamma_-$) that determine, respectively, the possible boost and reduction factors at each step. Passing from top to bottom is analogous to the redistribution of turbulent energy fluxes in a cascade. We see that most of the flux is in small spiky regions (the result of many boosts).[94]*

The α model is the simplest multifractal model and it was developed explicitly as a model of turbulent intermittency. The connection with turbulence and multifractality is perhaps more obvious if we make an α model in two-dimensional space by starting with an initial square and dividing it into four subsquares (Fig. 2.13). As before, each subsquare is multiplied by either 2 or ½ chosen at random with probabilities selected to ensure the average is one: On average, the area under each curve is constant as the construction (cascade) proceeds to smaller scales.
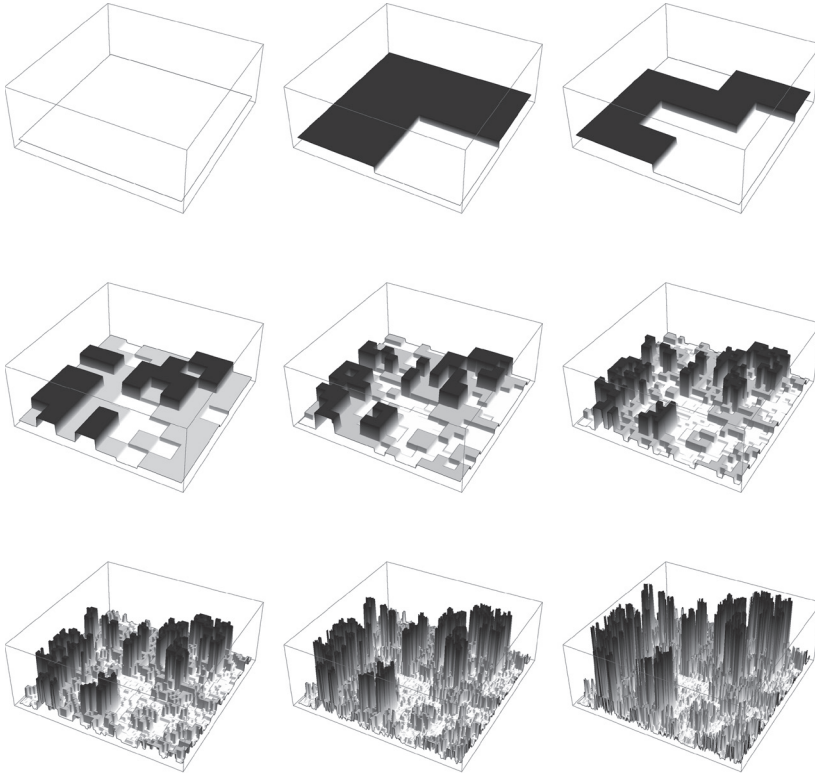


FIGURE 2.13 *This figure shows the first steps in the construction of an α model in two dimensions. The initial state (all white, all equal to one) is in the upper left. Moving to the right and from top to bottom, squares are divided in two in both horizontal directions. The vertical scale is fixed. For each subsquare, the value 1.31 (=2) chosen with a probability of 0.66; otherwise (probability 1 – 0.66 = 0.34), the value 0.38 (= 2) is chosen. On average, the value is 1. These values are then multiplied by the value given by the earlier steps in the process. At each level of activity (the vertical direction), the smallest scale is reduced by a factor of 2 so that, at the bottom right, there are 256 × 256 spikes. To show them clearly, the levels have been shaded from the lowest value (white) to the highest value (black). The result is a multifractal. The regions that exceed a given level of activity each have different fractal dimensions decreasing as we move to higher and higher levels (the higher regions are thus more sparse).*

If we compare this with Figure 2.14—the Richardson-inspired "dead" or "alive" model for the region of turbulent activity—we see that this two-state (dead or alive) model is a special case obtained when the multipliers are either 2 ⬚ost) or decrease to zero (i.e., 2 ⬚ n γ ⬚). Although the dead-or-alive case yields a fractal set of active regions, the α model is multifractal because each level of activity—the spikes at a given vertical level in Figure 2.12 or Figure 2.13—has a different fractal dimension that decreases at more and more extreme levels.



Parent eddy

Daughter eddies

Homogeneous

Intermittent

Grand-daughter eddies

$$D = \frac{Log 4}{Log 2} = 2$$

$$D = \frac{Log 3}{Log 2} = 1.58...$$
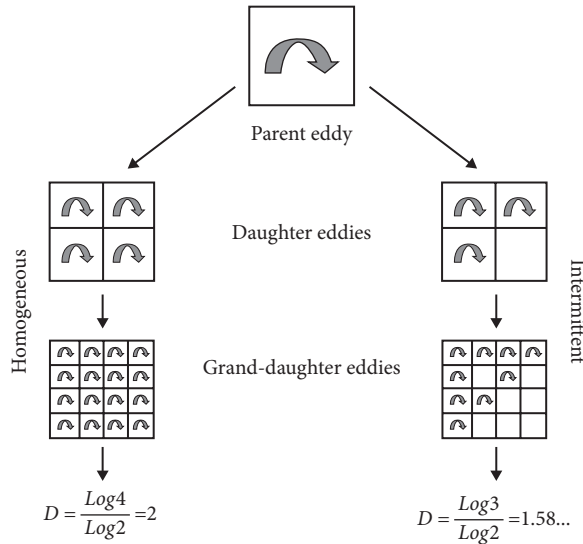
FIGURE 2.14 *A schematic diagram showing the first few steps in a Richardson-inspired cascade process. At each step, the parent eddy (Richardson's "big whirl," top) is broken up into "daughter" eddies, each reduced by a factor of 2 in scale, indicated as squares. The left shows a homogeneous cascade (corresponding to Kolmogorov's 1941 homogeneous turbulence) in which the energy flux (the energy rate density) is simply redistributed from large to small structures while keeping its average density constant. The right-hand side shows an improvement: "On/off" intermittency is modeled by an "alive/dead" alternative at each step (here, only the bottom right subeddy becomes dead). The result is a fractal set of active areas. The fractal dimension is shown at the bottom. For pedagogical reasons, the alternative displayed is purely deterministic, but could be randomized easily (see text).[95]*

In the limit of a large number of steps, the value at each point is multiplied by an infinite number of factors so that it isn't obvious that the process converges at all. It turns out that there is no convergence in the usual sense (i.e., at a geometric point in the horizontal plane).[rr] However, spatial averages over finite regions do converge, but they have extreme probabilities that are so far from the bell curve they can be called

---

[rr] This is why multifractals cannot be thought of as geometric sets of points.

*black swans* (see Box 3.1). Because the small scales are at the origin of the extreme behavior, this has been called[96] the "multifractal butterfly effect."[97]

The α model is the simplest multifractal model. It is an example of what physicists call a "toy model." It has many of the key features of a realistic model, but is simple to understand. To be a realistic model of real-world turbulent processes, the α model must be extended in several ways. One way is to allow for an arbitrary number of different multipliers: a continuous distribution rather than only the two possibilities allowed in the α model. A second generalization is needed to avoid the ugly straight construction lines in Figure 2.13. These lines are a consequence of only allowing reductions in scale by factors of 2. But nature doesn't operate in powers of two; we need continuous-in-scale multifractals. It turns out that, in such multifractals, one obtains a simplification called "universal multifractals" [98] that depend on only three parameters.[99] Finally, we need to remember that these pure multiplicative models are only appropriate for processes (such as the energy flux) that are, on average, conserved from one scale to another (large to small). To model directly observable fields, such as wind or temperature, we need a mixed multiplicative–additive-type process—for example, a kind of marriage between the α model and the *H* model discussed in Box 2.3. Chapter 3 shows many examples of clouds modeled in this way.

These cascade models are produced by large scales modulating smaller scales multiplicatively, and with the smaller scales modulating even smaller scales, repeating over a wide range so that the variability grows from large to small scales, with each level of activity characterized by a different exponent. It turns out there is a straightforward method of displaying this systematic growth of the variability with scale and directly estimating the "outer scale" at which the cascade started. For examples of the method (called "trace moments"[100]) applied to both Earth and Mars for pressure, wind, and temperature, see Figure 4.14. There, it is found that the outer scales of the cascade—the scale at which the variability begins to grow—are (as expected) very near the largest possible distances on each planet.[101]

---

### 2.2.4 RICHARDSON

We have examined several of the famous nineteenth-century fractals: the Cantor set (Fig. 2.8A), the first set with a noninteger dimension; the Peano curve (Fig. 2.8B), the first line that could pass through every point in the unit square (a plane); and the Weierstrass function (Fig. 2.2A), the first continuous curve that doesn't have a tangent anywhere. But for a long time, these were considered to be essentially mathematical constructions—academic oddities without physical relevance. Mandelbrot rehabilitated them and provocatively called them "monsters."

Mandelbrot not only coined the term "fractal," but with his indefatigable energy put them squarely on the scientific map. Although he made numerous mathematical contributions,[ss] his most important one was as a towering pioneer in applying

---

[ss] I will let the mathematicians judge his contributions to mathematics. However, there is no question that Mandelbrot's contribution to science has been monumental and underrated. In any case (and despite Mandelbrot's efforts!), it is still a bit early to evaluate his place in the history of science. For a brief biography, see Barton, C. C., Lovejoy, S., Schertzer, D., & Turcotte, D. L. Benoit B. Mandelbrot

fractals and scaling to the real world. In this regard, his only serious scientific precursor was Lewis Fry Richardson[tt] (1881–1953). As a result of his Quaker beliefs, Richardson was a pacifist, which made his career difficult, essentially disqualifying him from academic positions. He instead joined the Meteorology Office, but quit temporarily to drive an ambulance during the first world war. Afterward, he rejoined the Meteorology Office, but resigned in 1920, when it was militarized by being merged into the Air Ministry.

Richardson worked on a range of topics and is remembered for the nondimensional Richardson number that characterizes atmospheric stability, the Richardson 4/3 law (Fig. 2.6, left), the Modified Richardson Iteration, the Richardson Acceleration techniques of numerical analysis, and the Richardson divider's method. The latter is a variant on box counting (discussed earlier) that he used notably to estimate the length of the coastline of Britain, demonstrating that it followed a nontrivial power law. Mandelbrot's famous 1967 article that initiated fractals[102],[uu] took Richardson's graphs and interpreted the exponent in terms of a fractional dimension.[103] Fully aware of the problem of conceptualizing wide-range atmospheric variability, Richardson was the first to propose explicitly that the atmosphere might be fractal. A remarkable subheading in his 1926 article on turbulent diffusion is titled "Does the wind possess a velocity?" and is followed by the statement: "this question, at first sight foolish, improves upon acquaintance."[104] He then suggested that a particle transported by the wind might have a Weierstrass functionlike trajectory that would imply that its speed (tangent) would not be well defined.[vv]

Richardson is unique in that he straddled the two main—and superficially opposing—threads of atmospheric science: the low-level deterministic approach and the high-level statistical turbulence approach. Remarkably, he was a founding figure for both. His seminal book *Weather Forecasting by Numerical Process*[105] ,[ww] inaugurated the era of numerical weather prediction. In it, Richardson not only wrote down the modern equations of atmospheric dynamics, but also he pioneered numerical techniques for their solution. He even attempted a laborious manual integration.[xx] Yet, this work also contained the seed of an alternative: Buried in

(1924–2010). *EOS Trans.* **93** (4), 44 (2012). Interested readers may consult his posthumously published autobiography: Mandelbrot, B. B. The Fractalist. (First Vintage Books, 2011).

[tt] I have given some historical examples of early geophysical fractal models (Fig. 2.9A and the right side of Fig. 2.9B), but other notable precursors were Jean Perrin (1870–1942), who questioned the differentiability of the coast of Brittany [Perrin, J. *Les Atomes.* (NRF-Gallimard, 1913).] and Hugo Steinhaus (1887–1972), who questioned the integrability of the length of the river Vistula [Steinhaus, H. Length, shape and area. *Colloq. Math.* **III**, 1–13 (1954).]. Lack of differentiability and its converse, integrability, are typical scaling, fractal features. Looking back, these examples are significant, but are isolated. In contrast, Mandelbrot initiated a whole body of work.

[uu] This was still nearly a decade before Mandelbrot coined the word "fractal."

[vv] It turned out that the problem was not the velocity, but the acceleration.

[ww] Lacking support, he paid for the publication out of his own pocket.

[xx] Near the war's end, he somehow found six weeks to attempt a manual integration of the weather equations. His estimate of the pressure tendency at a single grid point in Europe turned out to be badly wrong (as he admitted), but the source of the error was only recently identified. See the fascinating account by Lynch, P. *The Emergence of Numerical Weather Prediction: Richardson's Dream.* (Cambridge University Press, 2006).

the middle of a paragraph, he slyly inserted the now iconic poem describing the cascade idea: "Big whirls have little whirls that feed on their velocity, little whirls have smaller whirls and so on to viscosity (in the molecular sense)."[yy] Figure 2.14 is a schematic showing a modern interpretation that is the basis of turbulent models of intermittency and is the basic multifractal model (see Box 2.2).

Richardson's book was soon followed by the first turbulent law, the Richardson 4/3 law of turbulent diffusion,[106] today celebrated as the starting point for modern theories of turbulence, including the key idea of cascades and scale invariance. Unencumbered by later notions of mesoscale,[zz] and with remarkable prescience, he even proposed that his scaling law could hold from dissipation up to planetary scales (Fig. 2.6, left). Richardson is the precursor of much of the work described in this book, including the area–perimeter analysis (Fig. 2.6, right) and the large body of results that we examine later. Today, he is honored both as the father of numerical weather prediction by the Royal Meteorological Society's Richardson prize and as grandfather of turbulence by the European Geosciences Union's Richardson medal.[aaa]

As a humanist, Richardson worked to prevent war. With his article "The Problem of Contiguity: An Appendix of Statistics of Deadly Quarrels,"[107] he founded the mathematical (and nonlinear!) study of war. He was also anxious that his research be applied to improve the situation of humanity directly, and proposed the construction of a vast "weather factory" that would employ tens of thousands of human "computers" and would make real-time weather forecasts. Recognizing (from personal experience) the tedium of manual computation, he foresaw the need for the factory to include social and cultural amenities.

Let me now explain a deep consequence of Richardson's cascade idea that didn't fully mature until the nonlinear revolution in the 1980s. We have seen that the alternative to scalebound thinking is scaling thinking, and that fractals embody this idea for geometric sets of points. For example, the Koch curve was a model of a turbulent interface, the set of points bounding two different regions; the Peano curve was a model of a hydrographic network. However, to apply fractal geometry to the set of bounding (perimeter) points on a cloud, we were already faced with a problem: We had to reduce the gray shades to white or black (cloud or no cloud). Because atmospheric science does not often deal with black/white sets—but rather with fields, such as cloud brightness or temperature, which have numerical values everywhere in space and that vary in time—something new was necessary.

(Re)consider Figure 1.6, the aircraft temperature transect. We could repeat the treatment of the Weierstrass function to try to fit the transect into the

---

[yy] Richardson, L. F. *Weather Prediction by Numerical Process.* (Cambridge University Press, 1922), p. 66. (Republished by Dover, 1965). This poem was a parody of a nursery rhyme by mathematician Augustus de Morgan, the "Siphonaptera": "Big fleas have little fleas, Upon their backs to bite 'em, And little fleas have lesser fleas, and so, ad infinitum." Quoted from Teskey, G. *The Poetry of John Milton.* (Harvard University Press, 2015).

[zz] That predicted a strong break in the scaling (see Chapter 4).

[aaa] The highest honor of the Nonlinear Processes Division.

framework of fractal sets by simply considering the points on the top graph as the set of interest. But, this turns out to be a bad idea, because as we also saw in Figure 1.6 (bottom), the figure was actually hiding some incredibly variable, spiky (intermittent) changes, and this behavior requires something new to handle it: multifractals.[bbb] Indeed, multifractals were first discovered precisely as models of such turbulent intermittency.[108]

Focus on the "spikes" at the bottom of Figure 1.6. Rather than treating all the points on the graph as a wiggly fractal set, instead consider the set of points that exceed a fixed threshold—for example, those above the level of one standard deviation, as indicated by the horizontal line in the figure—as a kind of Cantor set. If the spikes are scale invariant, then this set will be a fractal with a certain fractal dimension. Now, move the horizontal line a little higher to consider a different set: the spikes that exceed this higher level. We find that the fractal dimension of this different set is lower. Indeed, in this way, moving to higher and higher levels, we could specify the fractal dimension of all the different-level sets, thus completely characterizing the set of spikes by an infinite number of fractal dimensions. The absolute temperature changes (the spikes)—and indeed the temperature transect itself—are thus *multifractals*. It turns out that multifractals are naturally produced by cascade processes that are physical models of the concentration of energy and other turbulent fluxes into smaller and smaller regions. More information about this is included in Box 2.2.

Mathematically, although fractals are scale-invariant geometric sets of points, they are black or white; you are either on or off the set of points. In contrast, multifractals—at least when averaged over small segments and time intervals—are scale-invariant fields. Like the temperature, they have numerical values at each point in space, at each instant in time.

BOX 2.3 Fluctuations and the fractal *H* model

We have seen that both sparse sets of points as well as wiggly lines can be fractal sets, and that fractal dimensions can characterize both sparseness and wiggliness. Usually, we are not very interested in sets of points, but rather in functions—such as the temperature as a function of time. How can fractals help us there? An obvious way is to graph the temperature as a function of time. At every instant in time, plot the value of the temperature.[109] In the simplest "monofractal" model discussed here, if the temperature is scale invariant with respect to changes in timescales, then the set of points on the graph is a fractal set.[110]

We have already given the example of the Weierstrass function, where we introduced the exponent *H* to characterize the fluctuations. It turns out that fluctuations are fundamental. In particular, there is a basic distinction between

---

[bbb] Mathematically, the nontrivial point is that although the Weierstrass function is continuous (i.e., well defined at each instant *t*, a mathematical point on the time axis), a multifractal only converges in the neighborhood of the instant. To converge, the multifractal must be averaged over a finite interval. This is the origin of the "dressed" properties that are related to the "multifractal butterfly effect" and extreme events discussed later in Box 3.1.

functions that have fluctuations that increase in size with scale ($H > 0$) and those that have fluctuations that decrease with scale ($H < 0$).

To understand more fully fluctuations that increase or decrease with scale, let's consider the simple fractal $H$ model.[111] Figure 2.15 shows the first two steps in the construction. The motif is a basic step of width (time interval) $\Delta t$ and amplitude (temperature variation) $\Delta T$. The step represents a basic fluctuation at scale $\Delta t$. It is equal to $-\Delta T/2$ on the left half and $\Delta T/2$ on the right half. The second iteration in the construction is achieved by reducing the timescale by a factor of 1/2, but the amplitude (the vertical direction) by a different amount: $(1/2)^H$. When $H > 0$, the amplitude of the shorter duration fluctuations is less than one half; when $H < 0$, it is larger. We then flip a coin to decide if this reduced-scale fluctuation is left as it is or whether it is flipped upside down (equivalent to multiplying it by a factor of $-1$). We then repeat this a second time and place this second (reduced) replica to the right of the first one as shown (in the example, only the second one is flipped). Notice that the average over an entire fluctuation is zero and, similarly, that the difference in amplitude between the beginning and end of a fluctuation is $\Delta t^H$. We return to this when we consider analyzing the $H$ model.
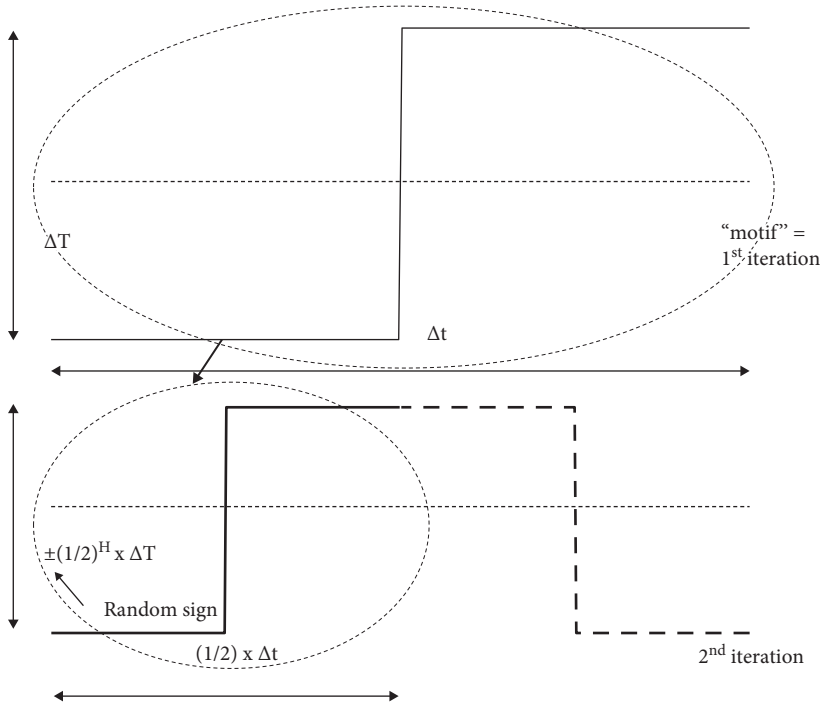


FIGURE 2.15 *The first two steps in the construction of the H model. The basic shape is a step function (top), two copies of which are produced by reducing in scale by a factor of 2 in the horizontal and by a factor of $(1/2)^H$ in the vertical. This reduced-scale copy is then either left as it is (e.g., the left-hand side in the bottom row, the solid line) or "flipped" (as in the right-hand side in the bottom row, the dashed line). Figure 2.16 shows several more stages in the construction.[112]*

We then continue this process (Fig. 2.16), generating a series of levels, each with fluctuations of shorter and shorter duration. To the left of Figure 2.16, we see an example with $H > 0$; on the right, $H < 0$. Finally, the fractal is made by summing the individual iterates (bottom). We see that when $H > 0$ (left), the resulting signal wanders, drifts up or down; on the right, it converges, displaying lots of cancelations, with origins that can be seen in the iterates.
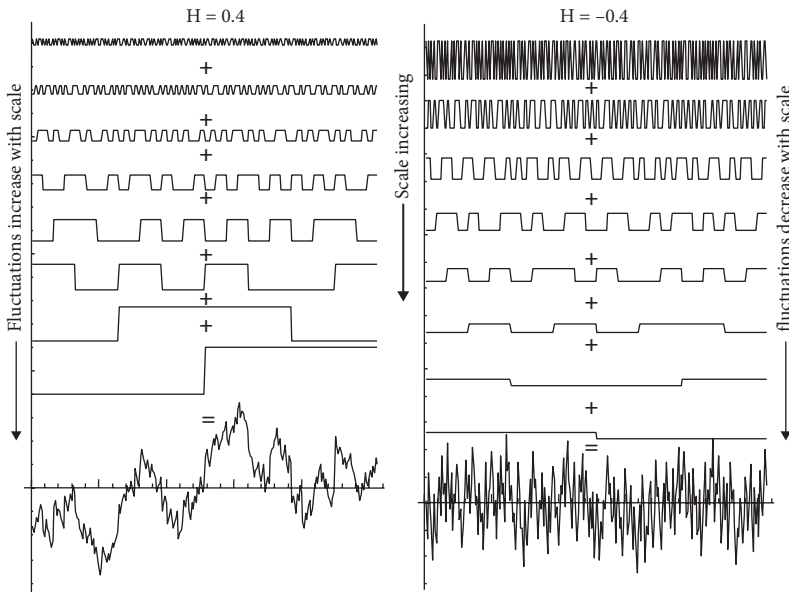


FIGURE 2.16 *Examples of the first seven steps in the construction of H models, with H > 0 (left) and H < 0 (right). The final result is obtained by summing over all the reduced-scale series. The left column corresponds to the weather, the climate, and megaclimate regimes (see Fig. 2.15) whereas the right column corresponds to macroweather and macroclimate.*[113]

If we were given the series on the right or the left, how could we deduce the exponent $H$ that was used in its construction? The easiest to understand is the right-hand simulation ($H < 0$). If we average the simulation over a duration $\Delta t$, we can see that we will average out all the shorter duration iterates so that the remaining variation depends only on the larger scales. Increasing the averaging duration will eliminate more and more of the small iterates, so that we could use the rate that the averaging eliminates the smaller scale fluctuations to estimate $H$.

For the left-hand side, with $H > 0$, the basic idea is the same, although instead of averaging, we must now consider the effect of taking differences between values of the process over longer and longer intervals ($\Delta t$). The differences will be dominated by the fluctuations smaller than $\Delta t$—those that come from the larger $\Delta t$ (lower level iterates) will typically not change the temperature difference over durations of $\Delta t$.

The Haar fluctuation (Fig. 2.17) combines averaging and differencing into a single analysis tool that works for either positive or negative *H*. It gives good estimates for either the left- or right-hand simulations.
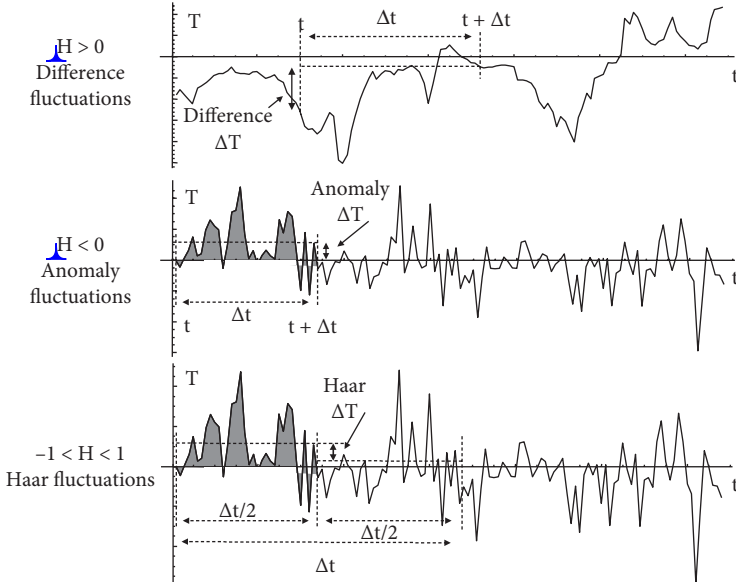


FIGURE 2.17  *(Top) Schematic illustration of difference fluctuations for a multifractal simulation of the atmosphere in the weather regime (0 ≤ H ≤ 1). (Middle) Illustration of the anomaly fluctuation for a series in the macroweather regime with –1 < H < 0. (Bottom) Illustration of Haar fluctuations (useful for processes with –1 ≤ H ≤ 1), applied here to the middle series. The Haar fluctuation over the interval Δt is the mean of the first half subtracted from the mean of the second half of the interval Δt.[114]*

AQ: Heading 2.3 is pro- [...] "Fl[...]tions as a micro- scope" in the table of contents. Please make this consistent.

## 2.3  Fluctuation analysis as a microscope

When confronted with variability over a huge range of space and timescales, we have argued that there are two extreme opposing ways of conceptualizing it. We can either assume that everything changes as we move from one range of scale to another—every factor of 10 or so—or, on the contrary, we can assume that, at least over a wide range (factors of hundreds, thousands, or more), blowing up gives us something that is essentially the same. But this is science; it shouldn't be a question of ideology. If we are given a temperature series or a cloud image, how can we analyze the data to distinguish the two? To tell which is correct? We have already introduced two methods: spectral analysis, which is quite general; and the area–perimeter relation, which is rather specialized. Although spectral analysis is a powerful technique, its interpretation is not so simple. Indeed, had the interpretations been obvious, we would never have missed the quadrillion, and the distinction between macroweather and the climate would have been clarified long ago!

It is therefore important to use an analysis technique that is both easy to apply and easy to understand—a kind of analytical microscope that allows us to zoom in and to compare systematically a time series or a transect at different scales to test the scalebound or scaling alternative directly—fluctuation analysis, which was already briefly discussed in Figures 2.3C and 2.4, and Box 2.1.

We probably all have an intuitive idea about what a fluctuation is. In a time series, it's about the how the values change over an interval in time. Consider a temperature series. We are interested in how much the temperature has fluctuated over an interval of time $\Delta t$. The simplest fluctuation is simply the difference between the temperature now and at a time $\Delta t$ earlier (Fig. 2.17, top). This is, indeed, the type of fluctuation that has traditionally been used in turbulence theory and was used in the first attempt to test the scaling hypothesis on climate data (Fig. 2.18). To make Figure 2.18, first, two instrument series were analyzed: the Manley[115] series from central England starting in 1659[ccc] (open circles) and an early northern hemisphere series from 1880 (black circles). The former was essentially local (regional) and the latter was global in scale. The other series were from early paleoisotope series, using the official calibrations to transform them into temperature values.[ddd]
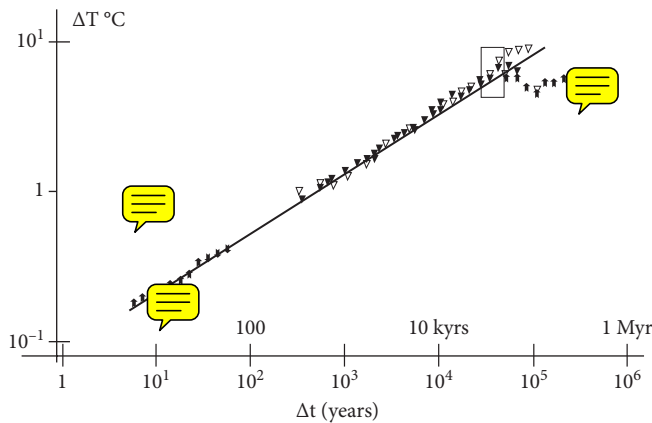


FIGURE 2.18 *The square root of the average squared differences (the root mean square, or RMS, differences; the "structure function") estimated from local (central England) temperatures since 1659 (open circles, upper left), northern hemisphere temperature (black circles), and from paleotemperatures from Vostok (Antarctic, solid triangles), Camp Century (Greenland, open triangles), and from an ocean core (asterisks). For the northern hemisphere temperatures, the (power law, linear on this plot) climate regime starts at about ten years. The rectangle (upper right) is the "glacial–interglacial window" through which the structure function must pass to account for typical variations of ±2 to ±3°C for cycles with half periods of ≈50,000 years.[116]*

---

[ccc] This monthly series from the Greater London area is famous for the being the longest complete series based on real thermometer measurements.

[ddd] Long before the Internet, scanners, and publicly accessible data archives, as a postdoc at the Météorologie Nationale in Paris, I recall taking the published graphs, making enlarged photocopies, and then using tracing paper to digitize them painstakingly.

To make the graph for a given time interval $\Delta t$, one calculates systematically all the nonoverlapping differences in each series and averages the squares of these differences. The typical values shown in the plot are the square root of these (the "root mean squares"). One then plots the results on logarithmic coordinates because, in that case, scaling appears as straight lines and can be identified easily. Reading the graph, one can see, for example, that at ten-year intervals, the typical northern hemisphere temperature change is about 0.2°C and that, over about 50,000 years, the typical temperature difference is roughly 6°C (±3°C), which corresponds to the typical difference of temperature between glacials and interglacials. Hence the box, which allows for some uncertainty, is the "glacial–interglacial window." These fluctuations are therefore straightforward to understand.

In Figure 2.18, a reference line with slope $H = 0.4$ is shown corresponding to the scaling behavior $\Delta T \approx \Delta t^H$, linking hemispheric temperature variations at ten years to paleovariations at hundreds of thousands of years. Although this basic picture is essentially correct, later work provided a number of nuances that help to explain why things were not fully cleared up until much later. Notice, in particular, the two essentially flat sets of points in the figure—one from the local central England temperature up to roughly three hundred years and the other from an ocean core that is flat from scales of 100,000 years and longer. It turns out that the flatness is an artifact of the use of differences in the definition of fluctuations. We need something better.

Before continuing, let us recall the scaling laws that have been introduced up until now:

$$\text{Spectrum} \approx (\text{Frequency } \omega)^{-\beta} \approx (\text{Scale})^{\beta}$$

$$\text{Number of boxes} \approx (\text{Size L})^{-D} \approx (\text{Scale})^{-D}$$

$$\text{Probability} \approx (\text{Scale})^{C}$$

$$\text{Fluctuations} \approx (\text{Interval } \Delta t)^{H} \approx (\text{Scale})^{H},$$

where $\beta$ is the spectral exponent, $D$ is the fractal dimension of a set, $C$ is the codimension, and $H$ is the fluctuation exponent.[117] Fluctuations are potentially widely applicable and easy to interpret. A nonobvious problem with defining fluctuations as differences is that, on average, differences cannot decrease with increasing time intervals.[118] This is fine when $H$ is positive, but it means that it is useless whenever $H$ is negative (implying that $\Delta t^H$ decreases with increasing $\Delta t$).[119] But do regions of negative $H$ exist? One way to investigate this is to try to infer $H$ from the spectrum, which does not suffer from an analogous restriction; its exponent, $\beta$, can take any value. In this case there is an approximate formula[120] we can use: $\beta = 1 + 2H$. This formula implies that negative $H$ corresponds to $\beta < 1$, and a check on the spectrum (Fig. 2.3A) indicates that several regions

are indeed flat enough to imply negative *H*. How do we fix the problem with difference fluctuations and use fluctuations to estimate the correct *H* when it is negative?

It took a surprisingly long time to clarify this issue. To start with, the turbulence community was fond of difference fluctuations for which it had developed many convenient theoretical results. Turbulence theorists had been the first to use fluctuations as differences, effectively introducing the first fluctuation exponent *H* = 1/3 as the exponent in Kolmogorov's famous law[eee] a decade before Edwin Hurst (see Chapters 3 and 4). In classic turbulence, all the *H* values are positive, so that the restriction to positive *H* was not a problem. Later, in the wake of the nonlinear revolution during the 1980s, mathematicians invented an entire mathematics of fluctuations called "wavelets."[fff] Although, technically, difference fluctuations are indeed wavelets, mathematicians mock them, calling them the "poor man's wavelet," promoting the more sophisticated ones. Wavelets turned out to have many beautiful mathematical properties and often have colorful names such "Mexican Hat," "Hermitian Hat," or the "Cohen-Daubechies-Feauveau wavelet." For mathematicians, it was irrelevant that the corresponding physical interpretations were not evident. The mastery of wavelet mathematics also required a fair intellectual effort, and this limited the number of their adepts.

This was the situation in the 1990s, when scaling started to be applied to geo time series involving negative *H* (essentially, to any macroweather series, although at the time this was not at all clear). It fell upon statistical physicist Chung-Kang Peng to develop an *H* < 0 technique that he applied to biological series: the detrended fluctuation analysis (DFA) method.[121] Also at this time, another part of the scaling community (including my colleagues and I) were focusing on multifractality and intermittency, and these issues didn't involve negative *H*, so the problem was ignored. Over the following nearly two decades, there were thus several more or less independent strands of scaling analysis, each with its own mathematical formalism and interpretations. The wavelet community dealt with fluctuations directly but were unconcerned about the simplicity of physical interpretations; the DFA community[ggg] wielded a somewhat complex method, but one that could be readily implemented numerically and didn't require much theoretical baggage[122]; and the turbulence community focused on multifractal intermittency. In the meantime, mainstream geoscientists continued to use spectral analysis—focusing on spectral peaks that supposedly (and often spuriously) represented quasi-oscillating processes—rather than on the scaling or on the

---

[eee] Kolmogorov's law was very close to Richardson's 4/3 law; the 4/3 was *H* + 1.

[fff] Although wavelets can be traced back to Alfred Haar (1909, ~~discussed later~~), they really took off starting in the early 1980s with the continuous wavelet transformation by Alex Grossman and Jean Morlet.

[ggg] At last count, Peng's original article had more than 2,000 citations, an astounding number for such a highly mathematical paper.

interpretation of the amplitudes of the spectra, most of which were treated as un-interesting background noise.

Ironically, the impasse was broken by the first wavelet, the one that Alfred Haar (1885–1933) had introduced in 1910, even before wavelet formalism had been invented.[123] The Haar fluctuation is beautiful for two reasons: the simplicity of its definition and calculation, and the simplicity of interpretation.[124] To determine the Haar fluctuation over a time interval $\Delta t$, one takes the average of the first half of the interval and subtracts the average of the second half (Fig. 2.17, bottom). That's it![hhh] As for the interpretation,[125] it is easy to show that when $H$ is positive, that it is (nearly) the same as a difference, whereas whenever $H$ is negative, we not only recover its correct value,[iii] but the fluctuation itself can be interpreted as an "anomaly."[jjj]

When Haar fluctuations are substituted for difference fluctuations, and using the climate series discussed in Chapter 1, we obtain the composite Figure 2.4A, which covers a range of about five orders of magnitude more than that covered by Mitchell[126] (Fig. 2.3A). One can clearly make out five distinct regions, each—with the exception of macroclimate[127]—of which is scaling over ranges of roughly 1,000 in timescale.[kkk] We clearly see the alternation of the sign of the slope ($H$) from positive (the weather regime to about 10 days; Fig. 2.3A, left) through the longer macroweather regime,[128] with decreasing fluctuations and negative $H$, to the increasing climate regime, decreasing macroclimate, and increasing megaclimate. The numbers on the vertical axis of Figure 2.4A all make perfect sense and give a precise idea of typical fluctuations at the corresponding timescale. For example, reading the numbers off the graph, we see that typical temperature fluctuations at intervals of one second are about 0.1°C; at ten days (midlatitudes), about 10°C. These are the increasing part of the curve at the left, and they indicate that typical changes at these scales are ±0.05°C, ±5°C. At ten years—on the decreasing part of the curve, we again have fluctuations with amplitudes of about 1°C—in this case, indicating that typical consecutive ten-year averages differ by this amount. Continuing to longer time periods, we find that typical Ice Age variations (with half periods $\Delta t$ about 50,000 years) are roughly 6°C, and at 100 million years are about 12°C. Similarly, in space (Fig. 2.4B), kilometer-to-kilometer changes are on the order of 0.2°C, with typical changes at 100 km of 1°C. In contrast, we saw by direct comparison (Fig. 2.3C) that the quantitative implications of Mitchell's spectrum are quite implausible. For example, analysis shows that it implies that consecutive centennial average global temperatures would vary by only a hundredth

---

[hhh] I can recall a comment of a referee of an article in which I explained the Haar fluctuation using the same words. Expecting a complicated wavelet expression, he complained that he didn't understand the words and instead wanted an equation!

[iii] The Haar fluctuation is only useful for $H$ in the range –1 to 1, but this turns out to cover almost all of the series that are encountered in geoscience.

[jjj] In this context, an anomaly is simply the average over a segment length $\Delta t$ of the series after its long-term average has been removed.

[kkk] The weather regime apparently continues for a further factor of 100 down to millisecond dissipation scales.

of[III] a degree centigrade, that million-year averages would vary by a few dozen microdegrees.

But how reliable is Figure 2.4A? Would it have been different if it had been compiled at a different epoch? Box 2.1 explored its robustness by imagining that our hominid ancestor Lucy was a climatologist, and argued that her version of the figure would have been only a little—but not much—different. With only one or two caveats, the figure is likely to be representative of the past 800,000 years, and even of the past 65 million years.

Figure 2.4C shows how fluctuations in the different regimes look, confirming the canceling ($H < 0$) and wandering ($H > 0$) behaviors. By comparing the bottom two at 0.017-second and hourly resolutions, we see visual confirmation that although they have the same $H$ values, the characters are somewhat different, with the bottom being more "spikey"—a consequence of intermittency (Box 2.2). Last, Figure 2.4B is the spatial counterpart of Figure 2.3A, showing how data averaged over different time intervals (the different regimes) fluctuates in space. It shows excellent scaling (straight lines in the figure) including through the mesoscale discussed earlier (1–100 km).

By comparing the difference and Haar fluctuations (Fig. 2.18 and Fig. 2.4A, respectively), we can now comprehend the limitations of the difference-based analysis (Fig. 2.18), and understand why macroweather was not clearly discerned until so recently. As expected, the increasing parts of the two figures are quite similar: The flat parts of Figure 2.18 do indeed correspond to negative $H$—both macroweather and macroclimate. The remaining apparent divergence between the differences and Haar fluctuations (Figs. 2.18 and 2.4A) has to do with the difference between local and globally averaged temperatures, and the difference between industrial and preindustrial temperatures (resulting from anthropogenic warming). I defer discussion of this until Chapter 5 and Box 5.1.

## Notes

1. Recent historical research indicates that Robert Hooke may, in fact, have preceded van Leeuwenhoek, but the latter is usually credited with the discovery.

2. Boeke, K. *Cosmic View: The Universe in Forty Jumps*. (John Day, 1957).

3. Hutchins, R. M. & Adler, M. J. *Gateway to the Great Books*. (Encyclopædia Britannica, 1963).

4. Produced by the National Film Board of Canada.

5. By Charles and Ray Eames.

---

[III] The missing quadrillion refers to the spectrum—the amplitude of the fluctuations squared. Mitchell's error in the fluctuation amplitude is only by a factor of about a few million. The subsecond thermistor data analyzed in the lower left of Figure 2.4A extends Mitchell's range of timescales by a further factor of 100,000 (from two hours down to 0.03 seconds). If Mitchell had extended his roughly flat background spectrum to the corresponding high frequencies, the error would have compounded by a further factor of 100 million or so.

6. The re-release had the subtitle "A Film Dealing with the Relative Size of Things in the Universe and the Effect of Adding Another Zero" and was narrated by Philip Morrison. More recently, the similar *Cosmic Voyage* (1996) appeared in IMAX format.

7. Sharf, C. *The Zoomable Universe*. (Scientific American, 2017).

8. Cooley, J. W. & Tukey, J. W. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* **19** (90), 297–301 (1965).

9. The averaging is over a collection (ensemble) of similar processes. If we only have a single time series, then technically the result is a "periodogram."

10. Because the darker curve is the result of analyzing a single series, there is no averaging over a collection of series so that, technically, this is a periodogram. Rather, the spectrum is represented by the smooth gray curves at the bottom, which are obtained by averaging over many periodograms (one of these by averaging over 5,000; the other is the theoretical result for averaging over an infinite number).

11. The spectrum is actually the ensemble average of the squares of the absolute amplitudes. It was "windowed" to avoid spurious "spectral leakage" that could smear out the spectrum artificially.

12. For other examples, see Lovejoy, S. The spectra, intermittency and extremes of weather, macroweather and climate. *Nat. Sci. Rep.* **8**, 1–13. (2018).

13. The negative sign is used by convention so that, in typical situations, β is positive.

14. In the more usual case of continuous spectra, we have $\beta = 1 + 2H$, with corrections when intermittency is important (see ~~Box~~ 2.3).

15. I admit that, to make my point, I made 5,000 simulations of the multifractal process in Figure 1.3 and then searched through the first fifty to find the one with the most striking variation. But this was by no means the most extreme of the 5,000, and if the statistics had been from the bell curve, then the extreme point in the spectrum in Figure 2.3 would have corresponded to a probability of 1 in 10 trillion, so that my slight cheating in the selection process would still have been extremely unlikely to have caused the result!

16. According to standard assumptions (which this example shows are inappropriate), successive frequencies should be statistically independent of each other; there should be no relationship between them.

17. The standard statistical null hypothesis used for detecting extreme spectral peaks is based on the bell curve and this is inappropriate. Several examples of spurious periodicities emerged in the 1990s as a consequence of the use of sophisticated analysis techniques such as the multitaper method and the singular spectral analysis method combined with such inappropriate probability assumptions, see Chapter 4.

The search for significant oscillations continues. A recent climate example of a spectral analysis with a spectrum similar to Figure 2.2B (albeit with much smaller spikes, but still claimed to be significant) can be found in Galloway, J., Wigston, A., Patterson, R.T., Swindles, G.T., Reinhardt, E., & Roe, H.M. Climate change and decadal to centennial-scale periodicities recorded in a late Holocene NE Pacific marine record: Examining the role of solar forcing. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **386**, 669–689 (2013).

18. I could also mention the contribution of "Box-Jenkins" techniques (1970) to bolstering scalebound blinkers. Originally, they were engineering tools for analyzing and modeling stochastic processes based on the a priori scalebound assumption that the correlations decayed in an exponential manner. This contributed especially to scalebound thinking in precipitation and hydrology. See, for example, the following influential

publications: Zawadzki, I. Statistical properties of precipitation patterns. *J. Appl. Meteorol.* **12**, 469–472 (1973); Bras, R. L. & Rodriguez-Iturbe, I. Rainfall generation: A nonstationary time varying multidimensional model. *Water Resourc. Res.* **12**, 450–456 (1976); Bras, R. L. & Rodriguez-Iturbe, I. *Random Functions and Hydrology*. (Addison-Wesley Publishing, 1985).

19. Mitchell, J. M. An overview of climatic variability and its causal mechanisms. *Quaternary Res.* **6**, 481–493 (1976).

20. Dijkstra, H. & Ghil, M. Low frequency variability of the large scale ocean circulations: A dynamical systems approach. *Rev. Geophys.* **43** (RG3002), 1–38 (2005). Fraedrich, K., Blender, R., & Zhu, X. Continuum climate variability: Long-term memory, scaling, and 1/f-noise. *Intl. J. Modern Phys. B* **23**, 5403–5416 (2009). Dijkstra, H. *Nonlinear Climate Dynamics*. (Cambridge University Press, 2013).

21. Mitchell, J. M. An overview of climatic variability and its causal mechanisms. *Quaternary Res.* **6**, 481–493 (1976).

22. Mitchell also distinguished between regional and global spectra, with the latter being the lower curve contained completely within the gray area near the bottom.

23. Adapted from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015). The original is Mitchell, J. M. An overview of climatic variability and its causal mechanisms. *Quaternary Res.* **6**, 481–493 (1976).

24. Reproduced from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

25. Only the sloping straight-line background is quantitative. This line implies very low variability. For example, on the graph, we see that successive century-averaged temperatures would differ by only about one hundredth of a degree, and that successive million year averages by less than 100 microdegrees.

26. Reproduced from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

27. Figure 2.18 covers a range from 5 years to 1,000,000 years, but the full article had other analyses down to three minutes in scale. See Lovejoy, S. & Schertzer, D. Scale invariance in climatological temperatures and the local spectral plateau. *Annal. Geophys.* **4B**, 401–410 (1986).

28. In Figures 2.3C and 2.4A, we plot the same information but in real space and find that the root mean square fluctuations at $5.53 \times 10^8$ years are $\approx \pm 10$K, so that extrapolating Gaussian white noise over the range implies a value of $\approx 10^{-6}$K. In other words, it is in error by a factor of $\approx 10^7$.

29. If we attempt to extend Mitchell's picture to the dissipation scales (at frequencies a million times greater, corresponding to millisecond variability), the spectral range would increase by an additional factor of a billion.

30. These are root mean square Haar fluctuations: the square root of the average of the squares of the fluctuations.

31. It has the same information content. From the point of view of spectral analysis, the Haar wavelet is just a filter.

32. Adapted from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

33. Reproduced from Lovejoy, S. The spectra, intermittency and extremes of weather, macroweather and climate. *Nat. Sci. Rep.* **8**, 1–13 (2018).

34. From top to bottom, the ranges used for normalizing are 10.1, 4.59, and 1.61 (Veizer, Zachos, and Huybers, respectively, all $\delta_{18}$ and 6.87°C, 2.50°C, 25°C, and 0.8°C (Epica, Berkeley, Lander, thermistor).

35. Adapted from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

36. Neilsen, T., Rypdal, K., & Fredriksen, H.-B. Are there multiple scaling regimes in Holocene temperature records? *Earth Syst. Dynam.* **7**, 419–439 (2016).

37. Nielsen et al. and others mistakenly couch the issue as one of statistical stationarity, claiming that Earth's temperature is a nonstationary process [i.e., one with statistical properties that depend on (absolute) time (not simply on *timescale*)]. But as discussed in Section 4.1, statistical stationarity is a property of an infinite ensemble, and because all data are finite, it is only a property of models. We only have one realization of Earth's evolution and so we cannot conclude anything about its stationarity or nonstationarity simply by analyzing data. The more interesting question is whether a stationary or nonstationary stochastic model is physically the most justified. Which of the two best fits the observations? To justify nonstationary *models*, appeals are often made to the fact that, mathematically, if the variability continued to grow down to arbitrarily low frequencies (infinitely long times), a nonstationary *model* would be necessary. Yet this observation isn't very helpful because the longest time period that is accessible empirically (550 million years) is still not infinite in length, and it is a trivial matter to make a finite-length stationary stochastic process that reproduces any given spectrum.

38. Galeotti, S., DeConto, R., Naish, T., Stocchi, P., Florindo, F., Pagani, M., Barrett, P., Bohaty, S. M., Lanci, L., Pollard, D., Sandroni, S., Talarico, F.M., & Zachos, J.C Antarctic ice sheet variability across the Eocene–Oligocene boundary climate transition. *Science* doi:10.1126/science.aab0669 (2016).

39. Lucy's fluctuations would be about her own Pliocene climate, but their statistical variability would be the same. This is analogous to today's GCMs, each of which have their own climates, but nevertheless have very similar statistical variability about these climates (see Chapters 5 and 7).

40. Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

41. The amplitudes of the variability did change somewhat from one 10-million-year section to another, but the basic megaclimate behavior was robust.

42. Even here there are exceptions. We mentioned the diurnal and annual cycles, but astronomical cycles presumably play a role in driving climate macroclimate regimes (Box 5.4).

43. Cited in Shackleton, N. J. & Imbrie, J. The $\delta^{18}$O spectrum of oceanic deep water over a five-decade band. *Climat. Change* **16**, 217–230 (1990); quote, p. 228.

44. Pelletier, J. D. The power spectral density of atmospheric temperature from scales of $10^{-2}$ to $10^6$ yr. *EPSL* **158**, 157–164 (1998). Huybers, P. & Curry, W. Links between annual, Milankovitch and continuum temperature variability. *Nature* **441**, 329–332 (2006).

45. Wunsch, C. The spectral energy description of climate change including the 100 ky energy. *Climate Dynam.* **20**, 353–363 (2003).

46. They have been updated more recently with the help of stochastics: the "random dynamical systems" approach. Chekroun, M. D., Simonnet, E., & Ghil, M. Stochastic climate dynamics: Random attractors and time-dependent invariant measures. *Physica D* **240**, 1685–1700 (2010). Dijkstra, H. *Nonlinear Climate Dynamics*. (Cambridge University Press, 2013).

47. Lovejoy, S. & Schertzer, D. In: *Chaos, Fractals and Models 96* (eds. F. M. Guindani & G. Salvadori), pp. 38–52. (Italian University Press, 1998).

48. In the more recent "random dynamical systems" approach, the driving noise may be viewed as an expression of large numbers of degrees of freedom. This interpretation is only justified if there is a significant scale break between the scales of the noise and of the explicitly modeled dynamics. This framework is not trivially compatible with scaling spectra.

49. Orlanski, I. A rational subdivision of scales for atmospheric processes. *Bull. Amer. Metereol. Soc.* **56**, 527–530 (1975).

50. Schertzer, D., Lovejoy, S., Schmitt, F., Chigirinskaya, Y., & Marsan, D. Multifractal cascade dynamics and turbulent intermittency. *Fractals* **5**, 427–471 (1997).

51. Reproduced with permission from Schertzer, D., Lovejoy, S., Schmitt, F., Chigirinskaya, Y., & Marsan, D. Multifractal cascade dynamics and turbulent intermittency. *Fractals* **5**, 427–471 (1997).

52. The term "zoo" is borrowed from statistical physics, where it is sometimes used to describe the bewildering variety of shapes that occur, for example, in the study of phase transitions.

53. At the time, NWPs and GCMs were much too small to allow for proper statistical scaling analyses of their outputs, and the reigning turbulence theory turned out to be seriously unrealistic (Chapter 3). Even today, the fact that NWPs and GCMs should be—and are—scaling is practically unknown. Sometimes the absence of a scale break is even seen not as a model strength, but as a model *deficiency*! See the discussion in Chapter 4.

54. Mandelbrot, B. B. *Fractals, Form, Chance and Dimension*. (Freeman, 1977).

55. Mandelbrot, B. B. *The Fractal Geometry of Nature*. (Freeman, 1982).

56. Thompson, D. W. *On Growth and Form*. (Cambridge University Press, 1917).

57. The mathematical issue is their singular small-scale nature. The basic multifractal process is a cascade that does not converge to mathematical points, but only converges in the neighborhood of points. This precludes them from being represented as a geometric set of points. Mathematically, they are "generalized functions," such as the Dirac delta functions used in physics.

58. Mandelbrot, B. Scalebound or scaling shapes: A useful distinction in the visual arts and in the natural sciences. *Leonardo* **14**, 43–47 (1981); quote, p. 45.

59. My approach to rainfall modeling followed the method that Mandelbrot had used to make cloud and mountain models in his book, except that I used a variant that was far more variable [based on "Levy" distributions rather than the bell curve (see Box 3.1)].

60. The *Fractal Geometry of Nature* contained simulations of cloud "surfaces" based on turbulence theory (the Corrsin-Obukhov law), but they were not related directly to meteorology.

61. The FSP model was an extension and improvement over the Levy fault model that I had developed during my PhD thesis, but was nevertheless still mono-, not multi-, fractal. Lovejoy, S. & Mandelbrot, B. B. Fractal properties of rain and a fractal model. *Tellus* **37 A**, 209 (1985).

62. Richardson, L. F. Atmospheric diffusion shown on a distance–neighbour graph. *Proc. Roy. Soc.* **A110**, 709–737 (1926).

63. Adapted from Lovejoy, S. Area perimeter relations for rain and cloud areas. *Science* **187**, 1035–1037 (1982).

64. The use of the term "fractal object" is frequent in this context, but is sometimes confusingly vague. Mathematically, fractals are sets of points that have a scale symmetry

(scaling) whereas multifractals are fields—such as the temperature in the atmosphere—that have values at each point in space and in time and that are scaling.

65. Dewdney, A. K. A computer microscope zooms in for a close look at the most complicated object in mathematics. *Scientific American* **August**, 16–24 (1985); quote, p. 18.

66. To get an interesting result, the constant should be a complex number (i.e., one that involves the square root of –1).

67. Reproduced with permission from Peitgen, H.O. & Richter, P.H. *The Beauty of Fractals.* (Oxford University Press, 1986).

68. Cantor, G. Sur les ensembles infinis et linéaires de points. *Acta Math.* **2**, 381–408 (1883). It was apparently discovered a bit earlier by H. J. S. Smith in 1874.

69. Adapted from Lovejoy, S., Schertzer, D., & Ladoy, P. Fractal characterisation of inhomogeneous measuring networks. *Nature* **319**, 43–44 (1986).

70. Adapted from Lovejoy, S., Schertzer, D., & Ladoy, P. Fractal characterisation of inhomogeneous measuring networks. *Nature* **319**, 43–44 (1986).

71. Here—as almost always—the box-counting dimension is the same as the Hausdorff dimension that is sometimes used in this context.

72. In the upper left construction, we omitted the lowest frequency. If we don't remove the lowest frequency, then the result is only approximately self-affine. However, the construction *mechanism* itself is nevertheless self-affine.

73. The familiar geometric shapes studied by Euclid—points, lines, planes, and volumes—have "topological dimensions" 0, 1, 2, and 3. Topological dimensions have to do with the connectedness properties of a set. For fractal sets, the fractal dimension and the topological dimension are generally different.

74. Of course, the line in the previous example is scale invariant with $D = 1$, so according to this definition it is also a fractal. However, we generally reserve the term "fractal" for less trivial scale-invariant sets, usually with a fractal dimension greater than the topological dimension.

75. Sierpiński, W. Sur une courbe cantorienne qui contient une image biunivoque et continue de toute courbe donnée. *C. r. hebd. Seanc. Acad. Sci. Paris* **162**, 629–632 (1916). (in French).

76. The negative sign indicates that smaller segments are more numerous. Indeed, $N \times L$ = total length (in $d = 1$), $N \times L^2$ = total area (in $d = 2$), and so on.

77. This technique actually estimates the "correlation dimension" of the set. If, instead, one centers circles at points chosen at random on Earth's surface (instead of only on stations), then one obtains the box-counting dimension discussed earlier. It turns out that, in general, the two are slightly different. The density of points is an example of a multifractal measure. Indeed, one can introduce an infinite hierarchy of different exponents associated with the density of points.

78. The rule $L^D$ for the number of stations in a circle is a consequence of the fact that the number of boxes at scale $L$ decreases with $L$ as $L^{-D}$, because, on average, the number of points ~~per box~~ is independent of $L$: $L^{-D} \times L^D$ = constant.

79. The geographical locations of the stations in Figure 2.8C were only specified to the nearest kilometer, so it is possible that the curve actually extends to even smaller scales. For large $L$, it is valid up to several thousand kilometers, which is about as much as is theoretically possible given that were only 9,962 stations.

80. Both in space and—as a result of data outages and ship movements—also in time, the fractal dimensions and codimensions are nearly the same as for the meteorological network.

Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

81. The techniques for filling the "holes," such as "Kriging," typically also make scalebound assumptions (exponential decorrelations and the like).

82. When estimating global temperatures over scales up to decades, the problem of putting meteorological station data onto grids does indeed dominate the other errors (although this is not the same as dimensional resolution). The most important effect is from slight errors in the space–time resolution of the data: the "space–time reduction factor" (Box 5.1).

83. Koch, H. On a continuous curve without tangents constructible from elementary geometry. *Arkiv. Matematik Astronomi Fysik* **1**, 681–702 (1904).

84. Koch, H. On a continuous curve without tangents constructible from elementary geometry. *Arkiv. Matematik Astron. Fysik* **1**, 681–702 (1904).]. Reproduced from Welander, P. Studies on the general development of motion in a two dimensional, ideal fluid. *Tellus* **7**, 156 (1955).

85. Steinhaus, H. *Mathematical Snapshots*. (Oxford University Press, 1960).

86. Note that in the infinitely small limit, at each point, the Peano curve touches itself. This means that although it is a mapping of the line onto the plane, the mapping is not one-to-one.

87. Lovejoy, S. Area perimeter relations for rain and cloud areas. *Science* **187**, 1035–1037 (1982).

88. If the cloud area is itself a fractal set, then $P = A^{D/D_c}$, where $D_c < 2$ is the fractal dimension of the cloud image.

89. The area–perimeter relation was proposed in Mandelbrot, B. B. *Fractals, Form, Chance and Dimension*. (Freeman, 1977).

90. Richardson, L. F. Atmospheric diffusion shown on a distance–neighbour graph. *Proc. Roy. Soc.* **A110**, 709–737 (1926).

91. Adapted from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

92. Interestingly, a slightly more simple variant with less extreme behavior (now called the "*p* model") was first proposed by a geologist to describe the distribution of mineral ores. See de Wijs, H. J. Statistics of ore distribution: (2) Theory of binomial distribution applied to sampling and engineering problems. *Geol.Mijnbouw* **15**, 12–24 (1953). See also Schertzer, D. & Lovejoy, S. Elliptical turbulence in the atmosphere. In: *Fourth Symposium on Turbulent Shear Flows*, (ed. L.J.S. Bradbury, F. Durst, B.E. Launder, F.W. Schmidt, J.H. Whitelaw) pp. 11.1–11.8 (Karlshule, 1983); and Turbulence and chaotic phenomena in fluids, IUTAM. In: *On the Dimension of Atmospheric Motions* (ed. T. Tatsumi), pp. 505–512. (Elsevier Science Publishers B. V., 1984).

93. The name "α model" was given because—like all cascade processes—it generally displays extreme power law probabilities with "black swan" events and, at the time, the probability exponent was denoted by the symbol α. Today, it is often denoted "$q_D$" (see Box 3.1), because it depends on the dimension of space $D$ over which the process is averaged (and it should not be confused with the symbol α used in a similar context for Levy extremes).

94. Reproduced from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

95. Adapted from Schertzer, D. & Lovejoy, S. Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes. *J. Geophys. Res.* **92**, 9693–9714 (1987). ~~For the generalization of the (fractal) model to multifractals, see Box 2.2.~~

96. The expression "multifractal butterfly effect" was coined in Lovejoy, S. & Schertzer, D. In: *Chaos, Fractals and Models 96* (eds. F. M. Guindani & G. Salvadori), pp. 38–52. (Italian University Press, 1998). But, the original mathematics were discovered by Mandelbrot B. B. Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.* **62**, 331–350 (1974). This was generalized in Schertzer, D. & Lovejoy, S. In: *Turbulent Shear Flow* (eds. L. J. S. Bradbury, F. Durst, B. E. Launder, F. W. Schmidt, & J. H. Whitelaw), pp. 7–33. (Springer-Verlag, 1985).

97. This is different from the more familiar "butterfly effect" (Chapter 7), which refers to sensitive dependence on initial conditions so that a Brazilian butterfly might sufficiently destabilize a local—and then a larger and then even larger—region of the atmosphere, ultimately provoking a tornado in Texas, where there might not have otherwise been one.

98. Schertzer, D. & Lovejoy, S. Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes. *J. Geophys. Res.* **92**, 9693–9714 (1987).

99. As a special case, the singularities can, for example, follow the bell curve. In this special case, the field values (the activity levels) would follow "log-normal" distributions, a log-normal multifractal process.

100. Schertzer, D. & Lovejoy, S. Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes. *J. Geophys. Res.* **92**, 9693–9714 (1987).

101. For many more trace moment plots on satellite data, aircraft data, dropsondes, and NWP models, see Chapter 4 in Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

102. Mandelbrot, B. B. How long is the coastline of Britain? Statistical self-similarity and fractional dimension. *Science* **155**, 636–638 (1967).

103. Earlier, we saw that the length of the analogous fractal cloud perimeter varies as $L^{1-D}$, where $L$ is the length of the ruler and $D$ is the fractal dimension.

104. Richardson, L. F. Atmospheric diffusion shown on a distance–neighbour graph. *Proc. Roy. Soc.* **A110**, 709–737 (1926); quote, p. 717.

105. Richardson, L. F. *Weather Prediction by Numerical Process*. (Cambridge University Press, 1922). (Republished by Dover, 1965).

106. Richardson, L. F. Atmospheric diffusion shown on a distance–neighbour graph. *Proc. Roy. Soc.* **A110**, 709–737 (1926).

107. Richardson, L. F. The problem of contiguity: An appendix of statistics of deadly quarrels. *Gen Syst Yrbk.* **6**, 139–187 (1961).

108. It was actually a little more complicated than that. The key multifractal formula appeared independently in three publications in 1983—one dealing with turbulence and the other two in the field of deterministic chaos. Schertzer, D. & Lovejoy, S. On the dimension of atmospheric motions. In: *IUTAM Symposium on Turbulence and Chaotic Phenomena in Fluids* (ed. T. Tasumi), pp. 141–144. Grassberger, P. Generalized dimensions of strange attractors. *Phys. Rev. Lett.* **A 97**, 227 (1983). Hentschel, H. G. E. & Procaccia, I. The infinite number of generalized dimensions of fractals and strange attractors. *Physica D* **8**, 435–444 (1983).

Although the turbulent publication was admittedly only in a conference proceeding, the debate about the priority of discovery was soon overshadowed by Mandelbrot's claim to be

the "father of multifractals": Mandelbrot, B. B. Multifractals and fractals. *Physics Today* **39**, 11 (1986).

Soon after the initial discovery of multifractals, a major contribution was made by Parisi and Frisch, who were also the first to coin the term "multifractal": Parisi, G. & Frisch, U. A multifractal model of intermittency. In: *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics* (eds. M. Ghil, R. Benzi, & G. Parisi), pp. 84–88. (North Holland, 1985).

Recognizing the importance of multifractals, Mandelbrot subsequently spent a huge effort claiming its paternity. Ironically, Steven Wolfram, in his review of Mandelbrot's posthumous autobiography *The Fractalist* [Mandelbrot, B. B. *The Fractalist*. (First Vintage Books, 2011).], complained that Mandelbrot had "diluted" the fractals concept by insisting on multifractals: Wolfram, S. The Father of Fractals: The pioneering mathematician Benoit Mandelbrot spotted a common thread in complex shapes such as clouds, coastlines and Romanesco broccoli. *Wall Street Journal* (November 22, 2012).

109. It is actually an important assumption that the value of the temperature at a mathematical point does indeed converge to a well-defined value. In the more general multifractal case, the reduction of the temperature to a graph and hence to a set of points is already a kind of low-resolution approximation.

110. It will actually be an anisotropic ("self-affine") fractal set like the Weierstrass function (discussed in Chapter 3).

111. The *H* model is actually a variant of the pulse-in-pulse model and has some interesting properties, including (depending on the value of *H*) the divergence of high-order statistical moments. Lovejoy, S. & Mandelbrot, B. B. Fractal properties of rain and a fractal model. *Tellus* **37 A**, 209 (1985).

112. Reproduced with permission from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

113. Reproduced with permission from Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

114. Adapted from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. Cambridge University Press, 2013).

115. Manley, G. Central England temperatures: Monthly means 1659–1973. *Q. J. Roy. Meterol. Soc.* **100**, 389–495 (1974).

116. Reproduced from Lovejoy, S. & Schertzer, D. Scale invariance in climatological temperatures and the local spectral plateau. *Annal. Geophys.* **4B**, 401–410 (1986).

117. The symbol *H* is used in honor of Edwin Hurst, who discovered the "Hurst effect": long-range memory associated with scaling in hydrology. He did this by examining ancient records of Nile flooding where he introduced the "rescaled range" statistic and its scaling exponent *H*. It turns out that the fluctuation exponent is, in general, *not* the same as Hurst's exponent and that they are only the same if the data follow the bell curve, which they do only rarely! This distinction has caused much confusion. Additional confusion occurs because many authors define *H* in terms of the running sum of the process. This increases the value of *H* by one, leading to an explosion of definitions and incommensurate emprical results. Hurst, H. E. Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.* **116**, 770–808 (1951).

118. This is true for any series that has correlations that decrease with increasing interval $\Delta t$, as physically relevant series always do.

119. When $H < 0$, the difference-based fluctuations will simply give a roughly constant result—the flat parts of Figure 2.18. Because neither the dimension $D$ nor the codimension $C$ can be negative, this problem does not arise for them.

120. Valid if we ignore intermittency; otherwise, there are "intermittency corrections" (Box 2.2).

121. In retrospect, the key innovation was simply that the method started by taking the running sum of the original series, effectively adding one to the value of $H$. As long as the original $H$ was greater than $-1$, the new series thus had a positive $H$, allowing the usual differences and differencelike fluctuations to be used. Peng, C.- K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., & Goldberger, A. L. Mosaic organisation of DNA nucleotides. *Phys. Rev. E* **49**, 1685–1689 (1994).

122. In other words, the DFA method estimates fluctuations by the standard deviation of the residuals of a polynomial fit to the running sum of the series. The interpretation is so obscure that typical plots do not bother even to use units for the fluctuation amplitudes, thus throwing away much of the information.

123. Haar, A. Zur Theorie des orthogonalen Funktionsysteme. *Math. Annal.* **69**, 331–371 (1910).

124. Lovejoy, S. & Schertzer, D. Haar wavelets, fluctuations and structure functions: Convenient choices for geophysics. *Nonlinear Proc. Geophys.* **19**, 1–14, (2012).

125. The correspondence Haar = difference with $H$ positive and Haar = anomaly with $H$ negative is not exact numerically. It is usual to multiply the "raw" Haar fluctuation by a factor of 2 to make the correspondence closer. This has been done in Figure 2.18 and elsewhere in this book.

126. To avoid cluttering the figure, I did not show the curve for the globally averaged temperature. This is examined extensively in Chapter 5.

127. It is not clear that this is, indeed, a true scaling regime (see Chapter 5).

128. The figure shows macroweather at a high-latitude location (Greenland). Over the ocean or averaged globally, it is less steep; $H$ is less negative (see Chapter 5).

{ 3 }

# How big is a cloud?

## 3.1 Fractals: Where's the physics?

### 3.1.1 SYMMETRIES

We have discussed two extreme views of atmospheric variability: the scalebound view, in which every factor of 10 or so involves some new mechanism or law; and the opposing self-similar scaling view, in which zooming gives us something essentially the same—a single mechanism or law that could hold over ranges of thousands or more.

By considering time series and spatial transects, we saw that, over various ranges of scale in space and in time, atmospheric scaling seemed to work quite well. We looked at a complication: Interesting geophysical quantities are not simply black or white (geometric sets of points), but have gray shades; they have numerical values everywhere. To deal with the associated extreme variability and intermittency, we saw that we had to go beyond fractal sets to multifractal fields (Box 2.2). Understanding multifractals turned out to be important. Failure to appreciate their importance led to numerous deleterious consequences.[1] In this chapter, I want to consider something quite different: the morphologies of shapes in two or three dimensions.

Up until now, we have identified scaling with self-similarity, the property that, following a usual isotropic zoom (one that is the same in all directions), small parts resemble the whole in some way. Yet in Chapter 1 (Fig. 1.8A, B), we saw that zooming into lidar vertical sections uncovered morphologies that changed with scale. As we zoomed into flat, stratified layers, structures became visibly more "roundish" (compare Fig. 1.8A with Fig. 1.8B). Vertical sections are thus not self-similar. Their degree of stratification—anisotropy—changes systematically with scale.

But the vertical isn't the only place where self-similarity is unrealistic. Although it is not as obvious, the same difficulty arises if we zoom into clouds in the horizontal. We criticized Orlanski's powers of ten classification as being arbitrary and in contradiction with the scaling area–perimeter relation, but Orlanski was only trying to update an older phenomenological classification scheme, some of which

predated the twentieth century. Surely this contained some truth. If we viewed Earth from outer space and zoomed in, wouldn't cloud morphologies change systematically with the degree of magnification? We might spot a hurricane 1,000 kilometers across, but if we zoom in, we surely wouldn't see one 1,000 meters wide!

By the mid 1980s, the initial excitement about scaling ideas had begun to dissipate and the wider scientific community was starting to dismiss scaling and fractals as "just geometry," and an implausibly self-similar one at that. While Mandelbrot's group was busily promoting the geometry of fractal sets, Schertzer and I were developing tools for analyzing and modeling multifractal processes in which geometry played only a secondary role.[a] Caught between a rock and a hard place, by the end of the 1980s, our articles were being rejected by the fractalists for being insufficiently *fractal* while being shunned concomitantly by the mainstream for being *too* geometrical—in my case, guilt by association. At one point—only weeks after the rejection of a manuscript from a special journal issue on fractals[b]—I received a particularly negative (and costly) evaluation of a research grant application. The irate reviewer—apparently a geographer oblivious to the debate and unaware of the irony—exhorted me to "do physics, not geometry."

The question on everyone's mind was: Fractals: Where's the physics? This was the title of an influential article by statistical physicist Leo Kadanoff (1937–2015).[2,c] Kadanoff had known about scaling ever since the mid 1960s, when he had contributed to the development of (scaling) renormalization group theory that accounted for phase transitions in materials.[d] However, in statistical physics, there was also a diverse gamut of scaling phenomena, such as low-frequency circuit noise, referred to collectively as "1/f" or "flicker" noises. These continued to defy understanding, thus fueling the idea that scaling was something mysterious that required new physics—not new geometry. Kadanoff's position was nuanced: While admitting that one had to go beyond geometry to physics,[e] he

---

[a] As discussed in the introduction in Chapter 1, in an attempt to build our alternative approach, starting at McGill University in 1986 through to 1999 in Roscoff, France, we organized four workshops on nonlinear variability in geophysics, or NVAG.

[b] The paper had originally been accepted for publication in the special issue by the referees and regular editor, but was rejected at the last minute following the personal intervention of Mandelbrot, who was a guest editor of the special issue. The article was finally published in a later (regular) issue: Schertzer, D. & Lovejoy, S. Generalized scale invariance and multiplicative processes in the atmosphere. *Pageoph* **130**, 57–81 (1989).

[c] It was published in *Physics Today*. Its circulation was wide and included all the members of the American Physical Society.

[d] Although he shared the 1980 Wolff prize with Kenneth Wilson for this work, in 1982, it was Wilson who got the Nobel.

[e] He could hardly have thought otherwise. He had just coauthored an important article proposing the dimension formalism of multifractals. Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I., & Shraiman, B. Fractal measures and their singularities: The characterization of strange sets. *Phys. Rev. A* **33**, 1141–1151 (1986).

was nevertheless enthusiastic about fractals.[f] Although he did not say it, what was missing was a proper formulation of the scaling principle itself.

If scaling was limited to invariance under ordinary zooms (isotropic scale changes, self-similarity), then it would indeed have few real-world applications. But seemingly attractive scientific ideas and theories often fail to conform to reality. Most of the time, we simply abandon them and move on to something else, but occasionally it's possible to fix them by generalizing them to fit the facts. In its historical development, the law of conservation of energy was frequently faced with such failures, but rather than being abandoned, it was generalized successively from applying only to mechanical energy, to include heat, chemical, and electrical energies, and finally to mass-energy ($E = mc^2$).

Self-similar scaling is seductive because it is a symmetry. It says that when we zoom into an object, something is invariant, that a small part is similar in some way to the whole. To save the scaling idea, it must be generalized to account for systematic changes in morphologies with scale. We must formulate scale invariance as a symmetry principle more general than self-similarity.

It turns out that modern physics is literally based on symmetry principles. Ever since Emmy Noether (1882–1935) published her eponymous theorem[3] demonstrating the equivalence between symmetries and conservation laws, physicists have been obsessed with symmetry. To get an idea of the power of Noether's theorem, consider two examples: If the laws of physics are the same everywhere in space (a symmetry), then momentum is conserved; if the laws are the same at all times (another symmetry), then energy is conserved.[4] The conservation of momentum and energy are hugely important because they tell us that, although a system may be complicated—at least in some respects—we don't need to know the details. Two cars may crash and make a mess, but if we know the momentum and energy of each before the collision, then the subsequent trajectories of the mangled metal are tightly constrained. Similarly, for a self-similar fractal set, the fractal dimension is conserved under zooming—it is scale invariant—thus telling us precious information about the set's morphological structure. If it was associated with a dynamical system, then it would give information about the underlying physics. If scaling could be formulated as a general symmetry principle, then—thanks to Noether's theorem—the answer to Kadanoff's question would be: Scaling *is* the physics.

The thing about symmetry principles is that they represent a kind of maximal simplicity, and because "entities must not be multiplied beyond necessity," [g] physicists always assume that symmetries hold unless there is evidence for *symmetry breaking*. Again, consider the car crash, only this time one in which a gas tank explodes. Now, the sum of the mechanical energy of the cars before and after

[f] I met him two years later, when he participated enthusiastically as an invited speaker at the second NVAG workshop in Paris.

[g] This is a common statement of "Occam's razor," the principle that one should always adopt the simplest explanation of the facts, after William of Occam (1287–1347).

the collision would no longer be equal. The energy symmetry would be broken.[h] However, in the absence of information about an exploding gas tank, one assumes conservation.[i]

Thanks to Noether's theorem, conservation laws are not mysterious. They are simply another way of expressing a symmetry: Energy and momentum are conserved unless there is a specific source or sink that breaks the symmetry. The original question of "Why are momentum and energy conserved?" has thus been transformed into "Why are the laws of physics the same everywhere and at all times?" The answer to this question is: "Because we haven't found any reason to doubt it." In complex systems with variability over wide ranges of scale, one therefore *expects* scaling. It is expected to hold unless processes can be identified that act preferentially and strongly enough at a specific scale that could break it. This turns the table on scalebound thinking. If we can explain the atmosphere's structure in a scaling manner, then this is the simplest explanation and should, a priori, be adopted. The onus must be on the scalebound approach to demonstrate the inadequacy of scaling and the need to replace the hypothesis of a unique wide-range scaling regime by (potentially numerous) distinct scalebound mechanisms.

### 3.1.2  ZOOMING AND SQUASHING

In science, solutions to problems are often unexpected. In fall 1981, I had begun participating in the newly created turbulence group[j] at the Météorologie Nationale and I became acutely aware that, for the theorists, the wide-range scaling displayed by the area–perimeter relation (Fig. 2.6) was highly problematic.[k] The difficulty was the "mesoscale," the range from roughly 1 to ≈100 km, where a "dimensional transition"[5] supposedly divided the quite different small-scale, self-similar (isotropic) three-dimensional turbulence from the large-scale, self-similar two-dimensional turbulence (see Chapter 4).

It was therefore no accident that in spring 1982, my first joint project with Daniel Schertzer was to find out what was happening with the mesoscale. The service had just finished an experiment in Landes involving eighty high-quality, research-grade weather balloons—"radiosondes," with vertical resolutions of 50 m, that were launched at regular three-hour intervals. We therefore chose to exploit this unique data to investigate the vertical structure that was plausibly linked

---

[h] Of course the energy in the universe would still be conserved because the explosion was powered by chemical energy in the fuel tank, but the mechanical energy in the collision would not be conserved.

[i] Even with the explosion, momentum would be conserved.

[j] The group had been created at the Météorologie Nationale in the wake of the historic Left victory in the May and June elections that loosened bureaucratic restrictions on fundamental research in the service. It was headed by Jean-Claude André, with Daniel Schertzer and Gerard Thierry. The high-level civil servants viewed it as a concession to the political Left and made sure that it didn't survive the turn to the Right over the following couple of years.

[k] The work had just been accepted for publication is *Science*. Lovejoy, S. Area perimeter relations for rain and cloud areas. *Science* **187**, 1035–1037 (1982).

theoretically to the hypothetical mesoscale dimensional transition and that was more experimentally accessible.

We had recently become aware of Mandelbrot's important article (published in 1974),[6] "Intermittent Turbulence in Self-similar Cascades: Divergence of High Moments and Dimension of the Carrier," in which he proposed two major ideas: that turbulence was fractal (although he hadn't quite coined the term yet) and also the "divergence of high-order moments." Having already been convinced of the fractality, it was the latter that caught our attention. The "divergence of high-order moments" is another way of saying that the extreme events generated by turbulent cascades are far stronger than anything possible under classic bell curve statistics, so that—in turbulence—one expects black swan extremes (Box 3.1).

---

BOX 3.1  Multifractal butterflies, gray and black swans, extreme events, and tipping points

We have discussed scaling in space and in time, the idea that as we zoom to smaller spatial scales—or, in time, to shorter durations—fluctuations or other features are "scaling" (i.e., they follow power laws, with exponents such as $H$ and $D$) that are scale invariant. It turns out that not only are the structures and fluctuations themselves scaling, but also their *probabilities* of occurrence are scaling. Indeed, processes that are scaling in space and/or time are usually also characterized by scaling probabilities, and many mechanisms—such as the multifractal butterfly effect (Box 2.2)—linking the two have been proposed. In this box we examine the consequences, which include the unexpectedly frequent occurrence of extreme events and the difficulty of distinguishing them from qualitatively different (potentially catastrophic) "tipping points."[l]

The basic phenomenon of power law, "fat" probability tails[m] has a long history, going back at least to Vilfredo Pareto (1848–1923), who used it as an empirical distribution in economics, notably for describing income distribution (the "Pareto law"). In 1925, Paul Levy (1886–1971) discovered a general additive mechanism for producing such distributions, but the resulting "Levy distributions"[7] were restricted to exponents less than 2. In the equation that appears later in this box, $q_D < 2$. Such low exponents imply incredibly strong extremes and, empirically,

---

[l] The spurious classification of an event as a tipping point (when in reality it is simply an extreme but "normal" outcome, the same unique mechanism that produces weak fluctuations) is a species phenomenological fallacy (see Section 3.3.7). Phenomena with power law probabilities often suffer from this. For example, following Mandelbrot [Mandelbrot, B. B. The variation of certain speculative prices. *J. Business* **36**, 394–419 (1963).], Bunde et al. have argued that stock market crashes may be the consequence of extreme power law price fluctuations. Bunde, A., Kropp, J., & Schellnhuber, H. J. *The Science of Disasters: Climate Disruptions, Heart Attacks and Market Crashes*. (Springer, 2002). Later, I suggest that Dansgaard-Oeschger events may be explained analogously. Box 4.1 gives an example of aircraft turbulence.

[m] "Tails" refers to the low-probability, extreme tail-end of the probability distribution. Probability distributions can be classified on the basis of the relative importance of their extremes. Standard (e.g., bell curve) Gaussian distributions are then "thin tailed"; power law distributions are "fat tailed." There is also an intermediate "long-tailed" category that occurs, for example, when the probability of the logarithm of the variable is Gaussian.

relatively few real-world processes have been found with $q_D < 2$ (although now that Levy distributions are becoming better known, this is changing!). In any case, even if Levy distributions have relatively few *direct* applications, they are nevertheless important as the generators of "universal multifractals"[8] (Box 2.2).

A decisive advance was Mandelbrot's (1974) discovery that multiplicative cascade processes can lead to power law probability tails but without Levy's restriction. He called the more general phenomenon the "divergence of high-order moments."[9] Schertzer and I (1987) generalized Mandelbrot's result and showed that, physically, it arose because of perturbations at small scales.[10] This is analogous to Lorenz's famous butterfly effect (Chapter 7), in which the atmosphere is so unstable that the flapping of a single butterfly's wings can drastically change the evolution of the atmosphere. Here, it is as though there is a flock of (spatially distributed) butterflies, hence we called it the "multifractal butterfly effect."[11]

At the same time, Bak et al. (1987)[12] discovered a somewhat different but also scaling process that generated power law probabilities that generated "self-organized criticality" (SOC). [n] It turns out that there are deep links between multifractal and SOC routes to power laws.[13] In the framework of multifractals, the power law probability distributions can be theorized as "multifractal phase transitions."[14] More recently (2009), another mechanism for generating power law probabilities using correlated additive and multiplicative noise was discovered by Sardeshmuhk and Sura.[15] Regardless of the mechanism, from an empirical point of view, as discussed later, it seems simpler to call the phenomena themselves "black swans."

To make the idea of scaling probabilities[o] more precise, consider the equation for the probability that a random variable (such as, for example, a temperature fluctuation), $\Delta T$, exceeds a fixed threshold $s$:

$$\Pr(\Delta T > s) \approx s^{-q_D}; s \gg 1,$$

where Pr indicates probability and $q_D$ is an exponent[16] that characterizes the extremes. The condition "$s \gg 1$" indicates that this law is only expected to hold for very extreme thresholds[17] $s$ (the "tail"). To get an idea of how extreme the extremes can be, consider an example with $q_D = 5$ [as has been estimated for both the wind (Fig. 3.2[18]) and the temperature (Fig. 3.1)]. Temperature fluctuations ten times larger than typical fluctuations would therefore occur only 100,000 times less frequently. In comparison, for the bell curve—with probabilities that decay exponentially quickly— the corresponding probability would be $10^{-23}$. In Chapter 6, I show how this can be used to test (and reject) statistically the hypothesis that the global warming that has occurred since the nineteenth century is no more than a giant natural fluctuation, or GNF.

---

[n] The basic SOC example is the "sandpile model," in which grains of sand are added one at a time to a pile of sand. From time to time, this provokes avalanches. The spatial pattern of the regions affected by avalanches are fractal sets, the magnitude of the avalanches (the number of grains in the pile that fall in an avalanche) have extreme power law probability distributions. The term "self-organized criticality" is justified because, at any moment, the shape of the sand pile is determined precisely by extreme events—the avalanches—so the pile is "self-organizing."

[o] This might be confusing. Up until now, I've used the term "scaling" to denote an invariance under zooming in time and/or space. Here it refers to zooming in probabilities.
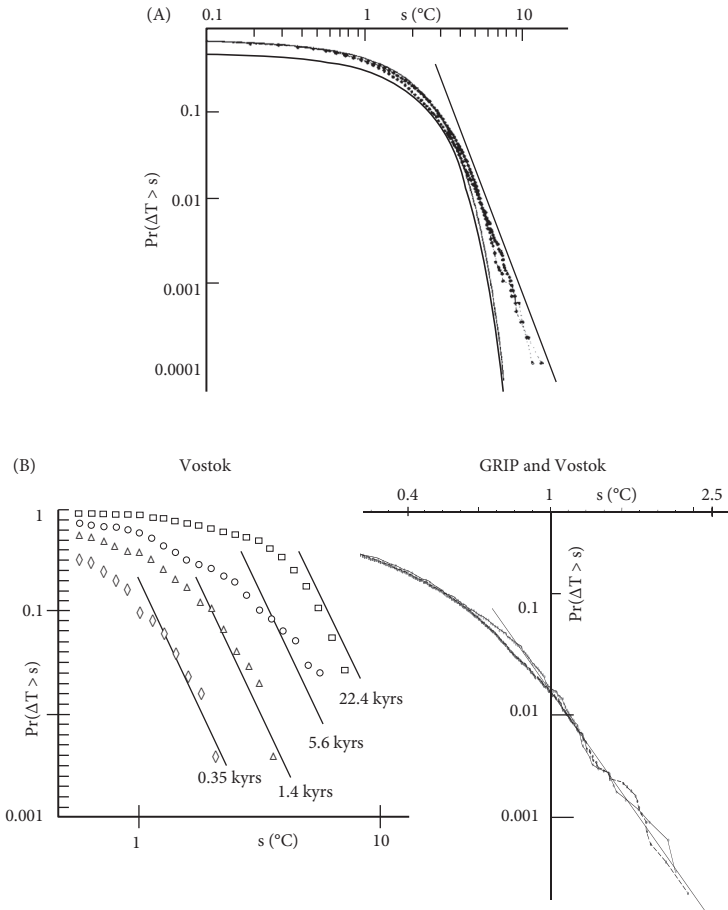
FIGURE 3.1 *(A) The probability distribution of daily temperature differences in daily mean temperatures from Macon, France, for the period 1949 to 1979 (10,957 days). Positive and negative differences are shown as separate curves. For reference, a best-fit Gaussian is shown as the solid curved line, indicating that the extreme fluctuations correspond to more than seven standard deviations. For a Gaussian, this has a probability of $10^{-20}$. The straight reference line (added) has a slope of $-q_D$, with $q_D = 5$.[19] (B) Probability distributions of paleotemperature changes for Vostok (left),[20] and a modern comparison of GRIP (Greenland Ice Core Project; Summit, Greenland) and Vostok (right).[21] The graphs differ not only because of the much improved sampling density of the more modern data, but also, the rightmost graph is at constant depth intervals (0.55 m for GRIP and 1 m for Vostok), which avoids issues of uncertain chronologies. In all cases, the straight reference lines indicate extreme $s^{-q_D}$ behavior with $q_D = 5$, where s is a temperature change. The reference lines in the left graph are spaced $H \log_{10} 4$ apart, with scaling exponent $H = 0.4$. (C) The total probability of random, absolute pre-1900 temperature differences exceeding a threshold s (measured in degrees Celsius), using three multiproxies to increase the sample size (the distributions are very similar in form for each of the multiproxies). To avoid excessive overlapping, the latter were compensated by multiplying by the lag $\Delta t$ (so that, for example, the far right curve must be divided by sixty-four to obtain the temperature fluctuations in degrees Celsius). The data are the pooled annual resolution multiproxies (Box 5.3) from 1500 to 1900. The (dashed) reference curves are Gaussians, with corresponding standard deviations, and the thin, straight lines starting at probabilities of 0.03 (Pr ≈ <3%) correspond to the bounding power laws $s^{-4}$ and $s^{-6}$.[22]*

(C)
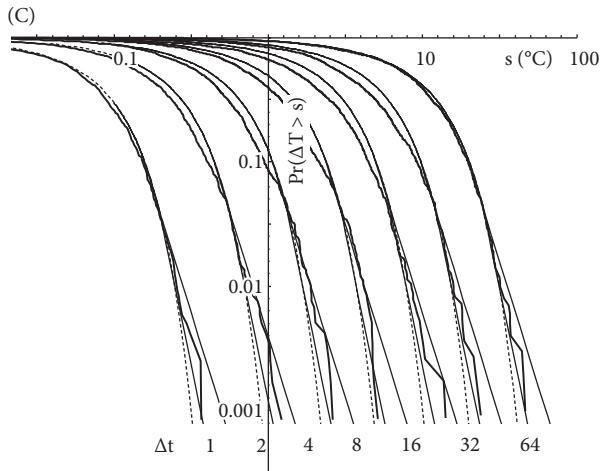
The "Gaussian," "normal," or "bell curve" introduced by Carl Friedrich Gauss (1777–1855) is known and loved by anyone who has taken a course in statistics. But why the obsession? The reason for its popularity is twofold. First, it can arise in fairly general circumstances—for example, as the result of the sum of many independent processes, including measurement errors.[23] Second, it is mathematically very convenient for calculations.

A basic property of the bell curve is that, for all practical purposes, it excludes extreme events. It doesn't do this simply by outlawing them, but rather by ensuring their probability of occurrence is tiny. We saw an example of this in Figure 1.6, where the extreme of the 8,192 temperature gradients was about seventeen times the "typical" gradient as quantified by the standard deviation (the spread about the mean value, see also Fig. 5.2). We noted that, had the temperature changes been generated by a standard bell curve process, the event would have had a probability of less than 1 in $10^{80}$.

Indeed, despite their ubiquity in the geosciences, very few geoprocesses follow the bell curve,[24] and extremes occur all too frequently—a fact that colleagues and I emphasized regularly starting in the 1980s.[25] Theoretically, the trouble is that, at best, the bell curve applies to additive processes—although not to the *H* model (Box 2.3)—and it fails completely for multiplicative processes[26] (Box 2.2).[27] One of Mandelbrot's first contributions—several years before he introduced fractal dimensions—was to apply strongly non-bell curve Levy distributions to financial series,[28] arguing that they accounted for extreme price changes, and later associating them with stock market crashes.[29]

This work in finance inspired experienced market trader Naseem Taleb to write an entire book popularizing the issue, introducing the term "black swan."[30] For a trader, Taleb has an unusual background (philosophy), and he put it to good work to argue that virtually all important historical events were unexpected. In his basic argument, the black swan issue is not one of probability; rather, a black swan event is something that is more than just improbable, it is *unthinkable*. The story of black

swans illustrates this nicely. Ancient Greek poets originally used black swans as metaphors for the impossible: It was a known fact that all swans were white. But then, in the seventeenth century, the British discovered Australia and the unthinkable happened: They discovered an entire black swan species (*Cygnus atratus*). For Taleb, a black swan event is thus one that is epistemologically extreme, one that could only be imagined with "out of the box"-type thinking.

In the second half of his book—inspired explicitly by Mandelbrot, whom he knew as a friend—Taleb introduces a variant—"gray swans"—to designate the extreme events associated with Levy distributions. He justifies this because extreme Levy events can be anticipated, but only on the basis of unconventional (non-Gaussian) theory. Hence their extremes were not truly black, only gray. The cascade extension from Levy to more general power law tails (i.e., with $q_D > 2$) should therefore rightly also be referred to as "gray swans," but this term never stuck. Hence, the term "black swans" is now often used for power law extremes in general.[p]

Let's now examine such black swan extremes with an example from the climate that will be relevant later. First, let's consider the probability distribution of daily temperature changes from a single station. Figure 3.1A shows the cumulative distributions of the temperature changes accumulating, starting from the largest value.[31] We see that for both positive and negative temperature changes, the distribution has far more extreme events than would be expected from the classic Gaussian distribution. Indeed, the extremes are at seven standard deviations, corresponding to events with Gaussian probabilities of less than $10^{-20}$. On the other hand, the data closely follow a power law with exponent $q_D \approx 5$. Moving to longer times (Fig. 3.1B), we see the same type of behavior, even in paleotemperatures. For the latter, modern data (far right) allow the tails to be examined more closely, yielding a more convincing result, again with $q_D \approx 5$. Note that taking differences over longer time scales shifts the tails by a constant factor corresponding to $H = 0.4$ (the left plots in Fig. 3.1B).

Last, to evaluate the statistics of natural temperature changes—and to avoid biases resulting from anthropogenic effects—consider global-scale preindustrial (1500–1900) temperatures. During the preindustrial period, global-scale temperatures can be estimated using "multiproxy" reconstructions (Section 5.7). As the name suggests, they combine data statistically from diverse sources ("proxies") that typically include tree rings, ice cores, and lake sediments to estimate temperature in the absence of instruments. Figure 3.1C shows the corresponding distributions for differences of 1, 2, 4, . . . , and 64 years. It also shows that scaling is a reasonable approximation to the tails of the (preindustrial) distributions of temperature changes. This result will be used later for estimating the probability that industrial warming is no more than a GNF.

Although the larger temperature changes at the bottom right of the plots in Figure 3.1C are extreme compared to anything that would be expected from the bell curve, they may nonetheless be small compared to changes that might occur if some qualitatively new nonlinear mechanism kicked in: a "tipping point." A tipping point denotes a change so strong as to be irreversible—a change that cannot be undone.

---

[p] Apologies to Taleb for this (slight) abuse of his excellent term!

In the context of climate change, a commonly invoked example is the release of methane ($CH_4$) from the thawing of Arctic permafrost caused by global warming. If enough $CH_4$ leaks to drive up temperatures significantly, it would accelerate itself through positive feedback, leading to a temperature increase far in excess of the black swans predicted on the basis of power law probabilities. This leads to a conundrum: With black swans already leading to exceptionally large extremes, how could we distinguish a "mere" black swan from a tipping point?

Before we started analyzing the Landes data, following the prevailing theory, we expected that the basic dimensional transition did indeed occur, but that it did not happen as originally theorized as a sharp break at a well-defined scale. Our hypothesis was that the transition was not seen in the area–perimeter relation because of its extreme variability. If the transition scale was variable enough from place to place and from time to time—for example, varying from hundreds of meters to hundreds of kilometers—then the horizontal statistics might effectively "blur it out" so that it would not be seen. The idea could be investigated by considering the atmospheric stability that should be connected to such a transition.

Ever since 1920, when the stability of stratified flows was studied by Richardson,[32] it has been known that stability depended on his eponymous number. The Richardson number is the (dimensionless) ratio of the buoyancy force gradient[33] to the square of the vertical shear[q]; classically, a stratified flow is unstable when this ratio exceeds some threshold conventionally taken to equal 1/4. Our idea was to use the radiosonde data to determine systematically the Richardson number over thicker and thicker layers. Because the Richardson number is notoriously variable,[34] we anticipated that vertical fluctuations would be extreme.

To capture the variability, we plotted the probability distribution of the Richardson number for layers increasing in thickness from 50 m to 3.2 km. Using log-log coordinates allowed us to investigate huge ranges of fluctuations and to test Mandelbrot's prediction about the black swans, which would appear as straight lines at the lower probability levels. The results were sensational (Fig. 3.2). First, the variability was so high that even the average Richardson number apparently diverged. This meant that when we tried to estimate it by averaging over more and more layers, the single layer (here, out of thousands) that happened to have the largest Richardson number had a value so large that it was roughly equal to the sum of all the others. No sensible average value could be obtained. Increasing the size of the sample to improve the estimate by averaging more values would simply give an even greater extreme with the same property. This huge variability is at the extreme end of the extreme and was itself a serious problem for the standard theory.

---

[q] The shear over a layer is the difference of the horizontal wind over the layer divided by the layer thickness.
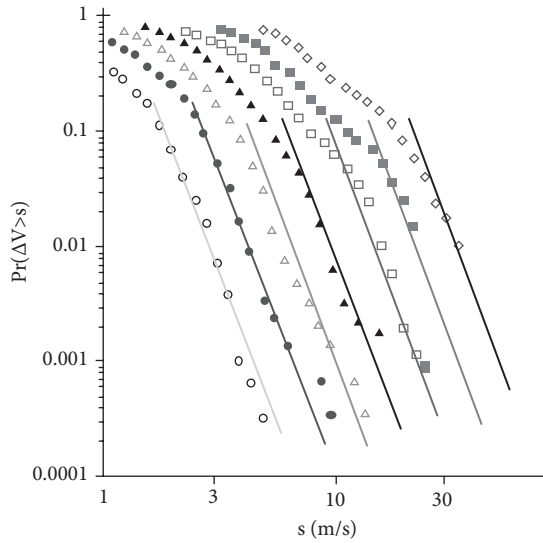
FIGURE 3.2 *The probability (Pr) that the change in horizontal wind ΔV over layers exceeds a given threshold s (measured in meters per second). Each set of points (and the reference line) is for layers of thickness increasing by factors of 2 from 50 m (left) to 3,200 m (right). For example, the extreme point on the farthest right-hand curve indicates that for layers 3,200 m thick, there is about a 1% chance (Pr = 0.01) that a difference of horizontal wind speed will exceed 30 m/s. The data were derived from eighty radiosondes from Landes, France. The reference lines have slopes of –5, corresponding to black swan extremes. The left–right separation of the lines corresponds to a vertical fluctuation exponent H = 3/5.[35]*

This is where things stood in summer 1982. While I ruminated on the implications of the probability tails, Schertzer went off for a few weeks to the Aspen Center of Physics. In Aspen, he realized there was something even more significant. When we compared the probabilities of wind, temperature, and humidity changes over layers of increasing thickness, we found they were nearly perfectly scaling. The curves in Figure 3.2 were parallel and equally spaced as the thickness of the layers were doubled. For the key wind speed, the spacing between the curves implies the fluctuation exponent had the value $H = 3/5$ (Fig. 3.2). In comparison, self-similar three-dimensional turbulence has $H = 1/3$, and self-similar two-dimensional turbulence has $H = 1$. We had not anticipated an in-between value!

Surveying the literature, we discovered that at the end of the 1960s, there had been two other experiments[36] that had found $H = 3/5$. One of the two had duly acknowledged that the experimental result was close to the value 3/5, which had been predicted theoretically ten years earlier by Alexander Obukhov (1918–1989) and by Ralph Bolgiano (1922–2002)[37] for buoyancy-driven turbulence (Chapter 4). However, both Bolgiano's and Obukhov's theories had shared the prevailing self-similar mindset: They both were isotropic; their fluctuation exponent 3/5 was supposed to hold in both the horizontal and vertical directions! However, experiments in the *horizontal* had instead found the value 1/3 associated with Kolmogorov's

theory of self-similar isotropic three-dimensional turbulence (Chapter 4). Ten years after they were proposed, Bolgiano and Obukhov's theories had already been forgotten.

At the same time that we were rediscovering the Bolgiano-Obukhov law in the vertical, the results of the first large-scale aircraft campaigns to study the wind in the horizontal direction started to appear.[38] To the theorists' embarrassment, these invariably found the value predicted for self-similar three-dimensional turbulence: $H$ was *not* the expected two-dimensional value[39] of 1. Instead, out to scales of at least hundreds of kilometers, they found the supposedly three-dimensional value 1/3.[r] Because the turbulence could not extend hundreds of kilometers into outer space, it could not possibly be isotropic three-dimensional turbulence. The inescapable conclusion was that we had to deal with a single scaling, but with an *anisotropic* regime with $H = 1/3$ in the horizontal and $H = 3/5$ in the vertical. This clear case of anisotropic scaling could be dealt with by combining zooms with squashing; the 23/9 D model was born.[40] Formulating the whole problem in terms of scaling symmetries soon allowed for the generalizations discussed next, which include rotation, waves, and more.[41]

## 3.2  Zooming, squashing, and rotating, and the emergence of scale

### 3.2.1  THE SCALING OF THE LAWS

The previous argument about scaling as a symmetry is at a nearly philosophical, metascientific level. This is fine for a system about which we don't have much detailed understanding. It is directly pertinent for many geophysics problems, such as understanding coastlines, topography, geomagnetism, geogravity, and earthquakes. But, for the weather and climate, we have a (fairly) well-defined set of equations[42] so that we can (and must) do much better than simply expound upon general considerations of scientific methodology. The equations are essentially those that were first written by Richardson, and they form the basis of numerical weather and climate modeling. To argue convincingly for scaling in the atmosphere, we must show that the equations themselves have the necessary scaling properties.

Since the 1950s, it had been known that the simpler equations of incompressible hydrodynamics[s] are invariant under isotropic scale changes[43]—symbolically:

---

[r] In Chapter 4, I discuss the story of the aircraft measurements that—precisely because of the anisotropic scaling—turned out to be nontrivial!

[s] In incompressible fluids, gravity is irrelevant because there are no density differences upon which it can act. Water may often be treated as an approximately incompressible fluid, so this is by no means an academic situation. In the ocean, there are density variations resulting primarily from varying salt concentrations and temperatures, so that there, too, gravity is important.

$$Incompressible\ fluid\ laws \rightarrow Isotropic\ blowup\ (Factor\ \lambda)$$
$$\rightarrow \lambda^H\ (Incompressible\ fluid\ laws).$$

As we shall see in Chapter 4, this isotropic invariance of the laws is the theoretical basis of the main scaling theories that describe the weather and climate. But once again, the question is: What do we do about gravity and stratification? It actually took a surprisingly long time—and there are still outstanding issues—to show that the equations of the atmosphere (i.e., that include the anisotropy induced by gravity) are indeed invariant under anisotropic scale changes[44]:

$$Atmospheric\ laws \rightarrow Anisotropic\ blowup\ (Factor\ \lambda) \rightarrow \lambda^H\ (Atmospheric\ laws).$$

In any case, the systematic analysis of the outputs of weather and climate models discussed in Chapter 4 confirms they are scaling accurately. For theoreticians, the problem is that most of their theories are unrealistic. For the moment, only one or two of the exponents can be derived theoretically, and are given only by numerical calculation or from experimentation!

### 3.2.2 WHAT IS SCALE?

Starting with stratification, let's see how we can extend scaling from self-similarity to a more general scaling symmetry. To do this, we need to go back to the ideas of scale and of size. What do we mean when we say that something is a certain scale or that it is a certain size? First, the usual idea of the scale of an object is just its length:

$$Scale\ =\ Length,$$

where the length is a distance. Size is a related idea. It is a quantitative measure of the amount of "stuff" in the system.[45] For example in one, two, and three dimensions, it is[46]

$$Size = Length\ (one\ dimension)$$

$$Size = Area \approx (Length)^2\ (two\ dimensions)$$

$$Size = Volume \approx (Length)^3\ (three\ dimensions).$$

Putting all this together, we may write

$$Size \approx (Scale)^D\quad (D\ dimensions),$$

where $D = 1, 2, 3$ is the usual dimension. To generalize scaling, it turns out that, in addition to dimension (Chapter 2), we need to generalize both the notions of scale and size; but, we nevertheless retain this size–scale relation.

Consider the problem of determining the scale of a tree (Fig. 3.3). How would we do it if all we were given was a stick and a rule that told us how to make a blowup by a small amount, by a factor only a little larger than 1? Our first step might be to use the stick as a compass, to trace out a circle 1 unit in radius, just as we were taught in school. Using the blowup rule, we could then take this unit circle and blow it up by a small factor. If the rule was isotropic—the same in all directions—then we would obtain a slightly larger circle. In Figure 3.3, I used the factor 1.26 (i.e., 26%), yielding the "1.26 circle" (i.e., at scale 1.26), but we could imagine using a factor of 1.01 (1%) or less, just slightly larger than one. We could then repeat blowing up circles iteratively until we got to a circle that was just big enough for the tree to fit inside. By knowing the total blowup factor, we would then know the size of the tree in terms of the original unit ruler. If we wanted to determine the height, because the tree is much taller than it is wide, and because the scale would be the radius and the height would be the diameter, our scale would be an estimate of half the height. As long as we applied our method to trees that are taller than their width, we would always get the usual height by doubling this scale.
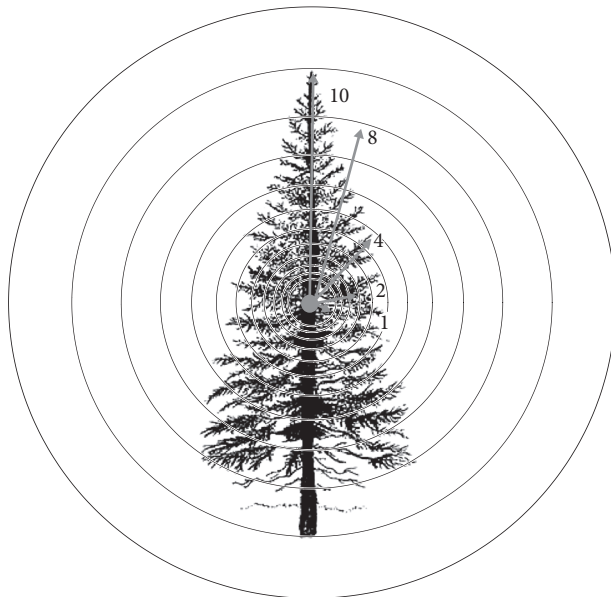


FIGURE 3.3  *To measure the size of a tree, we place a series of concentric circles at the midpoint and try to find the smallest circle that encompasses the tree. The radius of this circle is its scale. Here, it is 10 units.*

What about applying the method to a cloud? Figure 3.4 (top) shows a vertical cross-section of a cloud taken by the CloudSat satellite.[47] We can make out a large, horizontally stratified structure,[t] but when we look closely (especially toward the

---

[t] In this example, ignore the different shades of gray. Anything that is not dark black is taken to be a cloud.

lower right), we can see small vertically oriented ones. When we place our system of concentric circles on the image, we can already see a problem: For the large-enough clouds (structures), this way of determining the size will always end up measuring the horizontal extent, whereas for small-enough clouds, it will, on the contrary, end up measuring their vertical extents.



FIGURE 3.4  *A CloudSat vertical cloud cross-section with three different scale functions superposed (400 km left to right, 16 km bottom to top; aspect ratio, 6.5:1). The top scale function is isotropic at all scales (circles). In the middle scale function, the rule to change scales is isotropic so that, although the shape of the contours is now anisotropic, it is exactly the same at all scales. The thick gray contour line indicates the (arbitrary) unit scale, and, starting from this, each outer contour is a factor 1.58 enlargement; each inner contour is a factor 1.58 scale reduction. The bottom scale function is the same at the unit scale (thick gray line) as the previous, but now the rule itself is also anisotropic, so that in addition to the 1.58 enlargement, there is also a squashing of factor 1.23 in the vertical. This is the stratification measured for CloudSat vertical sections (Fig. 3.7A). It is close to the theoretical value corresponding to an elliptical dimension of 23/9.*

This manner of quantifying the cloud scale may seem academic. Why not simply define the scale as the horizontal length? The *distance* from one side to the other? But this would lead to the same problem. The horizontal length might be fine for quantifying the scale of a very wide, flat cloud deck, but it would be a lousy way of quantifying the scale of a system such as a tornado or a convective cell inside a thunderstorm, in which the physically important aspect is the vertical extent.

To get a physically meaningful measure of the scale of a cloud over the whole range from large to small, we could change the definition of scale to find a system of shapes that, when blown up or reduced, would fit all clouds (or at least fit "typical" or "average" clouds). The easiest way to modify our tree-measuring recipe is simply to change the definition of the unit scale. Rather than starting with a circle and blowing it up—or reducing it—in scale, we could start with a more appropriate shape, one adapted to the cloud in the image—for example, the "rounded triangles" in Figure 3.4 (middle). Using the same (isotropic) rule for blowing up as in the top case of Figure 3.4 (middle row) we would obtain a series of concentric rounded triangles. These contours represent all the points that are at the same scale. They are called "balls." Figure 3.4 shows that, at best, we will improve the fit of the balls to the larger cloud in the image, but there is no improvement in fitting the balls to the smaller vertically oriented ones.

Changing the definition of the unit scale didn't help much. What is needed is a change in the blow-up rule from one scale to another. Following what we learned from the analysis of the Weierstrass function, we could start with a conventional blowup, but follow it by a squashing in the vertical direction. By iterating the blowup and squashing, structures would become flatter and flatter at larger scales. Equivalently, going the other way—reducing—to smaller scales, the reduction would be accompanied by vertical stretching so that the structures would be more and more vertically oriented. In Figure 3.4 (bottom row), the blowup separating each of the balls is still by a factor of 1.58, but there is a final squashing by a factor of 1.23 in the vertical. Note that the same result would have been obtained by performing a usual blowup of 1.28 (= 1.58/1.23), followed by a stretch in the horizontal by a factor of 1.23. In both cases, the result is a blowup by a factor of 1.58 in the horizontal and by a factor of 1.28 in the vertical. We can see from Figure 3.4 (bottom row) that we are now doing a lot better. Not only are the large contours fitting the outlines of the large cloud, but inside the ~~red~~ unit circle, the small contours are now aligned vertically, apparently making a reasonable fit with the smaller clouds (this may be more obvious in Fig. 3.5, which is explained later).
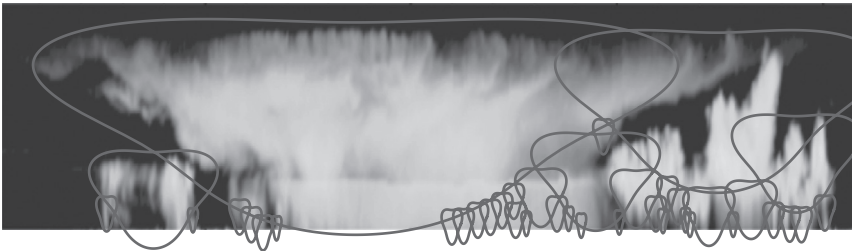
FIGURE 3.5  *A possible covering of the CloudSat cloud using the rounded triangular balls with the stratified rule for changing scales (Fig. 3.4, bottom row). The size can be estimated by summing the areas of all the balls (without double-counting the overlapping sections).*

The choice of factor 1.58 for each iteration was done for convenience. In this way, we get exactly five contours for every factor of 10 ($1.58^5 = 10$). But what about

factor 1.28? Where did that come from? To see what's going on, let's repeat the anisotropic blowup *n* times on the unit shape. The horizontal will be blown up by factor $1.58^n$ = (Horizontal scale), whereas the vertical will be blown up by factor $1.28^n$ = (Vertical scale). Eliminating *n*, we once again obtain a power law:

$$\text{Vertical scale} = \text{Horizontal scale}^{H_z},$$

with the new exponent $H_z = \log 1.28/\log 1.58 = 0.55$. If we had applied the same reasoning to the usual isotropic blowup, then we obtain

$$\text{Vertical scale} = \text{Horizontal scale},$$

and we recover $H_z = 1$. However, whenever $H_z$ is different from one, the aspect ratio of the contours change systematically with scale:

$$\text{Aspect ratio} = (\text{Horizontal scale} / \text{Vertical scale})$$
$$= (\text{Horizontal scale})^{1-H_z} = (\text{Vertical scale})^{1/H_z - 1}.$$

We can see that if the unit scale is fairly roundish, that whenever $H_z$ is less than one, the horizontal-to-vertical aspect ratio gets large for large scales so that the clouds become more and more horizontally stratified. Conversely, at small scales, the aspect ratio becomes large, so that the structures are vertically oriented. Figure 3.6 shows the result of applying this rule to circles using various values of $H_z$.
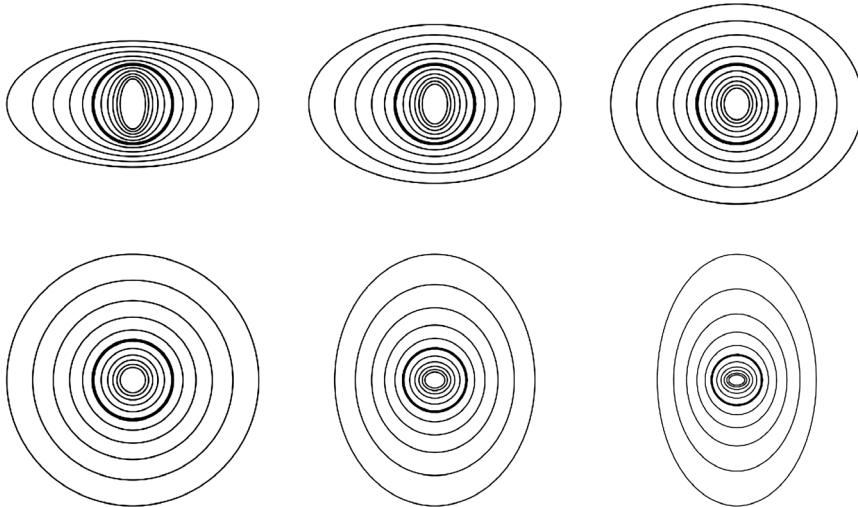


FIGURE 3.6 *This figure shows a series of ellipses, each separated by a factor of 1.26 in scale, with the thick black lines indicating the unit scales (here, circles). Upper left to lower right, $H_z$ increases from 2/5, 3/5, and 4/5 (top); and 1, 6/5, 7/5 (bottom, left to right). Note that when $H_z > 1$, the stratification at large scales is in the vertical rather than the horizontal direction.[48] Note that the size of each family of balls has been slightly rescaled to fit on the page (this can be gauged by the relative sizes of the thick black circles).*

We have discussed how to define the *scale* of a cloud in a way that takes into account its vertical stratification. We did this by finding a family of balls that defines all the points that have the same scale. Whether clouds are large or small, this definition of scale can be used to quantify their vertical and horizontal extents. But what about the *size* of a cloud? We already saw that, for simple balls (lengths, squares, cubes; or lengths, circles, spheres, for example) that the size is simply $\approx$(Scale)$^D$, where $D$ is the usual dimension: one for a line, two for a plane, three for a volume. What about these new stratified balls that take into account the flattening at larger scales?

Let's define the size of the unit ball as one. Then, every time we make a blowup and squashing operation to move from one ball to the next larger one, we can easily work out the factor by which the size increased. In the previous example, the blowup was by a factor of 1.58 in the horizontal and by $1.58^{Hz} = 1.23$ in the vertical ($H_z = 0.55$). Then, the increase in the size of the ball is by a factor of $1.58 \times 1.23 = 2.05$. After any number of iterations, we would have

$$\text{Size} = (\text{Horizontal blowup factor}) \times (\text{Vertical blowup factor})$$
$$= (\text{Horizontal blowup factor})^{1+H_z}.$$

Because we started at the unit scale, the horizontal blowup factor *is* the scale, so

$$\text{Size} = (\text{Scale})^{1+H_z}.$$

This suggests that we define a new "elliptical dimension"[49] $D_{el} = 1 + H_z$, so that this is the same as the usual definition of size:

$$\text{Size} = (\text{Scale})^{D_{el}}.$$

The new dimension is called "elliptical," although, as we saw, the balls are not necessarily ellipses. Applying this to the balls in Figure 3.6, we see that the areas of the ellipses increase as the horizontal scale raised to the following of powers $D_{el}$: 7/5, 8/5, 9/5, 2, 11/5, and 12/5 (upper left to lower right). The elliptical dimension thus quantifies the degree of stratification of the balls. When $D_{el}$ is close to one ($H_z$ is close to zero, upper left in Fig. 3.6), then the balls flatten rapidly at larger scales. When $D_{el}$ is two, $H_z$ is one and the balls are the usual circles. When the elliptical dimension is greater than two, $H_z > 1$ and the balls become more and more stratified in the vertical (vertically aligned, lower right).[50] We return to this relation later.

We've gone to a lot of trouble to define the scale and sizes of simple shapes (the balls), but cloud shapes are complex. What about them? Recall the discussion in Chapter 2 about covering a fractal set with "boxes" to estimate its fractal dimension. With our new language, the boxes are nothing but the balls, and we now know the size of each ball, so that the basic strategy is straightforward (Fig. 3.6). We can take the family of anisotropic balls from the bottom of Figure 3.4 and "cover" the cloud with them. Our estimate of the cloud size is simply the sum of the sizes of all the balls used in the covering. We can see that the covering used in Figure 3.6 includes a lot of overlap, so we could improve our estimate of the size of the cloud by using a better selection of balls—one that involves less overlap (if needed, we could even include infinitesimally small ones). Ultimately, the actual size of the cloud would be given by the "optimal" covering: the covering with balls that gives the smallest total sum.[51]

### 3.2.3  WHAT ABOUT REAL CLOUDS?

How can we tell what is the best definition of scale for real clouds? What is the best $H_z$ (and hence the elliptical dimension)? It turns out that we can use fluctuation analysis, which was described in Chapter 2. Averaging fluctuations over a CloudSat orbit enables us to obtain good statistics (Fig. 3.7A) and to use them to estimate statistically the relation between the horizontal and vertical extents. Let's say that we want to know the typical vertical extent of a structure 10 km across in the horizontal. First, we use fluctuation analysis to work out the typical change in the CloudSat radar reflectively over horizontal distances of 10 km. We then simply ask: How far must we go in the vertical so that typical fluctuations have the same amplitude? From Figure 3.7A, we see that, because the gradients are much stronger in the vertical, the answer is about 1.2 km. Pairing horizontal with vertical fluctuations systematically over a range of horizontal scales, we obtain Figure 3.7A. The figure shows that the results are close to the reference slope $H_z = 5/9 = 0.5555...$ . Also shown is the slope $H_z = 1$, corresponding to $D_{el} = 2$— in other words, equal to isotropic (self-similar) cloud sections. We see that isotropic behavior can clearly be rejected by the data. The reference line with slope $H_z = 5/9$ was hardly chosen at random. We mentioned earlier that, for wind, the fluctuation exponents were $H = 1/3, 3/5$ in the horizontal and vertical, respectively. The vertical must therefore be squashed using an exponent $H_z = (1/3)/(3/5) = 5/9 = 0.55...$ . We have just confirmed the anisotropic scaling model using CloudSat data!
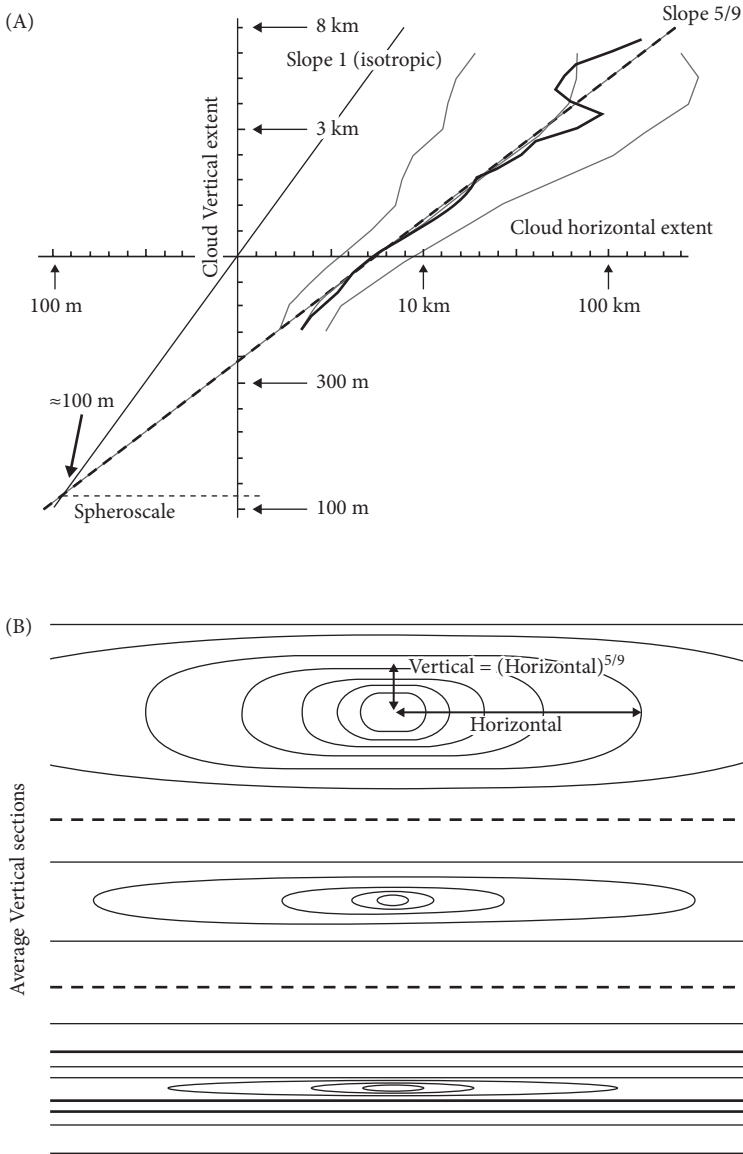
FIGURE 3.7 *(A) A space (horizontal)–space (vertical) diagram estimated from the average absolute reflectivity fluctuations from sixteen CloudSat orbits.[52] (B) The theoretical shapes of average vertical cross-sections using the CloudSat-derived mean parameters from (A): $H_z = 5/9$, with spheroscales 1 km (top), 100 m (middle), and 10 m (bottom), corresponding roughly to the geometric mean (100 m), and typical fluctuations away from the mean. In each of the three shapes (separated by dashed lines), the distance from left to right horizontally is 100 km, and from top to bottom vertically is 20 km. The top shape, in particular, shows that structures 100 km wide will be about 10 km thick whenever the spheroscale is somewhat larger than average.[53]*

The analysis of this CloudSat data is enough to tell us the exponent $H_z$, and hence the anisotropic blowup needed to go from one scale to another. However, to define the scale completely, we also need a "unit" ball to define all the balls and scales. Studying Figure 3.7A more carefully suggests a simple way of doing this. The lines with $H_z = 5/9$ and $H_z = 1$ meet at about 100 m, implying that moving 100 m in the horizontal direction would typically give the same change in radar reflectivity as moving 100 m in the vertical. The simplest shape that satisfies this is a circle 100 m in diameter. In the full horizontal–vertical space, this would correspond to a 100-m-diameter sphere, which is called the "spheroscale." At this unique scale, let's assume the cloud is isotropic.[u] However, there is no requirement that there be *any* ball that is exactly spherical—and even if one did exist, choosing it to define the unit ball would simply be a matter of convenience. Any of the family of balls that are related by blowups and squashings would be equally valid choices. With the assumption of isotropy at 100 m, we have defined implicitly all the balls (and scales) in the system. Figure 3.7B (middle) shows some of the corresponding balls. At the top and bottom are two other examples that show the system of balls when the spheroscale is larger or smaller by a factor of 10, corresponding roughly to the observed orbit-to-orbit variations in the spheroscale.

### 3.2.4 SIMULATING ANISOTROPIC SCALING, FRACTALS, AND CLOUDS

It may now be helpful to do the opposite of analysis: construction. Starting with something simpler than a cloud—a (black or white) fractal set—we can construct an anisotropic version of the Sierpinski carpet (Fig. 3.8), which can be compared with the isotropic version in Fig. 2.8. Although the isotropic Sierpinski carpet was constructed by dividing a square into nine smaller ones and then removing the central square, the anisotropic Sierpinski carpet is constructed by dividing the same original square by five in the horizontal and by three in the vertical to create $3 \times 5 = 15$ small rectangles, each of dimensions $1/5 \times 1/3$. Then, the middle three of the fifteen rectangles are removed, leaving only the twelve surrounding rectangles. The procedure is then iterated. After $n$ iterations, each box has Width $= (1/5)^n$ and Height $= (1/3)^n$. Eliminating $n$, we again find that

$$\text{Height} = (\text{Width})^{H_z},$$

with exponent $H_z = \log 3/\log 5 = 0.68$, which is not too far from the clouds with $H_z = 0.55$. Each rectangular box has Size $=$ Area $=$ (Height) $\times$ (Width) $=$ (Width)$^{1+H_z}$,

---

[u] Actually, there may be more than one scale that is isotropic—or they may never be such a scale. If the data permit, we can empirically determine the balls directly without any such assumptions. See, for example, Lovejoy, S., Schertzer, D., Lilley, M., Strawbridge, K. B., & Radkevitch, A. Scaling turbulent atmospheric stratification, part I: Turbulence and waves. *Q. J. Roy. Meteorol. Soc.* **134,** 277–300 (2008).

so that the elliptical dimension $D_{el} = 1 + H_z = 1.68$. We can also consider the fractal dimension of the carpet (the black points). The number of black boxes at the $n$th iteration is $12^n$, and since Width $= (1/5)^n$, eliminating $n$ we find Number $=$ Width$^{-D}$, with the fractal dimension $D = \log 12 / \log 5 = 1.54$. Notice that if there are no holes—meaning, the entire square remains black—then we would obtain $\log 15 / \log 5 = 1.68$, which is the elliptical dimension of the entire stratified space.[v] The latter is thus the maximum that is obtained when there are no holes. We can also notice that the holes (white) become more roundish eventually. At the smallest scales, they are even aligned vertically.
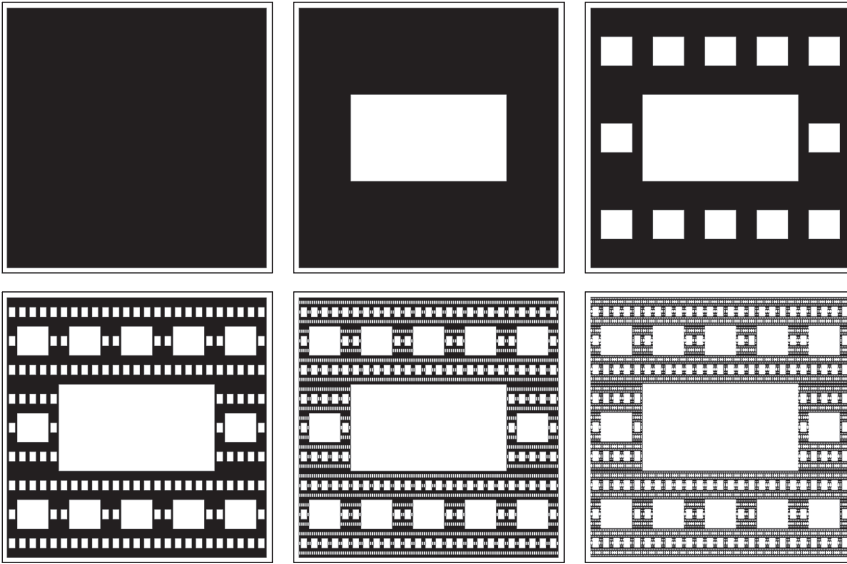


FIGURE 3.8  *An anisotropic Sierpinski carpet (black), produced by dividing the horizontal iteratively by five and the vertical by three, then removing the middle three rectangles (upper left to lower right).[54] Compare this with the original (isotropic) version of the carpet in Figure 2.8B. Notice that structures (holes) become more roundish, eventually even vertically aligned at the smallest scales.[55]*

Last, because we've fit data to clouds and estimated the exponent $H_z$, we can now do the opposite: make cloud simulations and see if they are realistic. Figure 3.9 compares the vertical sections of isotropic multifractal clouds (self-similar, left column), with the corresponding stratified, anisotropic multifractal clouds (right column). Notice that although the simulation starts as stratified cloud layers (upper right), at the highest zoom (lower right), the structures are roughly roundish.

---

[v] Of course, if we didn't remove any rectangle so that the final "fractal" was simply the same as the starting square, then we could use the usual definition based on square boxes, rather than rectangular stratified balls, and obtain the usual $D = 2$. Both dimensions are equally valid. They differ because of their different definitions of scale (the different series of boxes each uses).
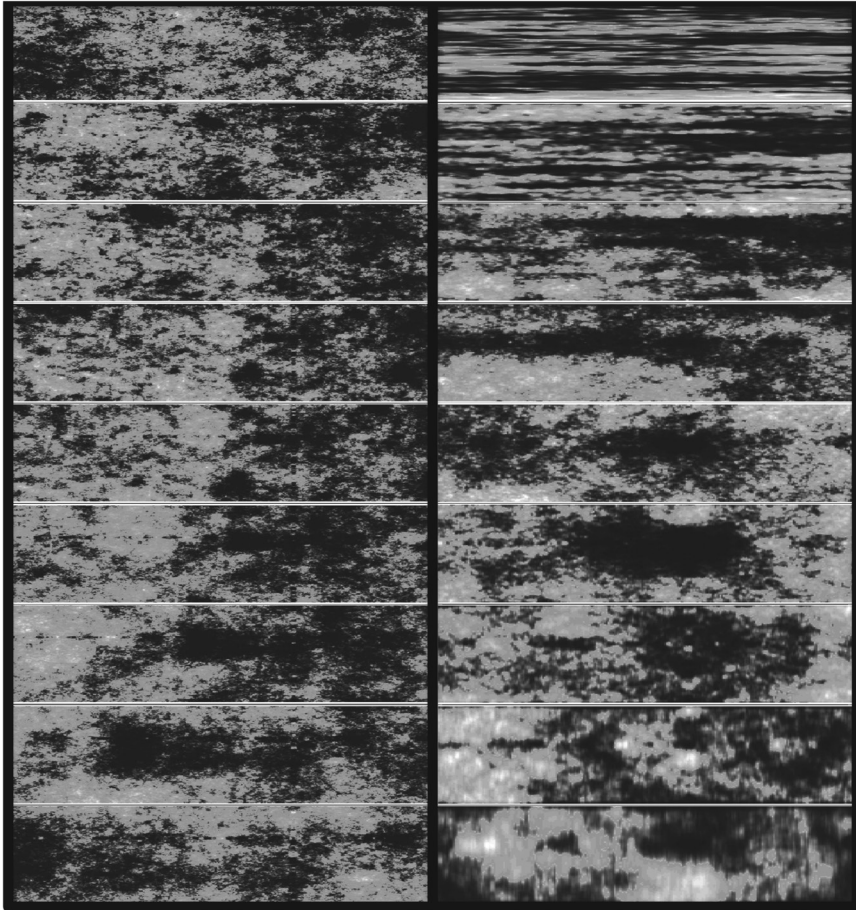
FIGURE 3.9 *(Left column, top to bottom) A sequence "zooming" into the vertical cross-section of an isotropic multifractal cloud (the density of liquid water was simulated and then displayed using a grayscale, with black sky below a low threshold). From top to bottom, we zoom in progressively by factors of 2.9 (total factor ≈ 1,000). We can see that typical cloud structures don't change. (Right column top to bottom) A multifractal cloud with the same statistical parameters as at left, but it is anisotropic. The zoom is still by factors of 2.9 in the horizontal, but now by $2.9^{H_z} = 1.80$ in the vertical, with $H_z = 5/9$. Notice that when at large scales, the clouds are strongly horizontally stratified; when viewed close up, they show structures in the opposite direction. The spheroscale is equal to the vertical scale in the rightmost simulation on the bottom row.*[56]

### 3.2.5 THE 23/9 D MODEL

Rather than restricting ourselves to vertical sections, let's consider the implications for cloud structures in the full three-dimensional space. Figure 3.10 shows typical "balls" that are obtained by assuming that the two horizontal directions are equivalent on average and by using various values of the exponent $H_z$. In three-dimensional space, the size of an object is its volume, and one can then see that because $(\text{Vertical}) = (\text{Horizontal})^{H_z}$, the typical volume of a three-dimensional object is

$$\text{Size} = \text{Volume} = \text{Horizontal} \times \text{Horizontal} \times \text{Vertical} = (\text{Horizontal})^{1+1+H_z},$$

so that the exponent $1 + 1 + H_z$ is the elliptical dimension $D_{el}$. If $H_z = 1$, then we would be back to the usual isotropic case and we see that the elliptical dimension would be equal to the usual dimension: $1 + 1 + 1 = 3$. Using $H_z = 5/9$, we obtain $D_{el} = 1 + 1 + 5/9 = 23/9 = 2.55$, hence the name the "23/9 D model."
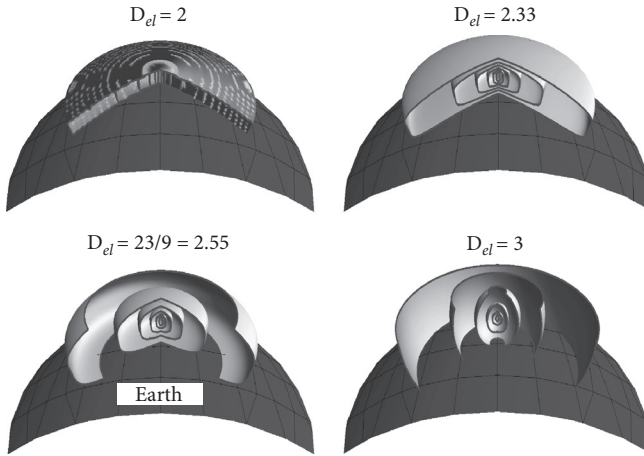


FIGURE 3.10 *A schematic diagram showing the change in shape of average structures that are isotropic in the horizontal (slightly curved to indicate Earth's surface), but with scaling stratification in the vertical. $H_z$ increases from zero (upper left) to one (lower right); $D_{el} = 2 + H_z$. To illustrate the change in structures clearly with scale, the ratio of tropospheric thickness to Earth's radius has been exaggerated by nearly a factor of 1,000. Note that in the case of $D_{el} = 3$, the cross-sections are exactly circles; the small distortion is an effect of perspective resulting from the mapping of the structures onto the curved surface of Earth.*[57]

Figure 3.10 shows several degrees of stratification ranging between two extreme cases: $H_z = 1$, the isotropic case we just examined, where the vertical and horizontal extents of clouds are always comparable; and $H_z = 0$, the completely flat case in which the thickness of the clouds is independent of their horizontal extent. These cases have elliptical dimensions three and two, respectively, and correspond to the predictions of classic isotropic turbulence in three- and two-dimensional space, respectively. Two intermediate cases are also shown—one corresponding to the CloudSat analysis, which we discovered turns out to be close to the theoretically predicted stratification, $H_z = 5/9$; as well as a fourth example, which corresponds to a different anisotropic theory based on gravity waves.[58]

In many respects, the atmosphere is like an onion skin surrounding the planet. The most obvious evidence is the pressure that falls off by a factor of 2 every 5 km in altitude, and the temperature that, on average, decreases from the surface to an inversion at around 10 to 18 km, where the temperature is a minimum—the "tropopause"—above which (in the stratosphere) it starts to increase again.[59]

With this atmospheric thickness in mind, we can see a problem with three-dimensional isotropic turbulence[w]: the case of $H_z = 1$. Figure 3.10 (upper left) shows

---

[w] Note that relative scales are exaggerated for graphical clarity.

that it can only hold over a narrow range, and we are forced to introduce a scale break at a scale somewhere in the vicinity of the thickness, at a scale of roughly 10 km (see Chapter 4). We can now begin to understand the significance of the mesoscale. It is defined conventionally to be over the range of horizontal distances comparable to the thickness[x]—roughly 1 to 100 km. Indeed, historically, the fundamental assumptions about turbulence have been scaling and isotropy—in other words, scaling is necessarily isotropic. This leaves only two choices: self-similar scaling in three dimensions or—if the layers are ever flat enough—self-similar scaling in two dimensions (i.e., with $H_z = 0$) (lower right in Fig. 3.10). However, if the stratification is scaling with $H_z$ between the extremes of zero and one, corresponding to an elliptical dimension between three and two, then—at least in principle—it would allow the scaling to hold over the entire range of horizontal scales.

We can check the relation between the horizontal and vertical directly by using the lidar smog data shown in Figure 1.8A, B. If we analyze it with the method discussed for CloudSat, we find that the full horizontal extent of 20,000 km does indeed correspond to layers about 10 km thick[60] (Fig. 3.11). The vertical stratification thus removes the need for a horizontal break in the scaling! At first sight, this result might seem surprising. How is it that there is no obvious signature of the 10-km vertical thickness in the horizontal direction? The answer is that the average atmospheric density and temperature are not dynamically important. What *is* dynamically important are the fluctuations, and they are not affected by the mean air density or mean temperature.
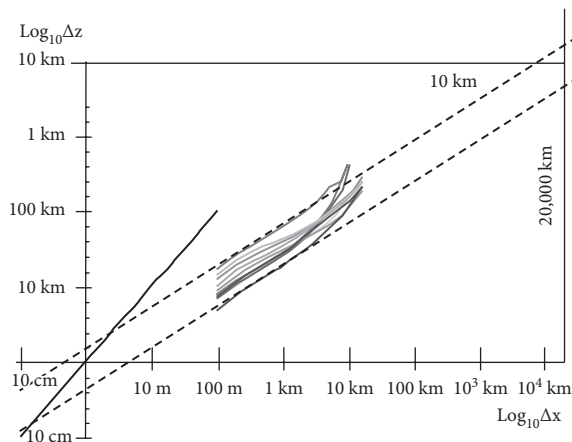


FIGURE 3.11 *The horizontal–vertical diagram for nine lidar smog data sets of the type shown in Figure 1.8A, B. The relationship was established using the same fluctuation analysis approach as for the CloudSat data in Figure 3.7A. It can be seen that the extrapolation to the size of the planet (dashed lines, up to 20,000 km) corresponds to about 10 km in the vertical. (The departure of three of the curves from linearity at the largest scales is an artifact of the poor statistics at these scales.) The steep solid line at the left shows the height–width relationship for isotropic clouds of smog, and the point where it intersects the dashed lines (at about 20 cm for the bottom and 2 m for the top dashed lines) is the estimate of the spheroscale, which depends on the relative strength of the vertical and horizontal gradients, thus varying from one smog cross-section to another.[61]*

---

[x] Recall that Richardson's diffusion data and the cloud area–perimeter relation (Fig. 2.14, left and right) showed that nothing special happened in the mesoscale. The classic expectation that the finite atmospheric thickness would imply a horizontal scale break is simply wrong, as explained in detail in Chapter 4.

### 3.2.6 THE EMERGENCE OF SCALE

What have we accomplished? Self-similar isotropic scaling was too restrictive to be of much use in the real world. Either we find a way to generalize it to take anisotropy into account or we return to business as usual and try to find different scalebound mechanisms to account for the behavior over every factor of 10 or so. Although the problem also afflicts horizontal cloud structures, we restricted our attention to vertical stratification. We discovered that to obtain a meaningful concept of cloud size—one that could apply over the whole range, from very small to very large (potentially from millimeters to thousands of kilometers)—we needed to take a lot of liberty with the way we defined size and scale. To be maximally flexible, we used only two ingredients. We started with a definition of the unit scale—so far we've considered only circles and "rounded triangles," but these were just simple examples—and followed this with an anisotropic blowup operation that could be thought of as either a usual blowup (isotropic, the same in all directions) followed by a squashing, or else simply a different blowup in the horizontal and vertical directions.

By considering the blowup rule to act over only very small factors and building up large factors by iterating many times, we generated a whole family of balls, each one composed of all the points obtained by blowing up the initial unit ball by an identical amount: the scale. The combination of the unit ball and the blowup operation was enough to define both scale and size in a way that is much more general than would have been possible if the blowups had been isotropic, which is appropriate only when the structures are self-similar. By definition, all the points on a ball have the same scale. Our recipe *defines* the notions of scale and size. We even were able to use satellite data to determine the notions of scale and size appropriate to real clouds. At least in such a stratified system—scale defined by atmospheric dynamics and size—they are *emergent* quantities.

## 3.3 Zooming with squashing and rotation, and the phenomenological fallacy

### 3.3.1 GENERALIZED SCALE INVARIANCE

So far, we have managed to save the scaling idea of atmospheric scaling even in the face of vertical sections with strongly anisotropic, nonself-similar behavior. It may seem that we have done little more than extend self-similarity to self-affinity à la the Weierstrass function—that by introducing families of balls we have needlessly complicated things. However, by turning our attention to the more challenging

horizontal direction, we'll again see the usefulness of using the balls to define scale and size.

Unlike the vertical, there is no evidence for any *overall* stratification in the horizontal analogous to the vertical,[62] but there is still plenty of evidence for the existence of different shapes at different sizes, and the fact that shapes commonly rotate[y] by various amounts at different scales. We thus need to go beyond self-affinity and (at least) add some rotation. The resulting framework that grew out of attempts to understand the vertical stratification has the insipid name "generalized scale invariance" (GSI),[63] but it allows scaling to be formulated as a very general symmetry principle, so that the notion of scale itself is effectively determined by turbulent dynamics.

It's now pretty easy to add rotation to the blowup and squashing that we discussed earlier. Again, start with some convenient definition[64] of the unit scale (ball, contour), and again, make a rule that tells us how to blow up by a factor close to one—a rule telling us how to make the contour a little bit bigger.[65] As before, the new rule will involve a regular blowup and squashing, but now add some rotation to the squashed, enlarged ball. In all of these examples, we have used a linear transformation: matrix multiplication of the scale ratio raised to the power of a matrix, called the "generator" of the scale transformation group. Figure 3.12A shows the result of a series of scale reductions of 28% applied to the acronym NVAG, which means nonlinear variability in geophysics. We see there is both a stretching and rotation (and hence deformation) of structures with scale.

---

[y] Here and later, when we refer to "rotation with scale," we mean the systematic change of orientation of structures with their size, not the rotation of a real structure changing its orientation as it evolves in time. Of course, the two phenomena are likely to be related to each other.

(A)



(B)
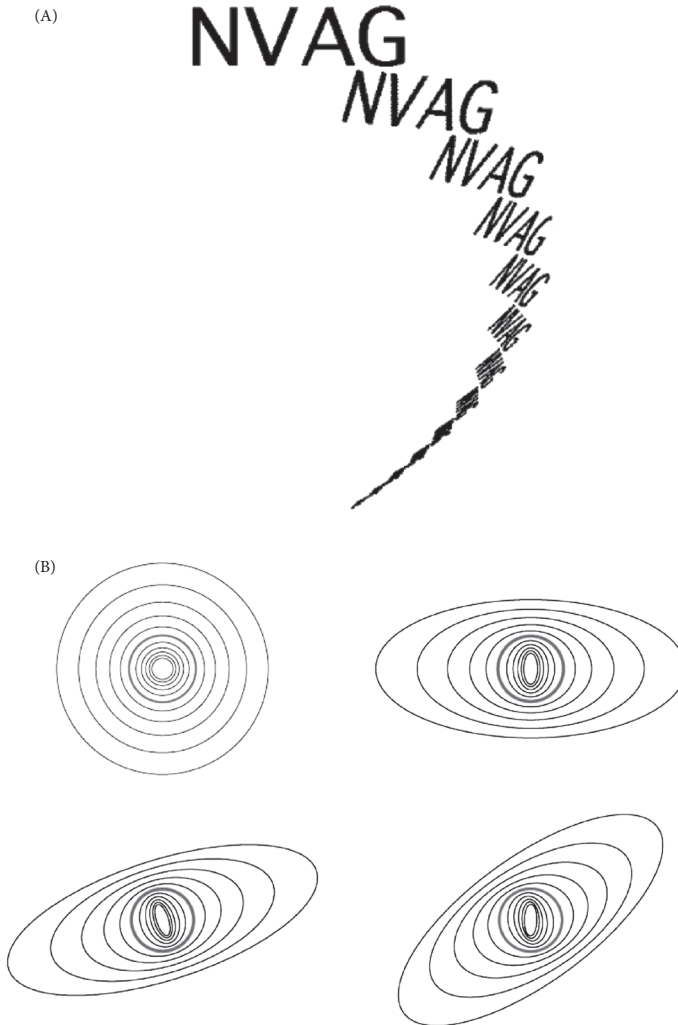


FIGURE 3.12 *(A) A generalized blow-down (or reduction) of the acronym NVAG, with each reduced in scale by 28%.[66] (B) Blowups and reductions by factors of 1.26, starting at circles (thick lines). The upper left shows the isotropic case, the upper right shows the self-affine (pure stratification case), the lower left example is stratified but along oblique directions, and the lower right example has structures that rotate continuously with scale while becoming increasingly stratified.[67] (C) The same as (B), except that now the unit ball is the rounded triangle. (D) The same blowup rule is used as in the lower right of Figure 3.7B, but this view shows an overall blowup by a factor of 1 billion. Starting with the inner thick gray ball in the upper left corner, we see a series of ten blowups, each by a factor of 1.26, spanning a total of a factor of 10 (the outer red ball). Then, that ball is shrunk (as indicated by the dashed lines) to show, conveniently, the next factor of 10 blowup (top middle). The overall range of scales in the sequence is thus $10^9 = 1$ billion. The scale-changing rule (matrix) used here is the same as that used for the lower right in (C). (E) A different example of balls with squashing but with only a little rotation. The maximum rotation of structures in this example, from very small to very large scales, is 55°.[68]*

(C)

(D)



FIGURE 3.12 *Continued*

(E)
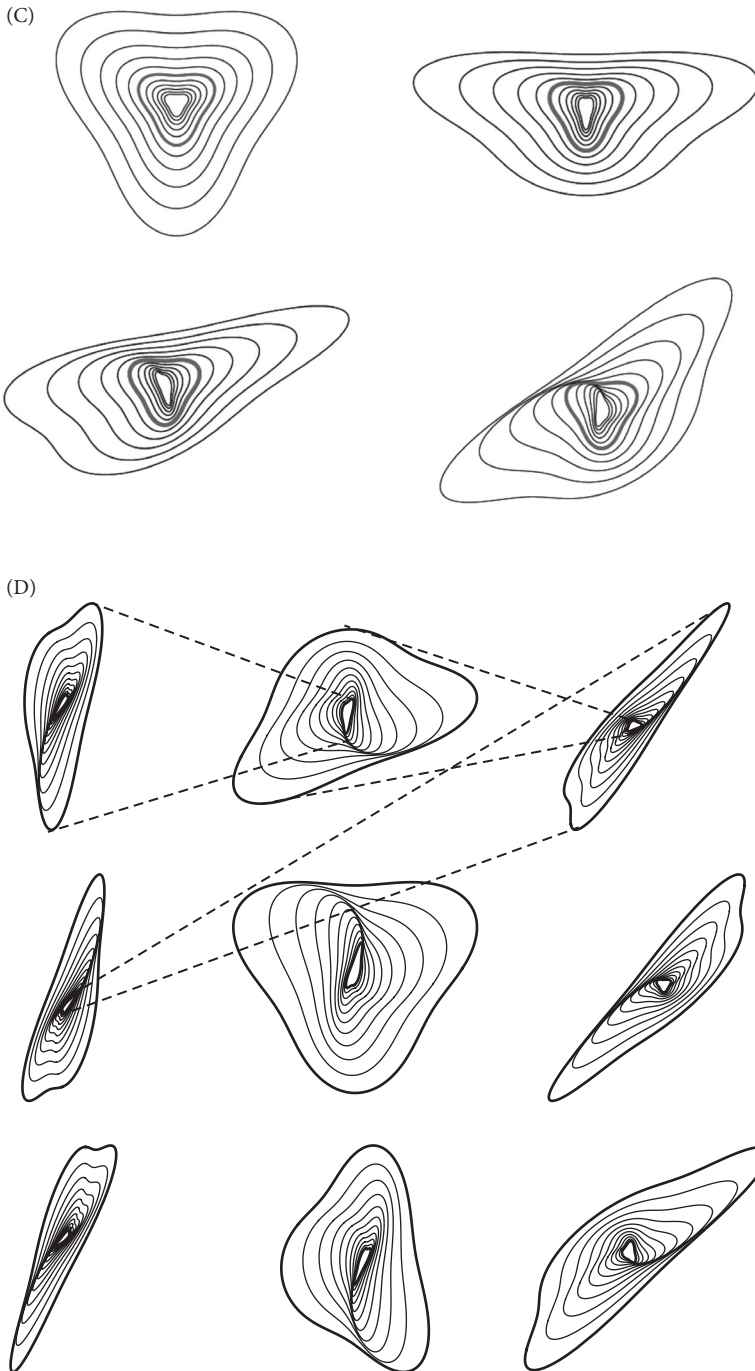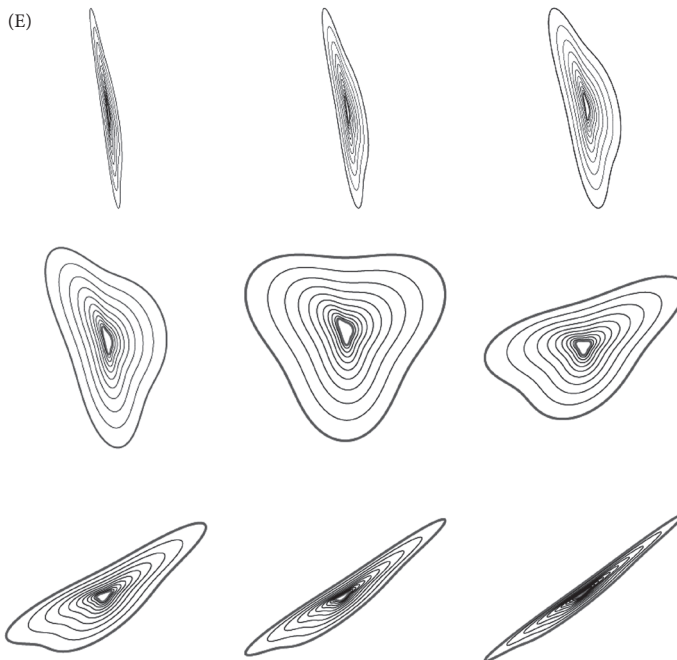
FIGURE 3.12 *Continued*

Figure 3.12B, C shows a few examples of contours at different scales, each representing the shapes of the balls at systematically varying scales. We can see that we have freedom to vary the unit balls (here, circles and rounded triangles) as well as the amounts of squashing and rotation. In Figure 3.12B, with unit balls taken to be circles, we show the self-similar case in the upper left, a stratified case in the upper right, a stratified case with a small amount of rotation in the lower left, and another case with lots of rotation in the lower right. Figure 3.12C shows the same, but with unit balls as rounded triangles. Figure 3.12D takes the lower right example and displays it over a factor of 1 billion in scale. Figure 3.12E shows an example with only a little rotation but over the same factor of 1 billion in scale. We can see that if these represent average morphologies of clouds at different scales, even though there is a single unique rule or mechanism to go from one scale to another, the average shapes change quite a bit with scale.

### 3.3.2 ANISOTROPIC MULTIFRACTALS

We have explored ways in which quite disparate shapes can be generated using blowups, squashings, and rotations. With the help of a unit ball, we generated families of balls, any member of which would have been an equally good starting point. The unit ball has no particular importance; it does not have any special physical role to play. If we have a scaling model based on isotropic balls, then replacing them with anisotropic balls is also scaling when we use the anisotropic rule to change scales. Any morphology made using such a system of balls is scale invariant. We already examined a self-similar and stratified example in which the balls were used to make a multifractal

cloud simulation of a vertical section (Fig. 3.9, left). Let's now take a quick look at a few examples of horizontal and three-dimensional multifractal cloud simulations.

The simulation of a cross-section of a stratified multifractal cloud in Figure 3.9 already shows that the effect of changing the balls can be quite subtle. Let's take a look at this by making multifractal cloud simulations with realistic (observed) multifractal parameters (which determine the fluctuation statistics, not the anisotropy), and vary the families of balls systematically (Fig. 3.13A). In Figure 3.13, all the simulations have the same random "seed,"[z] so that the only differences are the result of the changing definition of scale. First we can explore the effects of different degrees of stratification combined with different degrees of rotation. Let's consider two cases. In the first (Fig. 3.13A), there is roughly a circular unit ball within the simulated range; in the second (Fig. 3.13B), all the balls are highly anisotropic. Each figure shows a pair: the cloud simulation (left) and the family of balls used to produce it on the right.
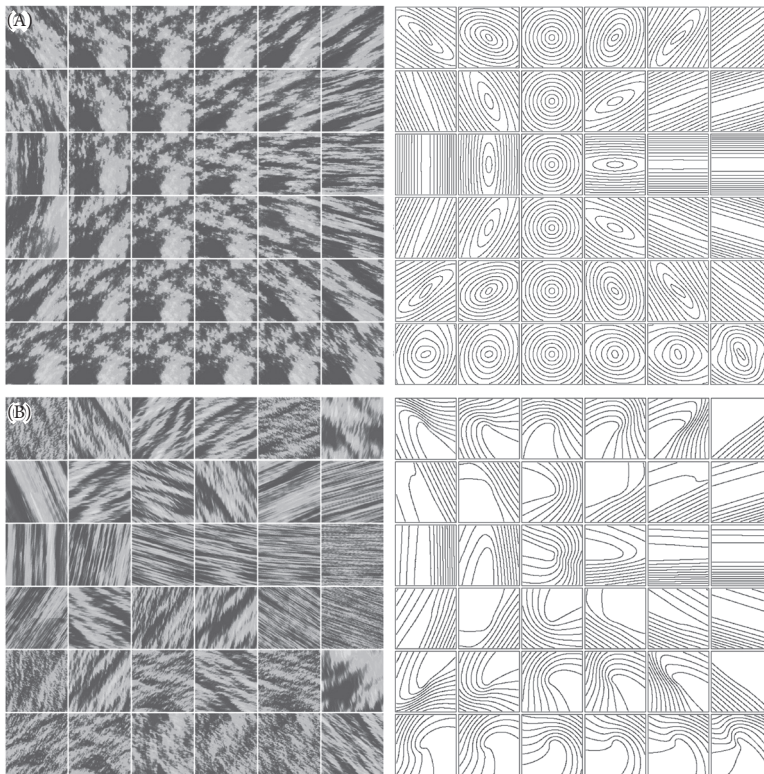


FIGURE 3.13 *(A) (Left) Multifractal simulations with nearly isotropic unit scales, with stratification becoming more important up and down away from the center line, and the rotation parameter (left to right) becoming more important as we move away from the third column. (Right) The balls used in the simulations to the left. This is an extract from the multifractal explorer website: http://www.physics.mcgill.ca/~gang/multifrac/index.htm. (B) The same as (A), except the initial ball is highly anisotropic in an attempt to simulate the effect of stretching that is a result of a wide range of larger scales.*[69]

---

[z] This means that the different shapes, structures, have the same set of random weights produced by the same series of random numbers in the computer generation algorithm.

From the third column in Figure 3.13A, with no stratification, we can note that changing the amount of rotation (moving up and down the column) changes nothing; this is simply because the circles are rotated to circles. Rotation is only interesting when combined with stratification. The simulations in Figure 3.13B might mimic small clouds (for example, 1 km across) produced by complex cascade-type dynamics that started rotating and stratifying at scales perhaps 10,000 times larger. In both sets of simulations, the effect of stratification becomes more important up and down away from the center line, and the effects of rotation vary from the left to the right, becoming more important as we move away from the third column.

Figure 3.14 shows examples in which rotation is strong and the scale-changing rule is the same everywhere. Only the unit ball is changed. By making the latter have some long narrow parts, we can obtain quite wispy-looking clouds.
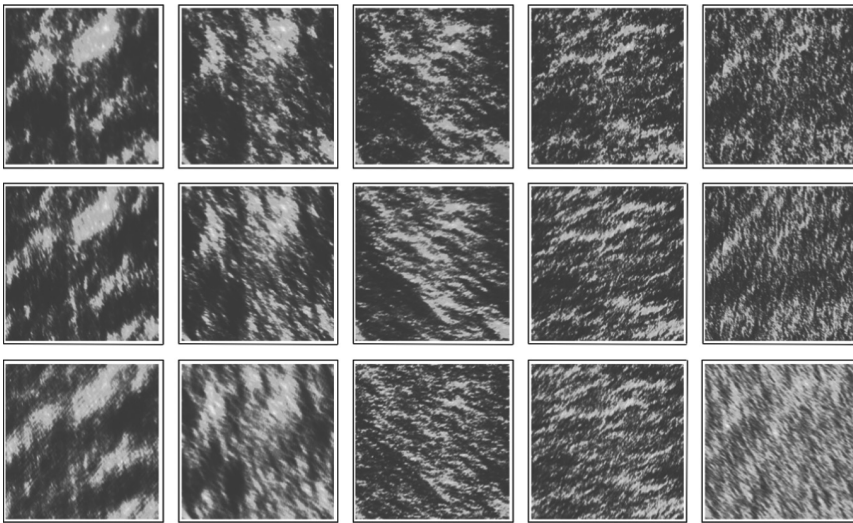


FIGURE 3.14 *Simulations of cloud liquid water density with the scale-changing rule the same throughout. Only the unit balls are modified systematically to yield more and more wispy clouds.*[70]

Figure 3.15A shows another aspect of multifractal clouds. In Chapter 2, we discussed the fact that, in general, the cascades occasionally produce extreme events (Box 3.1). If we make a sufficiently large number of realizations of the process (i.e., each time using a different set of random numbers), from time to time we will generate rare cloud structures that are almost surely absent on typical realizations. For example, a typical satellite picture of the tropical Atlantic Ocean would not have a hurricane, but hurricanes do appear there from time to

time. The multifractality implies that this could happen quite naturally, without the need to invoke any special scalebound "hurricane process." In the examples in Figure 3.15A, we see a set of balls rotating with scale (Fig. 3.15B). However, to simulate occasional, rare realizations, the process is "helped" by boosting artificially the values in the vicinity of the central pixel. The two different rows are identical except for the sequence of random numbers used in their generation. For each row, moving from left to right, we boosted only the central region to simulate stronger and stronger vortices that are more and more improbable.[71]

As we do this, we see that the shapes of the basic set of balls begin to appear out of the chaos.
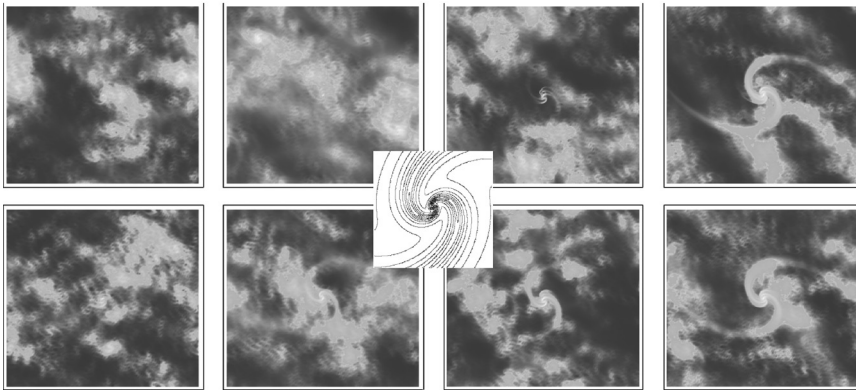


FIGURE 3.15  *(A) Each row has a different random seed, but is otherwise identical. Moving from left to right, we see a different realization of a random multifractal process, with the central part boosted by factors increasing from left to right, to simulate very rare events. The balls are shown in (B)₁ (B) The balls used in (A) exhibit the contours of the (rotation-dominant) scale function used in the simulations in (A).[72]*

The cloud simulations in Figure 3.15 are for the density of cloud liquid water; they use false colors to display the more- and less-dense cloud regions. Real clouds are, of course, in three-dimensional space, and the eye sees the light that has been scattered by the drops. Therefore, if we make three-dimensional cloud simulations, instead of simply using false colors, we can obtain more realistic renditions by simulating the way light interacts with the clouds.[73] Figure 3.16A, B shows the top and side views of a multifractal cloud with the usual false colors. Figure 3.16C, D shows the same cloud rendered by simulating light traveling through the cloud (in both top and bottom views). Last, in Figure 3.16E, a simulation of thermal infrared radiation emitted by the cloud is presented, similar to what can observed from infrared weather satellites. We see that quite realistic morphologies are possible.
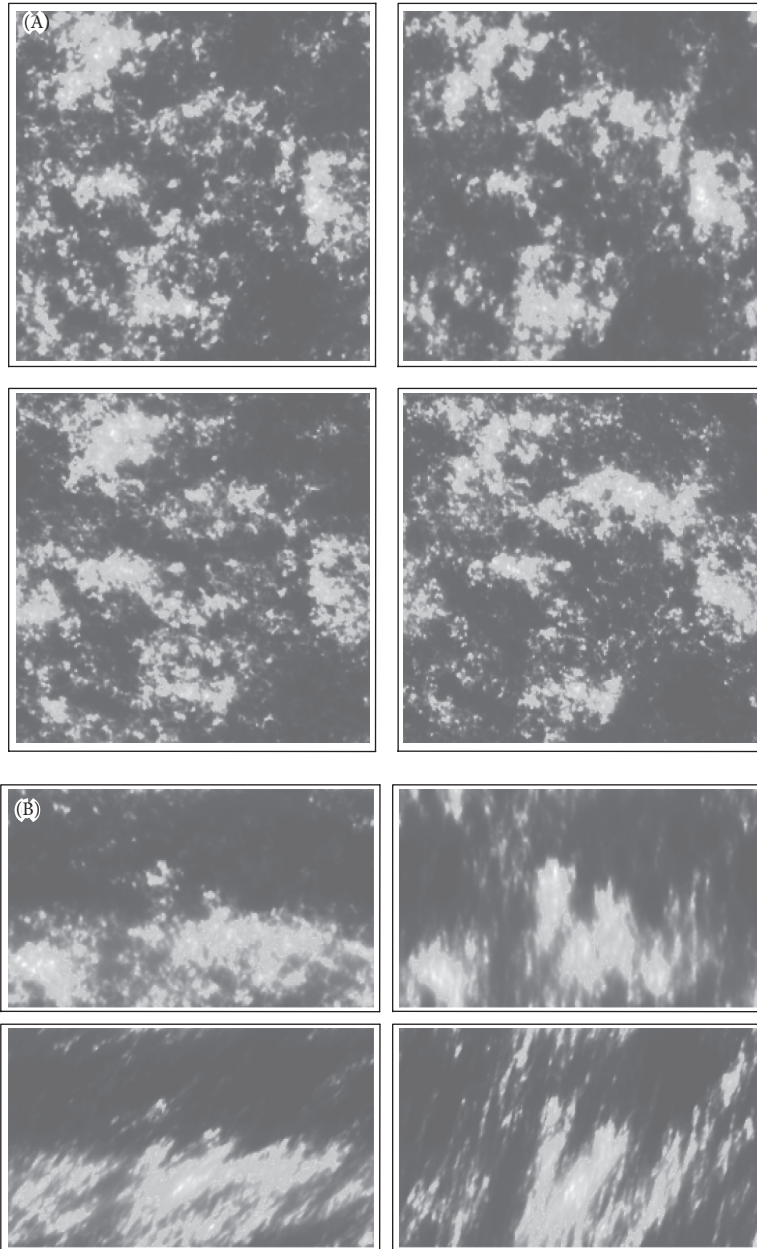
FIGURE 3.16  *(A) The top layer of cloud liquid water using a gray shade rendition. (B) A side view of (A). (C) The top view with light scattering from the sun (incident at 45° to the right). (D) The same as (C), except viewed from the bottom. (E) The same as (D), except for a gray shade rendition of a thermal infrared field, as might be viewed by an infrared satellite.[74] (F) Examples of simulations in space-time showing wavelike morphologies. The same basic shapes are shown, but with wavelike character increasing clockwise from the upper left.[75]*
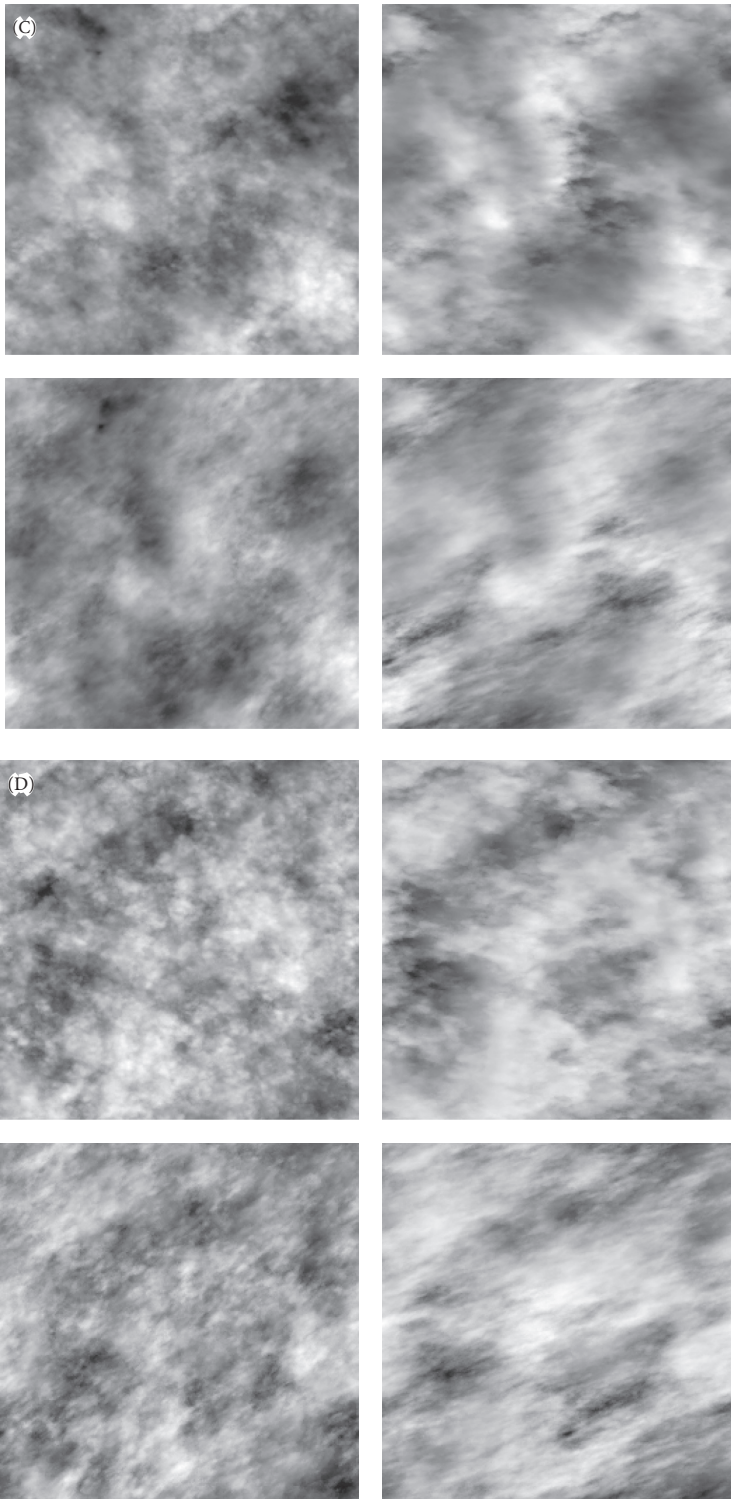
FIGURE 3.16 *Continued*
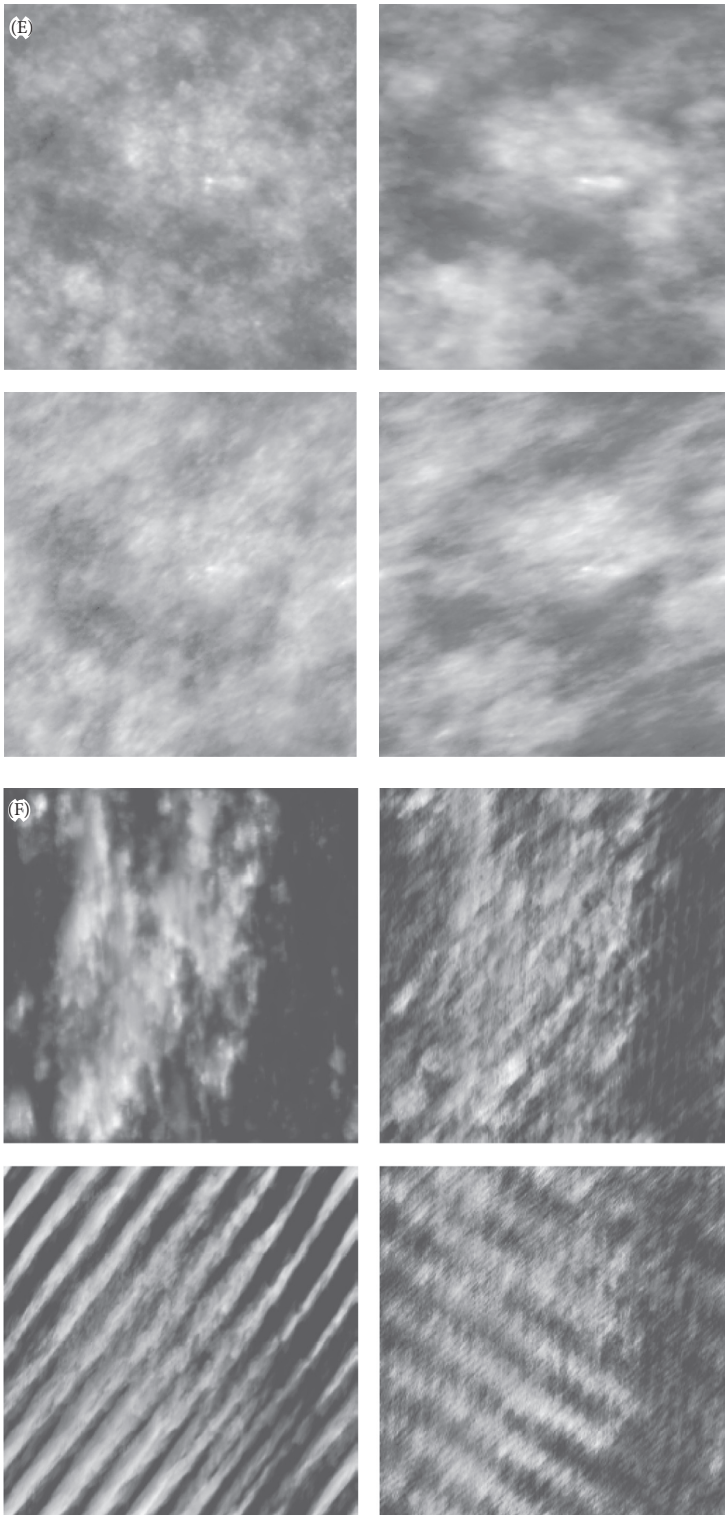
FIGURE 3.16 *Continued*

Up until now, we have only discussed space, but of course clouds and other atmospheric structures evolve in time. Because we have argued that the wind field is scaling—and the wind moves clouds around—it, effectively, couples space and time. We therefore have to consider scaling in space and in time: in space-time. The time domain opens up a whole new realm of possibilities for simulations and morphologies. Although the balls in space must be localized—because they represent typical spatial structures, "eddies"—in space-time they can be delocalized and form waves. In this case, it turns out that it is easier to describe the system using the Fourier methods mentioned in Chapter 2. Figure 3.16F shows examples of what can be achieved with various parameters.

### 3.3.3  ANISOTROPY THAT CHANGES FROM PLACE TO PLACE AS WELL AS SCALE TO SCALE: NONLINEAR GSI

To justify this rather lengthy excursion into GSI, I invoked the fact that zooming into clouds would display systematic changes of morphology with magnification, so that to be realistic, we needed to generalize the idea of self-similar scaling. The first step was to account for the stratification. This was straightforward, because gravity acts everywhere and at all scales. The fact that it also has a well-defined direction meant that the stratification was not rotating with scale and was in the horizontal (large) and the vertical (small). To model the horizontal plane, I needed to add rotation with scale. And, to a first approximation, we could think of the different cloud morphologies as corresponding to different cloud types: cumulus, stratus, cirrus, and so on.

But there is still a problem. We found a way of combining the scaling symmetry with morphologies that change with scale. But if this is all that we can do, then scaling is still implausible because cloud types and morphologies not only change with scale, they also change with spatial location (and in time). Figure 3.17A shows the problem with a real satellite infrared cloud picture. It seems clear that the textures and morphologies vary from one part of the image to another. Using a type of two-dimensional fluctuation analysis, we can try to estimate the corresponding "balls." When the image is broken up into an $8 \times 8$ array of subimages (Fig. 3.17B, with a fair bit of statistical scatter), we can confirm that the balls are quite different from one place to another.

(A)



(B)
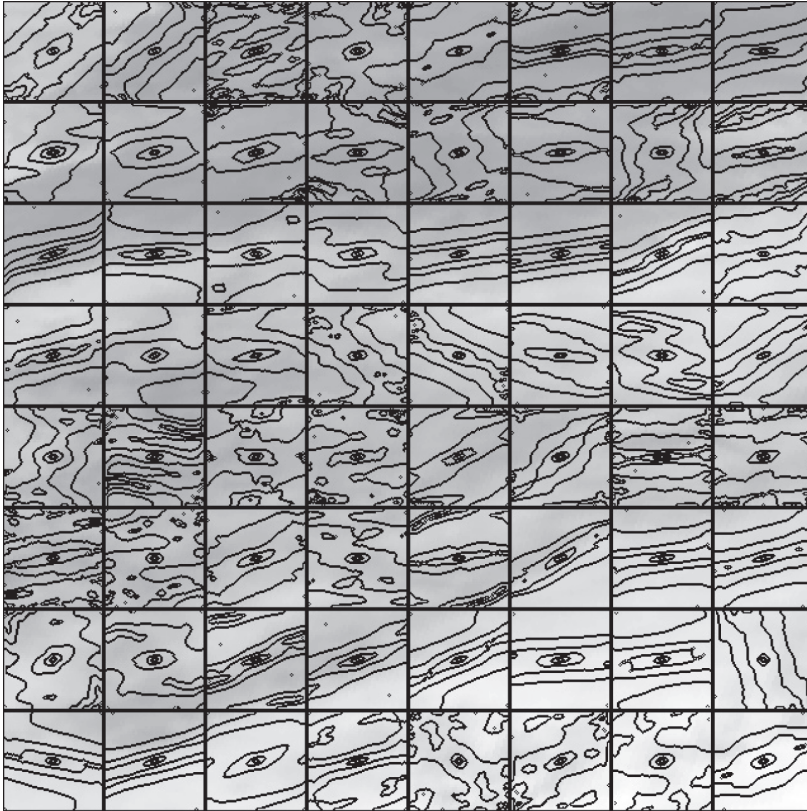


FIGURE 3.17  *(A) An infrared satellite image from a satellite at 1.1-km resolution (512 × 512 pixels). (B) Estimates of the shapes of the balls in each 64 × 64-pixel box from the image in (A).*[76]

What we have described so far is only a special case of GSI—the case in which the morphologies at a given scale are the same everywhere in space: "linear GSI" based on matrices. However, it turns out that it is possible to go beyond this to a more general "nonlinear" GSI, in which the notion of scale depends not only on the scale, but also on the location. Figure 3.18 shows an example. The physics behind this is analogous to those in Einstein's theory of general relativity. In the latter, the distribution of mass and energy in the universe determines the appropriate notion of distance (i.e., the metric). With GSI, nonlinear turbulent dynamics determine the appropriate notions of scale and size.[77]
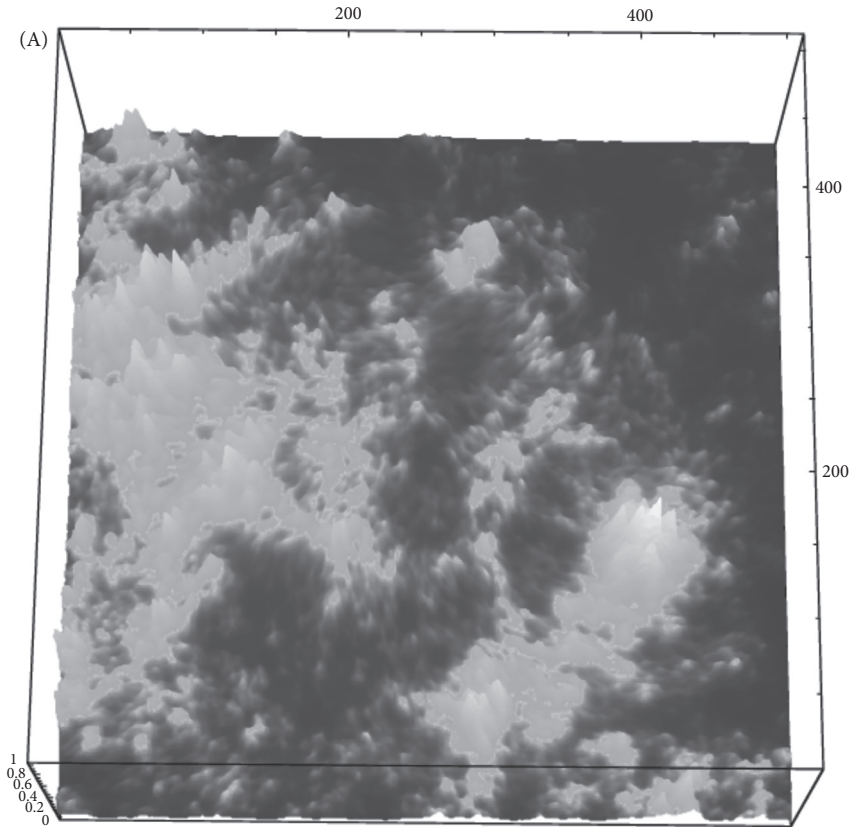


FIGURE 3.18 *(A) A multifractal simulation of a cloud with texture, morphology, varying in both location and scale, simulated using nonlinear GSI. The anisotropy depends on both scale and position according to the balls shown in (B). (B) The set of balls displayed according to their relative positions used in the simulation shown in (A).*[78]

(B)
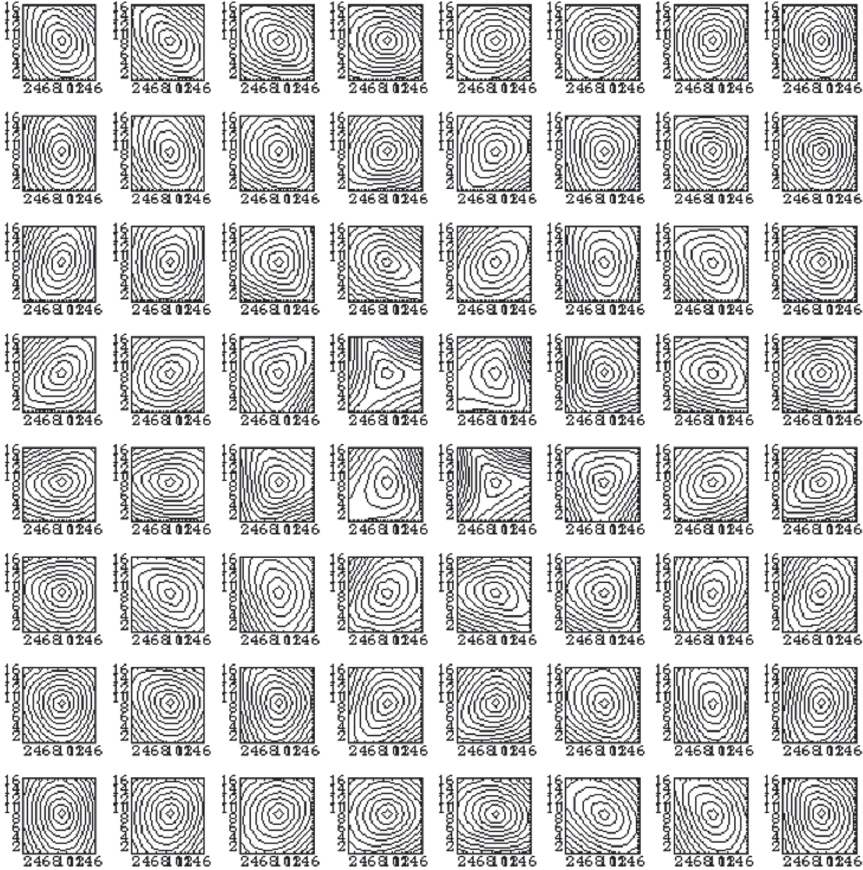


FIGURE 3.18 *Continued*

With nonlinear GSI, a bewildering variety of phenomena can be described in a scaling framework. The framework turns out to be so general that it is hard to make further progress. It's like saying, "The energy of the atmosphere is conserved." Although this is undoubtedly true—and it enables us to reject models that fail to conserve it—this single-energy symmetry is hardly adequate for modeling and forecasting the weather. One can imagine that if one must specify the anisotropy both as a function of scale and as a function of location, many parameters are required. At a purely empirical level, they are difficult to estimate because the process has such strong variability and intermittency. To progress much further, we'll undoubtedly need new ideas. However, the generality of GSI does make the introduction of scalebound mechanisms seem particularly superfluous.

### 3.3.4 ~~THE SCALEBOUND APPROACH AND~~ THE PHENOMENOLOGICAL FALLACY

I've given you a taste of the enormous diversity of cloud morphologies possible within the scaling framework. We examined morphologies that were increasingly stratified at larger scales, that rotated with scale (but only a bit), or that rotated many times. There were filamentary structures, there were structures with waves, and there were structures with characters that changed with position. Although all of these morphologies changed with scale, they were all consequences of mechanisms that were scale invariant. The scalebound approach is, therefore, not only scientifically unjustified, it is logically wrong. When scalebound mechanisms and models based solely on phenomenological appearances are invoked, they commit a corollary of the scalebound approach: the "phenomenological fallacy."[79] More concisely, the phenomenological fallacy is the inference of mechanism from phenomenology (from appearances).[aa]

AQ: Heading 3.3.4 is provided as "The phenomenological fallacy" under the table of contents. Please make this consistent.

## Notes

1. An important example of this failure was in the application of scaling approaches to solid Earth geophysics. Numerous, seemingly promising applications of the scaling symmetry in solid Earth geophysics, including coastlines and topography, were stymied by the unnecessary restriction of scaling to fractal sets of points. By the early 1990s, this led to a backlash against scaling and to the unfortunate rejection of scaling altogether. If the appropriate multifractal scaling concepts had been used, this might have been avoided. For a review, see Lovejoy, S. & Schertzer, D. Scaling and multifractal fields in the solid earth and topography. *Nonlin. Proc. Geophys.* **14**, 1–38 (2007).

2. Kadanoff, L. P. Fractals: Where's the physics? *Physics Today* **February**, 6–7 (1986).

3. Noether, E. Invariante Variationsprobleme. *Nachr. kgl. Ges. Wiss. Göttingen* **1918**, 235–257 (1918).

4. Strictly speaking, these symmetries apply to the Lagrangians that determine the fundamental forces. It turns out that at cosmic scales, the expansion of the universe breaks the time symmetry and there is no cosmological conservation of energy.

5. This term was coined a little later. Schertzer, D. & Lovejoy, S. The dimension and intermittency of atmospheric dynamics. In: *Turbulent Shear Flow* (eds. L. J. S. Bradbury, F. Durst, B. E. Launder, F. W. Schmidt, & J. H. Whitelaw), pp. 7–33. (Springer-Verlag, 1985).

---

[aa] In Box 3.1, we examined an extension of the phenomenological fallacy that arises when the probability of extremes is scaling so that a unique mechanism generates both weak and strong extreme events.

6. Mandelbrot, B. B. Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.* **62**, 331–350 (1974).

7. Levy, P. *Calcul des probabilites*. (Gautier-Villars, 1925).

8. Schertzer, D. & Lovejoy, S. Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes. *J. Geophys. Res.* **92**, 9693–9714 (1987).

9. Mandelbrot, B. B. Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.* **62**, 331–350 (1974).

10. Schertzer, D. & Lovejoy, S. Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes. *J. Geophys. Res.* **92**, 9693–9714 (1987).

11. Lovejoy, S. & Schertzer, D. Stochastic chaos and multifractal geophysics. In: *Chaos, Fractals and Models 96* (eds. F. M. Guindani & G. Salvadori), pp. 38–52. (Italian University Press, 1998).

12. Bak, P., Tang, C., & Weiessenfeld, K. Self-organized criticality: An explanation of 1/f noise. *Phys. Rev. Lett.* **59**, 381–384 (1987).

13. Schertzer, D. & Lovejoy, S. Multifractal generation of self-organized criticality. In: *Fractals in the Natural and Applied Sciences* (ed. M. M. Novak), pp. 325–339. (Elsevier, North-Holland, 1994).

14. Schertzer, D., Lovejoy, S., & Lavallée, D. Generic multifractal phase transitions and self-organized criticality. In: *Cellular Automata: Prospects in Astronomy and Astrophysics* (eds. J. M. Perdang & A. Lejeune), pp. 216–227. (World Scientific, 1993).

15. Sardeshmukh, P. D. & Sura, P. Reconciling non-Gaussian climate statistics with linear dynamics. *J. Climate* **22**, 1193–1207 (2009). This mechanism combines additive and multiplicative processes. It is similar to the multifractal butterfly effect.

16. One reason that scaling probabilities are disliked is that if we calculate the statistical moments of order higher than $q_D$, they do not converge. Common mathematical assumptions—used, for example, in "statistical closure" theories and approximations—are therefore inappropriate.

17. The power law can only hold for large fluctuations, $\Delta T$, because the total probability is equal to one and so is bounded. In practical applications, it isn't always obvious where the extreme power law probability "tails" start.

18. For the wind, the power law probabilities that Mandelbrot predicted theoretically more than forty years ago—and even the value $q_D = 5$, which was discovered empirically more than thirty years ago (Fig. 3.2)—has recently been numerically confirmed for the first time. This was possible because computer technology allowed the equations of fluids to be solved on an enormous grid—in a box of $8{,}192 \times 8{,}192 \times 8{,}192$ points—and the simulation generated histograms with more than a quadrillion events ($10^{15}$). Ironically, the authors of the article were apparently unaware of these earlier, decades-old results: Yeung, P. K., Zhai, X. M., & Sreenivasan, K. R. Extreme events in computational turbulence. *PNAS* **112** (41), 12633–12638 (2015).

19. Adapted from Ladoy, P., Lovejoy, S. & Schertzer, D. In: *Non-linear Variability in Geophysics: Scaling and Fractals* (eds. D. Schertzer & S. Lovejoy), pp. 241–250. (Kluwer, 1991).

20. Antarctica; paleotemperatures from $^{18}$O proxies. Reproduced with permission from Lovejoy, S. & Schertzer, D. Scale invariance in climatological temperatures and the local spectral plateau. *Annal. Geophys.* **4B**, 401–410 (1986).

21. Reproduced with permission from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

22. Adapted from Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014).

23. This insensitivity to details is an example of "universality" (Box 1.1). For example, measurement errors may be the result of a large number of different factors, and the individual contributions to the total error may be from distributions that are not at all Gaussian. This is a rather rough statement of a mathematical result called the "central limit theorem." But to be more precise, we require extra and not always obvious assumptions. To start with, for the result to be Gaussian, the variances of each contributing factor must be finite; otherwise, the result will be a (power law) Levy distribution. However, we also require that individual contributions be statistically independent of each other, or at least only weakly dependent. It is this assumption that is violated in multifractal processes; they have strong "long-range" statistical dependencies.

24. In Chapter 5, we look at a useful exception: several (nearly) nonintermittent macroweather time series (but not their spatial transects!). For these, the fractional Gaussian noise process (which follows a bell curve) can be a good approximation, although, even here, the extremes are not quite Gaussian and may in fact be power laws with $q_D \approx 5$.

25. Early works that found power law extremes in precipitation, wind, temperature, and potential temperature are as follows: Lovejoy, S. Analysis of rain areas in terms of fractals. In: Proceedings of the *20th Conference on Radar Meteorology*, pp. 476–484. (American Meteorological Society, 1981). Lovejoy, S. & Mandelbrot, B. B. Fractal properties of rain and a fractal model. *Tellus* **37A**, 209 (1985). Schertzer, D. & Lovejoy, S. The dimension and intermittency of atmospheric dynamics In: *Turbulent Shear Flow* (eds. L. J. S. Bradbury, F. Durst, B. E. Launder, F. W. Schmidt, & J. H. Whitelaw.), pp. 7–33. (Springer-Verlag, 1985).

26. Purely multiplicative random *variables* lead to somewhat less extreme log Levy and lognormal distributions (i.e., the logarithms are Levy or Gaussian; their tails are "long" but not "fat"). The enhanced variability of multiplicative processes when compared to multiplicative variables is a result of the singular small-scale limit of the former; it has been theorized in the framework of multifractal phase transitions. Schertzer, D. & Lovejoy, S. Hard and soft multifractal processes. *Physica A* **185**, 187–194 (1992).

27. An entire book by Aitchison and Brown popularized the idea that multiplicative random variables could simply be handled by taking logarithms to yield log-normal distributions [Aitchison, J. & Brown, J. A. C. *The Lognormal Distribution, with Special Reference to Its Uses in Economics*. (Cambridge University Press, 1957).]. But as a result of their "pathological" small-scale limiting behaviors, this reasoning doesn't apply to multiplicative cascade processes (Box 2.2).

28. Mandelbrot, B. B. The variation of certain speculative prices. *J. Business* **36**, 394–419 (1963).

29. Mandelbrot, B. B. How fractals can explain what's wrong with Wall Street: The geometry that describes the shape of coastlines and the patterns of galaxies also elucidates how stock prices soar and plummet. *Scientific American* **February**, 70–72 (1999).

30. Taleb, N. N. *The Black Swan: The Impact of the Highly Improbable*. (Random House, 2010).

31. It is one minus the usual cumulative distribution function that instead accumulates probability from the smallest value.

32. Richardson, L. F. The supply of energy from and to atmospheric eddies. *Proc. Roy. Soc. A London* **A97**, 354–373 (1920).

33. It is equal to the square of the Brunt-Väiäsala frequency.

34. So much so that several "calmer" versions were invented to yield less variable results in the hope that these would be more meaningful.

35. Reproduced with permission from Schertzer, D. & Lovejoy, S. In: *Turbulent Shear Flow* (eds. L. J. S. Bradbury et al.), pp. 7–33. (Springer-Verlag, 1985).

36. Endlich, R. M., Singleton, R. C., & Kaufman, J. W. Spectral analyses of detailed vertical wind profiles. *J. Atmos. Sci.* **26**, 1030–1041 (1969). Endlich, R. M. & Mancuso, R. L. Objective analysis of environmental conditions associated with severe thunderstorms and tornadoes. *Monthly Weather Rev.* **96**, 342–350 (1968).

37. Obukhov, A. Effect of Archimedean forces on the structure of the temperature field in a turbulent flow. *Dokl. Akad. Nauk SSSR* **125**, 1246 (1959). Bolgiano, R. Turbulent spectra in a stably stratified atmosphere. *J. Geophys. Res.* **64**, 2226 (1959).

38. Nastrom, G. D. & Gage, K. S. A first look at wave number spectra from GASP data. *Tellus* **35**, 383 (1983).

39. In terms of the spectral exponents, the three-dimensional and two-dimensional isotropic values are 5/3 and 3.

40. Schertzer, D. & Lovejoy, S. The dimension and intermittency of atmospheric dynamics, In: *Turbulent Shear Flow* (eds. L. J. S. Bradbury, F. Durst, B. E. Launder, F. W. Schmidt, & J. H. Whitelaw), pp. 7–33. (Springer-Verlag, 1985).

41. Schertzer, D. & Lovejoy, S. Generalised scale invariance in turbulent phenomena. *Phys. Chem. Hydrodynam. J.* **6**, 623–635 (1985).

42. The equations governing the basic attributes of a fluid—wind, pressure, and temperature—are well defined, but the equations are not so clear for clouds, precipitation, and radiation, which are heavily "parameterized." In addition, the equations are solved on numerical grids that are far, far larger than the dissipation scale (typically 100 km compared to millimeters), so that it is by no means a trivial matter to solve them numerically (see Chapter 5).

43. The scaling symmetries of fluid equations are at the root of the many "similarity" laws of classic fluids mechanics. Sedov, L. I. *Similarity and Dimensional Methods in Mechanics*. (Academic Press, 1959). Schertzer, D. & Tchiguirinskaia, I. Multifractal vector fields and stochastic Clifford algebra. *Chaos* **25**, 123127 (122015), doi:10.1063/1.4937364 (2015).

44. Actually, a slightly simplified set of atmospheric equations was used (the vorticity equations); but, thanks to numerical analysis of the outputs of weather and climate models, there is no doubt about the basic scaling result for NWPs and GCMs (see Chapter 4). Schertzer, D., Tchiguirinskaia, I., Lovejoy, S., & Tuck, A. F. Quasi-geostrophic turbulence and generalized scale invariance: A theoretical reply. *Atmos. Chem. Phys.* **12**, 327–336 (2012).

45. The mathematically rigorous notion of the size of a set is its "measure."

46. The approximately equal symbol ($\approx$) is used because there will generally be unimportant numerical factors. For example, for a circle, if we take Length = Diameter, then its area is $(\pi/4)(\text{Length})^2$, but we just write Area $\approx (\text{Length})^2$.

47. Launched in 2006, CloudSat orbits at an altitude of 700 km and has a radar that has about a 1-km resolution in the horizontal and about 250 m in the vertical. The satellite doesn't scan from left to right, but builds up a vertical section from the profiles below as it revolves in its orbit.

48. It turns out that this is required for modeling Earth's geological strata. See a review in Lovejoy, S. & Schertzer, D. Scaling and multifractal fields in the solid earth and topography. *Nonlin. Processes Geophys.* **14**, 1–38 (2007).

49. Schertzer, D. & Lovejoy, S. Generalised scale invariance in turbulent phenomena. *Phys. Chem. Hydrodynam. J.* **6**, 623–635 (1985).

50. This brings out the fact that our definition of scale and size is somewhat arbitrary. In this case, we obtained the relation $D_{el} = 1 + H_z$ by identifying scale with the horizontal enlargement factor. We could equally well have defined it as equal to the vertical enlargement factor. In the latter case, we would have obtained $\text{Size} = \text{Area} = (\text{Vertical})^{1/H_z} \times (\text{Vertical})$, so that, as before, we would have $\text{Size} = (\text{Scale})^{D_{el}}$, but with $D_{el} = 1/H_z + 1$. The elliptical dimensions in Figure 3.4 would thus be 7/2, 8/3, 9/4, 2, 11/5, and 3. This would be a perfectly valid system of scales and sizes.

51. I have just described in words the anisotropic generalization of the Haussdorf measure. See Schertzer, D. & Lovejoy, S. Generalised scale invariance in turbulent phenomena. *Phys. Chem. Hydrodynam. J.* **6**, 623–635 (1985).

52. Adapted from Lovejoy, S., Tuck, A. F., Schertzer, D., & Hovde, S. J. Reinterpreting aircraft measurements in anisotropic scaling turbulence. *Atmos. Chem. Phys. Discuss.* **9**, 3871–3920 (2009).

53. Adapted from Lovejoy, S., Tuck, A. F., Schertzer, D., & Hovde, S. J. Reinterpreting aircraft measurements in anisotropic scaling turbulence. *Atmos. Chem. Phys. Discuss.* **9**, 3871–3920 (2009).

54. The elliptical dimension over the overall space (if there were no holes) is therefore log15/log5 = 1 + log3/log5 = 1.68. The elliptical dimension of the black areas (the anisotropic fractal Sierpinski carpet) is log12/log5 = 1.54.

55. This figure is an elaboration on the original one that appeared in Schertzer, D. & Lovejoy, S. The dimension and intermittency of atmospheric dynamics. In: *Turbulent Shear Flow* (eds. L. J. S. Bradbury, F. Durst, B. E. Launder, F. W. Schmidt, & J. H. Whitelaw), pp. 7–33. (Springer-Verlag, 1985).

56. The film version of this (and other anisotropic space–time multifractal simulations) can be found at http://www.physics.mcgill.ca/~gang/multifrac/index.htm. Reproduced with permission from Blöschl, G., Thybo, H., Savenije, H., & Lovejoy, S. In: *A Voyage through Scales: The Earth System in Space and Time* (eds. G. Blöschl, H. Thybo, & H. Savenije), pp. 13–18. (Edition Lammerhuber, 2015).

57. Adapted from Lovejoy, S. & Schertzer, D. Towards a new synthesis for atmospheric dynamics: Space–time cascades. *Atmos. Res.* **96**, 1–52 (2010).

58. Dewan, E. & Good, R. Saturation and the "universal" spectrum vertical profiles of horizontal scalar winds in the stratosphere. *J. Geophys. Res.,* **91**, 2742 (1986). These are the atmospheric analogues of waves on the ocean. They have nothing to do with Einstein's undulations of the fabric of space–time itself.

59. In the lower troposphere, the atmosphere is not heated directly by the sun; rather, the sun warms the surface, expanding the nearby air, which then rises and cools by expansion. Above the tropopause, the air is heated directly by ultraviolet radiation being absorbed by ozone, so that the stratosphere is warmer.

60. The smog data have a smaller spheroscale than the CloudSat data. Figure 3.11 graphically shows that it is more typically about 10 cm to 1 m instead of 10 m to 1 km. The reason for the difference isn't clear, except that that the spheroscale is highly variable and the lidar

data were acquired under calm conditions, whereas the CloudSat clouds are associated with more violent conditions.

61. Adapted from Lovejoy, S. & Schertzer, D. Towards a new synthesis for atmospheric dynamics: Space–time cascades. *Atmos. Res.* **96**, 1–52 (2010).

62. Structures 10,000 km in the east–west direction are typically "squashed" to a size about 1.6 times smaller than in the north–south direction, reflecting the importance of the larger north–south temperature gradients. However, there is no evidence of a *systematic* change in this aspect ratio as we move to smaller scales, nor is there a plausible theory that might explain one. Although this statement is true of the data, it turns out that one of the limitations of NWPs and GCMs is that they *do* have *horizontal* stratifications that are presumably spurious. If the east–west direction is taken as the reference, then the structures in the models for north–south direction follow $(\text{North}-\text{South}) = (\text{East}-\text{West})^{H_y}$, with $H_y = 0.80$. For this, and an explanation, see Lovejoy, S. & Schertzer, D. Space–time cascades and the scaling of ECMWF reanalyses: Fluxes and fields. *J. Geophys. Res.* **116**, doi:10.1029/2011JD015654 (2011).

63. Schertzer, D. & Lovejoy, S. Generalised scale invariance in turbulent phenomena. *Phys. Chem. Hydrodynam. J.* **6**, 623–635 (1985). Lovejoy, S. & Schertzer, D. Generalized scale invariance and fractal models of rain. *Water Resourc. Res.* **21**, 1233–1250 (1985).

64. The definition of the unit scale is quite flexible, but not completely so. When combined with the scale-changing rule, it must not imply that any of the balls cross each other. If they did, then at the crossing point, the notion of scale would become ambiguous; it would not be unique.

65. Mathematically, the rule telling us how to make an infinitesimal scale change is the generator of the (Lie) group of scale changes. Up until now, all the blowup rules were matrix multiplications and they were linear. The anisotropic blowup was defined by the scale change ratio to the power of a matrix, the latter being the "generator."

66. The blowdown is based on the matrix $\begin{pmatrix} 1.3 & -1.3 \\ 0.3 & 0.7 \end{pmatrix}$. Adapted from Lovejoy, S. & Schertzer, D. Course Notes: Multifractals in Geophysics, AGU-CGU-MSA Spring Meeting, May 11, Montreal (1992).

67. The matrices used are $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1.35 & 0 \\ 0 & 0.65 \end{pmatrix}$, $\begin{pmatrix} 1.35 & 0.25 \\ 0.25 & 0.65 \end{pmatrix}$, and $\begin{pmatrix} 1.35 & -0.45 \\ 0.85 & 0.65 \end{pmatrix}$ (upper left to lower right).

68. The matrix used here was $\begin{pmatrix} 1.1 & 0.02 \\ 0.18 & 0.9 \end{pmatrix}$.

69. Both views reproduced from Lovejoy, S. & Schertzer, D. In: *Nonlinear Dynamics in Geophysics* (eds. J. Elsner & A. A. Tsonis), pp. XX–XX. (Elsevier, 2007).

70. Reproduced from Lovejoy, S., Tuck, A. F., Schertzer, D., & Hovde, S. J. Reinterpreting aircraft measurements in anisotropic scaling turbulence. *Atmos. Chem. Phys. Discuss.* **9**, 3871–3920 (2009).

71. Each simulation was made using $256 \times 256 = 65{,}536$ random numbers to form the "subgenerator." Only a single value was boosted artificially to obtain the simulation shown here.

72. Both views reproduced from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

73. The study of radiative transfer in multifractal clouds is in its infancy. See, however, the following works: Naud, C., Schertzer, D., & Lovejoy, S. Fractional Integration and radiative transfer in multifractal atmospheres. In: *Stochastic Models in Geosystems* (eds. W. Woyczynski & S. Molchansov), pp. 239–267. (Springer-Verlag, 1997). Schertzer, D., S., Lovejoy, F. Schmitt, C. Naud, D. Marsan, Y. Chigirinskaya, & C. Marguerit. New developments and old questions in multifractal cloud modeling, satellite retrievals and anomalous absorption. In: *7th Atmos. Rad. Meas. (ARM) Meeting* (ed. US Department Energy), pp. 327–335. (US Department of Energy, 1998). Watson, B. P., Lovejoy, S., Grosdidier, Y., & Schertzer, D. Multiple scattering in thick multifractal clouds part I: Overview and single scattering in thick multifractal clouds. *Physica A* **388**, doi:10.1016/j.physa.2009.05.038 (2009). Lovejoy, S., Watson, B., Grosdidier, Y., & Schertzer, D. Scattering in thick multifractal clouds, part II: Multiple scattering. *Physica A* **388**, 3711–3727, (2009).

74. (A) through (E) reproduced from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

75. Reproduced from Lovejoy, S., Schertzer, D., Lilley, M., Strawbridge, K. B., & Radkevitch, A. Scaling turbulent atmospheric stratification, part I: Turbulence and waves. *Q. J. Roy. Meteorol. Soc.* **134**, 277–300 (2008).

76. Both views reproduced from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

77. The GSI notion of scale is generally *not* a metric. It is *not* a distance in the mathematical sense.

78. Both views reproduced from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

79. Lovejoy, S. & Schertzer, D. Scale, scaling and multifractals in geophysics: Twenty years on. In: *Nonlinear Dynamics in Geophysics* (eds. J. Elsner & A.A. Tsonis), p. 311–337. (Elsevier, 2007).

{ 4 }

# The weather: Don't mind the gap

**4.1** A convenient gap

"This afternoon, the sky will start to clear, with cloud shreds, runners, and thin bars followed by flocks."

If Jean-Baptiste Lamarck (1744–1829) had had his way, this might have been an uplifting early-morning weather forecast announcing the coming of a sunny day.[a] Unfortunately for poetry, in 1803, several months after Lamarck proposed this first cloud classification, the "namer of clouds," Luke Howard (1772–1864), introduced his own staid Latin nomenclature that is still with us today and includes terms such as "cumulus," "stratus," and "cirrus." Howard not only had a more scientific-sounding jargon, but was soon given publicity in the form of a poem by Goethe; Lamarck's names didn't stand a chance.

For a long time, human-scale observation of clouds was the primary source of scientific knowledge of atmospheric morphologies and dynamics. This didn't change until the appearance of the first weather maps based on meager collections of ground station measurements around 1850. This was the beginning of the field of "synoptic" (literally "map-scale") meteorology. Under the leadership of Wilhelm Bjerknes (1862–1951), it spawned the Norwegian school of meteorology that focused notably on airmasses, the often sharp gradients between them called "fronts," and the stability of the airmass interfaces. This was the dominant view when, in the mid 1920s, Richardson proposed his scaling 4/3 diffusion law. The spatial resolution of these "synoptic-scale" maps was so low that features smaller than 1,000 kilometers or so could not be discerned. Between these and the kilo-metric human "microscales," virtually nothing was known. Richardson's claim that a single scaling law might hold from thousands of kilometers down to millimeters didn't seem so daring. Not only was it compatible with the scale-free equations that he had elaborated, but also there were no scalebound paradigms to contradict it.

---

[a] English translations fail to do justice to Lamarck's poetry. Here are some of his cloud types in the original French: *brumeux, en voiles, lambeaux, boursouflés, attroupés, en balayures, en barres, pommellées, coureurs, diablotins, groupes*, and *en montagnes*.

By the late 1940s and '50s, the development of radar finally opened a window onto the intermediate range. During the second world war, the first radars had picked up precipitation as annoying noise that regularly ruined the signals. In 1943, in an attempt to understand the problem better, the Canadian Army Operational Research Group initiated "project stormy weather." After the war, the team— headed by John Stuart Marshall—set up the Stormy Weather Group at McGill University, which—thanks to the "Marshall-Palmer relation[b]"—soon established the quantitative basis for interpreting radar precipitation scans: the famous "Z-R" relation (reflectivity–rain rate).[c] Beyond this quantification of precipitation, the key advance of radar was the ability to image the first weather patterns in the range of 1 to 100 km in size—the discovery of structures and motions in the middle ("meso") scales between the human micro- and the synoptic-map scales.[1]

As this window opened onto the mesoscale, the path pioneered by Richardson's statistical theories of turbulence was rapidly advancing. The idea of turbulence theory was to derive high-level statistical laws governing the collective statistical behavior of strongly nonlinear flows, such as those in the atmosphere, where the nonlinear terms were typically a thousand billion times larger than the linear ones.[2] To make progress, three important simplifications were made. First, only incompressible fluids were considered. Because gravity acts on density varia- tions, this had the effect of eliminating the main real-world source of anisotropy and stratification at the very outset.[3] Second, boundaries and walls, and—for the atmosphere—Earth's surface and north–south temperature gradients are also sources of anisotropy. Therefore, an additional assumption of statistical isotropy

---

[b] This is still the name used by meteorologists for the humble exponential distribution of raindrops as functions of drop size. In 1948, Marshall's graduate student Walter Palmer had used chemically coated blotting paper to relate the size of a drop to the diameter of a "blot." Marshall and Palmer had used many such small pieces of blotting paper, placed in the bottom of jars, to establish the rel- ative number of small and large drops. This information was needed to interpret radar microwave backscatter. But they had assumed the drops were distributed uniformly in space whereas—thanks to turbulence—they were in fact distributed in a hierarchical (cascade)-like manner. Forty years later, using a huge (128- × 128-cm) piece of blotting paper, a student and I recalibrated similarly coated paper in the same McGill staircase, only this time showing that the spatial distribution of drops was not uni- form, but rather a fractal set. See Marshall, J. S. & Palmer, W. M. The distribution of raindrops with size. *J. Meteorol.* **5**, 165–166 (1948); Lovejoy, S. & Schertzer, D. Fractals, rain drops and resolution dependence of rain measurements. *J. Appl. Meteorol.* **29**, 1167–1170 (1990).

A decade later, images each containing tens of thousands of drops in volumes of 10 m³, were analyzed using stereophotography. This confirmed that, as a result of turbulence, Marshall's homo- geneity assumption is only valid up to about 40 to 50 cm, not up to kilometers, as is still routinely assumed. See Desaulniers-Soucy, N., Lovejoy, S., & Schertzer, D. The continuum limit in rain and the HYDROP experiment. *Atmos. Res.* **59–60**, 163–197 (2001); Lovejoy, S. & Schertzer, D. Turbulence, rain drops and the l^{1/2} number density law. *New J. Physics* **10**, doi:075010.071088/071367-072630/075010/ 075017/075017 (2008).

[c] After Marshall retired in 1977, my PhD supervisor, Geoff Austin, succeeded him as leader of the Stormy Weather Group and as director of McGill's (later baptized) J. S. Marshall radar observatory, which was then attached to the physics department. When, in 1980, I gave my first seminar on fractal models of rain in the McGill meteorology department, Marshall attended as a still active professor emeritus.

was made: that the flow itself was, *on average*, the same in all directions.[4] This was an approximation believed to hold at small-enough scales and far enough from surfaces. Third, although at any instant in time the actual turbulent flow would be highly variable from one place to another, it was assumed that, on average, the turbulence was the same everywhere—that it was statistically *homogeneous.*

It is important to take a moment to examine the notions of homogeneity and isotropy more closely. In common parlance, something that is homogeneous is spatially uniform—the same everywhere, constant. Similarly, something isotropic is the same in all directions; it is spherically symmetrical. If the atmosphere was literally—in this deterministic sense—both homogeneous and isotropic, then wind, temperature, pressure, and other atmospheric parameters would have identical values everywhere, and so this would be a useless approximation.

The notion of a turbulence that is *statistically* homogeneous and *statistically* isotropic is much more subtle than this. It has to do with the same symmetries—translational and rotational invariance—but over statistical averages. A statistical average is neither a spatial nor a temporal average; rather, it is an average over a *statistical ensemble*. To understand an ensemble, one must imagine reenacting (almost) exactly the same experiment a large number of times under identical conditions. For each experiment, the details of the resulting turbulent flow would be different because infinitesimally small differences are amplified by the strongly nonlinear character of the flow (the "butterfly effect"; see Chapter 7). Statistical averages would then be obtained by averaging the flow over this huge (in principle, infinite) ensemble of experiments.

Each member—"realization"—of such a statistically homogeneous and statistically isotropic ensemble could easily be extremely inhomogeneous in space and could have a strong preferred direction.[d] However, the preferred locations of turbulent "hotspots," or the preferred orientations of vortices, would be different for each experiment, so that the average over all the experiments would be a constant everywhere and would display no preferred direction. The problem with testing this idea empirically is that no one ever performs an infinite number of identical experiments. And when it comes to the weather and climate, there is only one planet Earth (although, in many respects, Mars comes pretty close, as discussed later!). Often, we, somehow, have to figure out what "typical" inhomogeneities and "typical" anisotropies might be expected on single realizations, even of processes that are known to be statistically homogeneous and isotropic. This underlines the importance of multifractals generated by cascades: They can be constructed easily to be statistically homogeneous and isotropic, but nevertheless, on each concrete instance—realization—they are much more wildly variable than anyone had ever imagined!

Although these assumptions may sound academic, they are not unreasonable approximations to appropriately stirred water in a tank—or even the coffee in

---

[d] Indeed, the breakthrough resulting from cascades and multifractals was precisely the understanding that we should expect extreme variations from one realization of a turbulent process to another.

your cup. Of course, in practical terms, it is impossible either to stir your coffee in exactly the same way throughout the cup (homogeneously) or to do so in a way that is the same in all directions (isotropically). However, there are reasonable arguments to the effect that if one was far enough from boundaries and at small-enough scales, these anisotropies and inhomogeneities would no longer be important.[5] This emboldened theorists to apply these ideas to the atmosphere—even if only over limited ranges of scales. Unfortunately for the isotropy assumption, gravity acts at all scales, so that even if the boundaries affect only the nearby flow, and even if the north–south temperature gradients are important only at the largest scales, the presence of gravity *is* sufficient to render the isotropic theories of the atmosphere academic.

We have considered statistical constancy in space and in direction, but what about in time? A vigorous stirring of your coffee might lead to some approximation of statistical homogeneity and isotropy, but if the stirring stopped, then—as a result of friction (viscosity)—the motion would die down. Therefore, an even simpler situation was usually considered: "quasi-steady" homogeneous and isotropic turbulence in which the fluid was stirred constantly so that the stirring energy would, on average, be dissipated as heat at the same rate at which it was input by the stirring.[6]

Because large structures ("eddies") tended to be unstable and to break up into smaller ones, it was enough for the stirring to create large whirls and let the turbulence do the rest: to create smaller and smaller structures until, eventually, dissipation took over. This hierarchical transfer of energy from large to small was what Richardson had referred to in his poem: "the big whirls have little whirls that feed on their velocity"; it was the basic cascade idea.[7] Such a quasi-steady state means, on average, everything is the same at all times. It is an approximation to the temporal equivalent of statistical homogeneity: statistical "stationarity."

The paradigm of "isotropic, homogeneous turbulence" emerged by the end of the 1930s.[8] During this time, Soviet mathematician and physicist Andrey Kolmogorov (1903–1987) was axiomatizing probability theory,[9] thus laying the mathematical basis for the treatment of random processes. By the end of the 1930s, Kolmogorov had begun to turn his attention to turbulence. The breakthrough was the recognition that the key parameter controlling the flow of energy from the large-scale stirring to the small-scale dissipation was the energy rate density[10] (Chapter 1). Using this quantity, one obtains[e] the Kolmogorov law,[11] which relates the turbulent velocity fluctuations across a structure to its scale[f]:

$$(\text{Velocity Fluctuations}) = (\text{Energy Rate Density})^{1/3} \times (\text{Scale})^{1/3}.$$

---

[e] The basic result follows from dimensional analysis, although various arguments and models—such as cascades—are increasingly convincing.

[f] This effectively explained Richardson's 4/3 law of turbulent diffusion, which Richardson had proposed largely on empirical grounds.

Sometimes, scientific ideas are so ripe that they are "in the air." The Kolmogorov law is a classic example. It was discovered independently no less than five times! One of them was at almost exactly the same time—by another Soviet, Alexander Obukhov (1918–1989)[12]—but in the (equivalent) spectral domain, where it has the form $k^{-5/3}$, where $k$ is an inverse length called the "wavenumber" (the spatial equivalent of a frequency). As a consequence, the law is also referred to as the "5/3 law" or the "Kolmogorov-Obukhov" law. During the second world war, scientific exchanges were limited, so the next discovery was made several years later by Lars Onsager (1903–1976) in 1945.[13] Onsager was the first to link the law explicitly to a cascade of energy flux from large to small scales.[g] But Onsager's American publication only a short abstract; it was no more visible than the earlier Soviet articles had been. This led, in 1948, physicists Werner Heisenberg (1901–1976)[1] Carl Freidrich von Weizacker (1912–2007)[14] to their own rediscoveries.

The pattern of independent Soviet and nearly concurrent western discoveries continued with the discovery (in 1951 and 1949, respectively) of the closely analogous turbulent laws of turbulent mixing, the (also scaling) "Corrsin (1920–1986)-Obukhov (1949) law"[15]:

$$(\text{Temperature fluctuations}) = (\text{Turbulent fluxes}) \times (\text{Scale})^{1/3}.$$

This happened again in 1959, with the Bolgiano-Obukhov[16] law for buoyancy-driven turbulence[h]:

$$(\text{Velocity fluctuations}) = (\text{Turbulent fluxes}) \times (\text{Scale})^{3/5}.$$

By 1953, the theory of isotropic homogeneous turbulence had evolved to the point that it was already the subject of the landmark book *The Theory of Homogeneous Turbulence*[17] by George Batchelor (1920–2000). By then, the role of isotropy had subtly changed. Although it had been introduced originally as a way of simplifying theoretical treatments of turbulence, it had now taken on a life of its own. And although the main application of the theory was to the atmosphere,[i] Kolmogorov noted that the rather stringent "inertial range[j]" assumptions that he had used to derive it (including the neglect of gravitational forces) would only be

---

[g] Looking back many years later, in an important update on his law (taking intermittency into account), Kolmogorov explained that—although they had not mentioned it explicitly—during the period 1939 to 1941 both he and Obukhov had been inspired by Richardson's cascades: Kolmogorov, A. N. A refinement of previous hypotheses concerning the local structure of turbulence in viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.* **83**, 349–369 (1962).

[h] See the left–right shift in Figure 3.2.

[i] The theory was only valid at a very large Reynold's number (i.e., strong nonlinearity, only barely attained in huge wind tunnels).

[j] So-called because the law was valid only when the inertial terms in the equations dominated the dissipation/friction terms.

valid up to scales of several hundred meters—a conclusion amplified by Batchelor, who speculated that the range might only be between 100 m and 2 mm.

Writing twenty years later in the influential book *Weather Forecasting as a Problem in Physics*,[18] Andrei Monin (1921–2007) reproduced (with embellishments) Richardson's original 4/3 law figure (Fig. 2.6), commenting—in accord with Richardson—that it "is valid for nearly the entire spectrum of scales of atmospheric motion from millimeters to thousands of kilometers."[19]

Yet, on the opposite page, he claims that the almost identical Kolmogorov law should hold only to about 600 m! In a later publication,[20] the contradiction is noted with the following mysterious explanation: "in the high frequency region one finds unexpectedly, that relationships similar to those valid in the inertial subrange of the microturbulence spectrum are again valid."

Although Richardson had been blissfully ignorant of isotropic theory and had dared to propose that his scaling law would hold over the whole range of atmospheric scales, the nearly equivalent Kolmogorov law was now claimed to be limited to a tiny range. This drastic limitation was neither the result of any evidence nor the discovery of any scale-breaking mechanism. Rather, the restriction and its implied scale break were hypothesized because atmospheric stratification contradicted isotropy. Isotropy had come to dominate the theory, and the reason for its introduction had been forgotten: theoretical simplicity! Rather than finding the best theory to fit reality, the theorists were trying to find realities to fit their theories.

To be savored, the full irony had to wait more than fifty years for the analysis of dropsonde data (Fig. 1.7). It turned out that scientists had been so confident in the isotropy assumption that they hadn't bothered to check the Kolmogorov law in the *vertical* direction. And when they finally did so, at the end of the 1960s (see Chapter 3), the contrary Bolgiano-Obukhov law resulted. This awkward experimental result had simply been ignored rather than investigated and explained.[k] Isotropy was too beautiful to be disturbed by ugly data!

Figure 4.1 shows the analysis of the dropsonde data. The change (difference) of the horizontal wind was calculated over layers of increasing thickness, first only for the near-surface region (bottom of the plot), then from the surface to higher and higher altitudes (with the upper curves offset for clarity). The clusters of points show the result using all the layers between the ground and 12.8 km (roughly the tropopause). For all altitude ranges, one obtains nearly perfect straight lines, indicating scaling over its internal layers, covering layers with thicknesses ranging from 5 m to nearly 10 km. Even at small scales, the Kolmogorov law (the line with the 1/3 slope at the bottom of Fig. 4.1, next to the letter K) is completely unrealistic, with the real data being very close to the Bolgiano-Obukhov law (slope 3/5).

---

[k] Actually, their confidence in isotropy was so strong that, often, they didn't even bother checking it in the horizontal. They measured the spectrum of a time series taken at a fixed location and then converted the result from time to space, as described later.
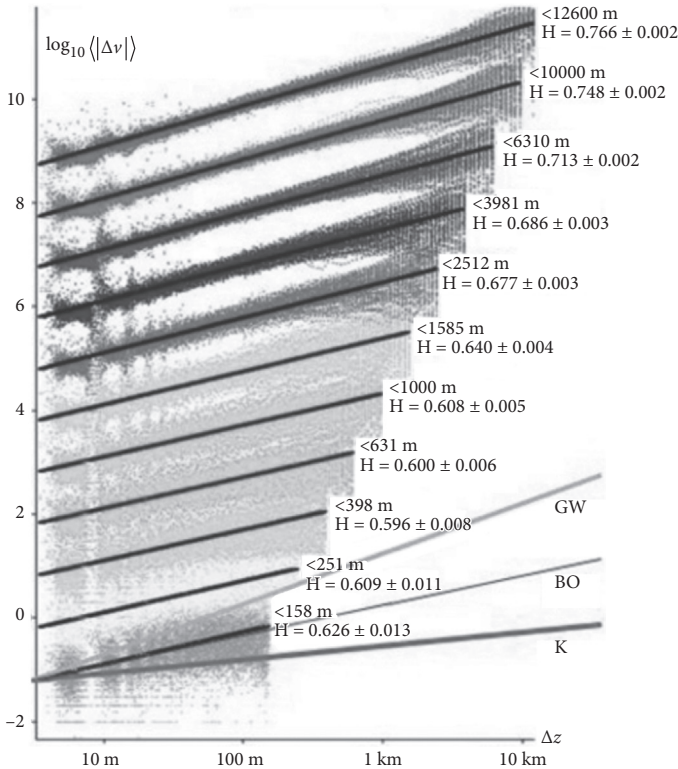
FIGURE 4.1 *The average mean absolute difference in the horizontal wind from 238 dropsondes over the Pacific Ocean taken in 2004. The data were analyzed over regions from the surface to higher and higher altitudes. From bottom to top, the different lines are each separated by a factor of 10 for clarity. Layers of thickness $\Delta z$, increasing from 5 m to the thicknesses spanning the region, were estimated, and lines were fit corresponding to power laws with the exponents as indicated. At the bottom, reference lines with slopes 1/3 (Kolmogorov, K), 3/5 (Bolgiano-Obukhov, BO), and 1 [gravity waves (GW) and quasi-geostrophic turbulence] are shown for reference.[21]*

Although the slopes in the figure increase a little at higher altitudes, even the theories predicting a slope of one can be rejected (see the line marked GW, for "gravity waves"). This slope is predicted by gravity wave theories[22] as well as by Charney's quasi-geostrophic turbulence theory, both of which are seen to be quite unrealistic. The conclusion from the data is unequivocal: The original (isotropic) Kolmogorov law simply does not hold anywhere in the atmosphere (unless it is hiding at scales less than 5 m!). Kolmogorov's and Batchelor's speculation that the Kolmogorov law would hold up to hundreds of meters was doubly wrong. In reality, it holds much *less* in the vertical, but much *more* in the horizontal: up to planetary scales (as discussed later).

A corollary fundamental discovery is that the atmosphere is divided into a fractal hierarchy of stable and unstable layers. The traditional (low-resolution) view is that the atmosphere was generally unstable near the surface and then stable at higher altitudes. However, the high-resolution dropsondes showed that, within

each apparently stable layer, there were unstable sublayers; and within the unstable sublayers, there were stable subsublayers; and so on.[23] This discovery proved to be difficult for various theories (in particular, of the propagation of gravity waves) that assumed that wide, stable layers existed.[24]

## 4.2 ~~The standard two-dimensional/three-dimensional model~~ and Richardson's posthumous vindication

By the mid 1950s, empirically based synoptic-scale meteorology had already relegated the microscales to mere turbulence, but this had been done mostly for practical reasons. Similarly, the new mesoscale was viewed pragmatically as the connection between the two, while simultaneously promising a better understanding of thunderstorms and other previously inaccessible meteorological phenomena. The emerging synoptic-, meso-, and microscale regimes were thus not ordained theoretically; rather, they were practical distinctions awaiting theoretical clarification. Yet the theorists were loath to drop their isotropy assumptions and were happy to find convenient justifications for dividing up the range of scales into small-scale isotropic three-dimensional turbulence and something stratified—albeit not yet clearly discerned—at the larger scales. It was already tempting to knit all this together and to identify the microscales with three-dimensional isotropic turbulence, and the weather with a different, larger scale stratified turbulence.

That the larger stratified scales might be a different type of turbulence was already suggested by the discovery made by Fjorthoft (1953) that completely flat, two-dimensional turbulence was fundamentally different from three-dimensional isotropic turbulence.[1] Although Fjorthoft was cautious in interpreting his results in terms of real atmospheric flows, the seed had been planted for the isotropic two-dimensional/three-dimensional model that followed fifteen years later.

This was the situation when Panofsky and Van der Hoven began their famous measurements of the wind spectrum that they published between 1955 and 1957.[25] At this point, wind data at subsecond scales for durations of minutes had already confirmed the Kolmogorov law,[m] but data were lacking at the longer timescales. Given the lack of computers, the researchers averaged their data at ten-second intervals using "eye averages" and collected data in this way for an hour or so. The spectrum was then calculated laboriously by hand. Last, knowing the average wind speed allowed the scientists to make a rough conversion from time to space. For example, if the average wind over one minute was 10 m/s, then the variability at one minute was interpreted as information about the variability at spatial scales of $10 \times 60 = 600$ m. To investigate the mesoscale between 1 km and 100 km, such data were needed spanning periods of minutes to several hours.

The new element was the use of lower resolution series that could be eye-averaged at five-minute resolutions and that lasted several days. When the

---

[1] This was not yet the discovery of Robert Kraichnan's law of two-dimensional turbulence. Kraichnan, R. H. Inertial ranges in two-dimensional turbulence. *Phys. Fluids* **10**, 1417–1423 (1967).

[m] The confirmation was indirect because the measurements were in time, not in space.

spectrum from this analysis was plotted on the same graph as a one-minute spectrum that had been taken from a completely different experiment under different conditions, Panofsky and Van der Hoven discovered there was a dearth or "gap" in the variability, roughly in the range of about 1 to 100 km, centered on timescales of thirty minutes, corresponding to 10 km. In the authors' words: "The spectral gap suggests a rather convenient separation of mean and turbulent flow in the atmosphere: flow averaged over periods of about an hour . . . is to be regarded as 'mean' motion, deviations from such a mean as 'turbulence.'"[26]

The mesoscale gap was born.

However, the first (1955) article was based on a single location and on only two experiments, and the figures were not very convincing. This led Van der Hoven to perform another series of four experiments that produced what later became the iconic mesoscale gap spectrum (Fig. 4.2). The gap cleanly and "conveniently" separated the synoptic weather scales from small-scale turbulence. With the development of the first computer weather models with resolutions—even today— that don't include the microscales, this gap became even more seductive. At first, it justified simply ignoring these scales; later, it justified "parameterizing" them. The gap idea was so popular that Van der Hoven's spectrum was reworked and republished many times, notably in meteorological textbooks throughout the 1970s. Soon, the actual data points were replaced with smooth artists' impressions (as in Fig 4.2), thus inadvertently hiding the fact that his spectrum was actually a composite taken under four different sets of conditions. Even today, his article is still cited frequently and approvingly.[27]
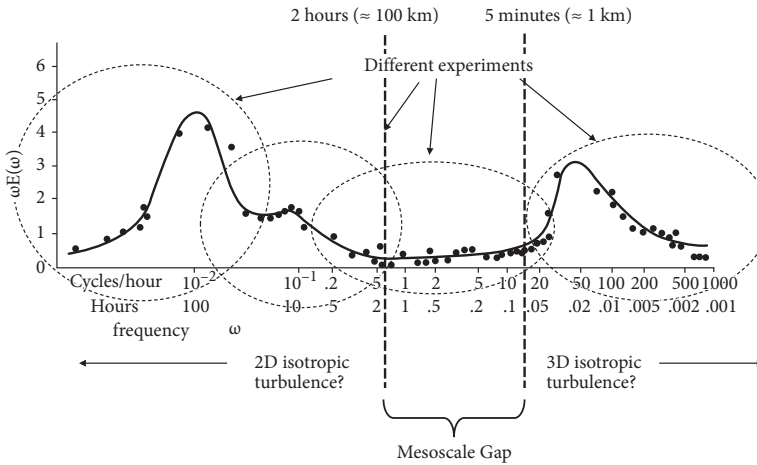


FIGURE 4.2 *The famous "mesoscale gap" is shown between the two dashed vertical lines. The ellipses show the rough ranges of the four experiments that were combined to give the composite spectrum (the actual data points had already been replotted from the original). The temporal limits of the gap (five minutes to two hours) corresponds roughly to 1 km and 100 km, respectively. The leftmost "bump" at about one hundred hours corresponds to the "synoptic maximum," which is none other than the weather–macroweather transition. It appears spuriously as a bump instead of as a smooth transition because the vertical scale is not the spectrum itself [E(ω)], but rather the spectrum multiplied by the frequency ω (for comparison, see, for example, the corresponding frequency range of the more usual log-log spectrum in Fig. 2.3A).[28]*

Yet within ten years, the gap was strongly criticized,[29] with critics pointing out that it was essentially based on a single high-frequency bulge (Fig. 4.2, near scales of one hundred seconds) as a result of the data in that part of the curve being acquired "under near-hurricane" conditions. By the end of the 1970s, the gap idea was scrutinized by satellites that routinely imaged interesting mesoscale features.[n] Practitioners of the nascent field of mesoscale meteorology[30] were skeptical of any supposedly barren "gap" that might relegate their entire field to a mere scientific footnote.[o]

Regardless of the gap, the mesoscale itself underwent a transformation. The new developments were of two-dimensional isotropic turbulence by Robert Kraichnan (1928–2000) in 1967,[31] and, in 1971, its extension to "quasi-geostrophic turbulence"[32] by Jules Charney (1917–1981).[33] Although modern data had filled in the gap with lots of structures and variability, these isotropic turbulence theories supported a new interpretation of the mesoscale as a regime transition between isotropic three-dimensional and isotropic two-dimensional (quasi-geostrophic) turbulence, supposedly near the atmospheric thickness, of about 10 km.

Although in some quarters the gap lived on, the development of two-dimensional turbulence changed the focus. Rather than searching for the gap, the new goal was to search for signs of large-scale isotropic two-dimensional turbulence. Such a discovery promised to transform the mesoscale from a barren gap into a regime transition from three-dimensional to two-dimensional turbulence: the site of a "dimensional transition."[34] Following Kraichnan, the two-dimensional isotropic turbulence idea was so seductive that claims of two-dimensional turbulence sprang up almost immediately. Although three-dimensional isotropic turbulence followed the $k^{-5/3}$ law, two-dimensional turbulence was expected to have a large-scale $k^{-3}$ regime,[35] so that anything resembling a $k^{-3}$ regime was considered to be a smoking gun for the purported two-dimensional behavior (recall that $k$ is a wavenumber: an inverse length scale). Even tiny ranges with only two or three data points that were vaguely aligned with the right slope were soon interpreted as confirmation of the theory.[p]

The first experiment devoted to testing the new two-dimensional/three-dimensional model was the EOLE experiment[q] (1974). It dispersed 480 constant-density balloons[36] (at about 12 km in altitude[37]) over the southern hemisphere, effectively an update[38] of methods used by Richardson (see Fig. 2.6). Because Kolmogorov's 5/3 law was essentially the same as Richardson's 4/3 law, the

---

[n] Because of the curvature of Earth, ground-based weather radars start looking above the weather at distances of around 100 km; satellites give a much more satisfactory range of scales.

[o] In his book on mesoscale meteorology, Atkinson approvingly quotes an early critic Gershen Robinson, who noted, "I find it unconvincing the argument that disturbances on scales between the cyclone and the thunderstorm do not exist because we do not see them on synoptic charts" (p. 417). Robinson, G. D. Some current projects for global meteorological observation and experiment. *Q. J. Roy. Meteorol. Soc.* **93**, 409–418 (1967).

[p] Some early authors admitted to "eyeballing" their spectra over a mere factor of 2 in scales to back up such claims: Julian, P. R., Washington, W., M., Hembree, L., & Ridley, C. On the spectral distribution of large-scale atmospheric energy. *J. Atmos. Sci.* **27**, 376–387 (1970).

[q] After the Greek god of wind.

dispersion of the balloons was an indirect test of the former. By then, there had been several confirmations of Richardson's law in the 1950s and '60s (including in the analogously turbulent ocean). However, the original analysis of the EOLE data by Morel and Larchevesque[39] contradicted Richardson's law and concluded that the turbulence in the 100-km to 1,000-km range did *not* follow his law, but instead followed those predicted by two-dimensional turbulence.

Yet even in the mid 1970s, internal discrepancies in the EOLE analysis were noted.[r] More importantly, the EOLE conclusions soon contradicted those of the Global Atmospheric Sampling Program (GASP) (1983) and, later, Measurement of Ozone by Airbus In-service Aircraft (MOZAIC) (1999) analyses that found Kolmogorov turbulence out to hundreds of kilometers.[s] Two decades later, the original (and still unique) EOLE data set was reanalyzed by Lacorte et al.,[40] and it was concluded that the original EOLE conclusions were not founded, and that, on the contrary, the data vindicated Richardson over the range from 200 to 2,000 km.

But the saga was still not over. Strangely, despite their support for Richardson at the largest scales, the reanalysis by Lacorte et al. contradicted him at the *smallest* EOLE scales—from 200 km down to the smallest available scales (50 km)—over which they claimed to have validated the original two-dimensional turbulence interpretation! Nearly a decade later, this conclusion prompted a re-revisit that found an error in this smaller scale analysis, thus eliminating any evidence for two-dimensional turbulence up to the largest scale covered by EOLE—2,000 km—thus (finally!) vindicating Richardson nearly 90 years later![41]

## 4.3  Science: A human enterprise

Science is a quintessentially human activity. At each epoch, it depends on the available technology, the reigning scientific theories, and the key scientific problems—on its historical level of development. Yet it also depends on society's attitude and on its willingness and ability to allocate resources. To understand the trajectory of atmospheric science following the heady nonlinear decade of the 1980s, we need to examine briefly the changing fortunes of fundamental science in scientifically advanced countries. Although the following account is tainted by my own situation in Canada, it reflects the Friedmanite economic policies that have been dominant in advanced industrial nations during the past few decades.

The post second world war élan of scientific optimism was expressed vividly in the title of Vannevar Bush's famous report: *Science, the Endless Frontier*.[t] It was the beginning of the era of "Big Science," of science being harnessed directly by big

---

[r] Between the relative diffusivity and the velocity structure function results.

[s] The interpretation of aircraft data itself turned out to be nontrivial (see Box 4.1).

[t] The title of a report in July 1945 by Vannevar Bush, Director of the Office of Scientific Research and Development, US. Bush, V. *Science: The Endless Frontier: A Report to the President on a Program for Postwar Scientific Research*. (National Science Foundation 1960 [1945]).

business, albeit in the form of a partnership between publicly and privately funded efforts. It was recognized that investment in inappropriate scientific concepts or unrealistic models could squander huge sums of money, and that no matter how urgent a problem might be, its solution required a balance between fundamental and applied research.

It was understood that fundamental research leads to new knowledge and, as Bush observed, "it creates a fund from which practical applications of knowledge must be drawn."[42]

By the 1990s, in the wake of the atom bomb, DDT, Bhopal, Chernobyl, and other human and environmental catastrophes, the mood had changed radically and technology was no longer seen as a progressive, beneficial force. In the words of a prominent cultural critic, the "look" of the future "is a survivalist one, governed by the dark imagination of technological dystopias."[43] In his ~~book~~ *Science: The End of the Frontier?*[44] Nobel Prize winner Leon Lederman counterpointed Bush and described a reigning "mood of deep depression in the research community,"[45] concluding "that it raises serious questions about the very future of science in the United States."[46]

Then, in 1993, signaling the end of a half century of support for pure research, the US Congress canceled the Superconducting Super Collider.[u]

Disenchantment with the technological fruits of science was only one contributor to the unraveling of the "endless frontier." A more pernicious factor was the growing influence of cognitive relativism that undermined science's core intellectual basis. Cognitive relativism is the philosophical position that claims that all truth—including scientific truth—is relative and ultimately no more than a subjective system of belief. Accordingly, scientific theories of global warming are intrinsically no more valid than those of professional climate skeptics. Each simply represents a system of beliefs shared by different communities. Neither "belief" has any greater or lesser worth. Although cognitive relativism can be traced back to the Greek Sophist philosophers, during the 1990s, in the wake of the academic fashion for postmodernism, it gained a new lease on life[v]. Taking advantage of experience, tactics, and even personnel from the tobacco-denial industry, the oil industry launched its own climate denialist machine.[47,w] Then, as today, it gives

---

[u] According to analysis, since 1992, the use of the term "basic research" in the journals Science and Nature has steadily declined. See: Pielke, R. J. In retrospect: Science: The endless frontier. *Nature* **466** (7309), 922–923 (2010).

[v] For a more thorough discussion, see Gillott, J. & Kumar, M. *Science and the Retreat from Reason*. (Merlin Press, 1995).

[w] The origins of industrial climate denial can be traced to a speech made by Exxon Chief Executive Officer Lee Raymond to the world Petroleum Congress in October 1997, on the eve of the Kyoto meeting that gave rise to the Kyoto Accord. The subsequent American Petroleum Industry "Global Climate Science Communications Plan" marked the beginning of an all-out campaign against climate science. Ironically, it also marked a reversal for Exxon, which—until then—had funded its own genuine climate science research. For more details, see Otto, S. L. *The War on Science: Who's Waging It, Why It Matters, What We Can Do About It.* (Milkweed Editions, 2016).

license to the media to present "both sides" of the climate debate and, more generally, it provides a context in which "alternative facts" can flourish.

Cognitive relativism attacked science's intellectual foundation; its ability to produce truths of general, universal validity; to produce knowledge. By the mid 1990s, this was already perceived as a veritable "war on science,"[x] skillfully analyzed in Sean Otto's book with the same title.[48] In climate science, rather than asking, "Is Earth warming?" and "Why is its temperature rising?" postmoderns ask, "Who is making the claim?" and "Why are *they* making that claim?"[y] They will further draw attention to the fact that the scientists making these claims are mostly white Anglo-Saxon men, the same group that holds a position of power, with claims that are therefore suspect and probably biased. In Canada, the situation has now evolved to the point where some granting agencies encourage environmental scientists to incorporate "indigenous knowledge" in their proposals, effectively placing the latter on a par with scientific knowledge.[z]

None of this would have mattered so much if fundamental science had continued to be highly profitable, but by this time it had become increasingly expensive and—worse still—accountants were unable to determine the corresponding rates of return on investment.[aa]

Corporations were happy to let academic or other publicly funded institutions[bb] pick up the bill. By the 1990s, high costs and risks had created a situation in which only a handful of giant corporations still carried out any fundamental research.

---

[x] Calling postmodern philosophy "the higher superstition," Levitt and Grossman defended the objectivity of scientific knowledge, but lumped inaccurately all the academic Left together with its extreme postmodern wing. Gross, P. R. & Levitt, N. *Higher Superstition: The Academic Left and Its Quarrels with Science*. (Johns Hopkins University Press, 1994). This prompted self-described Left-wing physicist Alan Sokal to publish his famous hoax "Transgressing the boundaries: Toward a transformative hermeneutics of quantum gravity" in the postmodern journal *Social Text*. See Sokal, A. Transgressing the boundaries: Toward a transformative hermeneutics of quantum gravity. *Social Text* **46/47**, 217–252 (1996). The text was, in fact, a hilarious mishmash of scientific gobbledygook purporting to show that reality was no more than a social construct. Through careful use of postmodern language, the editor was tricked into thinking the article was a genuine contribution to a new postmodern physics. In his later book Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science [Sokal, A. & Bricmont, J. *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*. (Picador, 1998).], Sokal and Bricmont scathingly criticized the bogus use of physics concepts by key postmodern thinkers, including Lacan, Latour, Deleuze, and Baudrillard: The emperor had no clothes.

[y] As described in Section 6.1, climate skeptics routinely accuse climate scientists of distorting their research for pecuniary gain: fat checks from granting agencies.

[z] In Chapter 7, I describe research that uses the atmosphere's memory to make long-range (macroweather) forecasts and climate projections (Sections 7.2 and 7.3). In an attempt to obtain funding, I applied for a grant from a federal source to develop these techniques in the Canadian north. Although the scientific aspect of the proposal was judged "excellent," funding was denied because of weak "partnerships." This included a failure to integrate Inuit knowledge into the stochastic forecast technique.

[aa] Although much attention has been paid to the various political, social, and ideological aspects of the war on science, there are surprisingly few analyses of trends in funding of fundamental science; my account is somewhat subjective.

[bb] In the United States, especially the military.

Marking the end of an epoch, in 1996, even the famous Bell labs were sold off,[cc] whereas others were downsized or refocused toward more practical matters.[dd] Today, the fundamental research required for technological advance is virtually entirely publicly funded.

This evolution came at a price. Until then, fundamental-sector scientists had been given free rein to investigate the areas of greatest scientific significance; it was "curiosity driven." Now, all research required an economic justification, and curiosity and kindred terms became the kiss of death. Businesses lobbied governments for direct control over public research, over both its priorities and the management of its funds. Violating its very nature as a long-term enterprise, funding of fundamental science was retargeted toward short-term corporate gain, with public research agencies reoriented accordingly. To smooth the transition and to reassure the public, officials mouthed the new mantra of "excellence" according to which special interests were also excellent and where doing more with less was particularly excellent.[ee]

At the same time, official research and development figures were doped by including tax shelters for businesses claiming to invest in high tech,[ff] effectively hiding this pirating of public funds from close scrutiny.

Concomitant with the industrial focus was a growing disinterest—even by scientists—in fundamental issues, including those that the nonlinear revolution had promised to solve. In atmospheric science, resources were tightly concentrated on the development of NWP models and GCMs,[gg] which included satellites and other NWP and GCM inputs. The only acceptable justification for funding became the promise of improving model performance. In the past, it had been possible to obtain support for an applied science project and—thanks to deliberately loose controls—scientists would regularly siphon off some of the funds (illicitly) to do "real science." But in the brave new world of excellence, sponsors required excellent accountability. Not only were research priorities imposed from

---

[cc] In 1996, the parent company, AT&T, sold Bell labs to Lucent Technologies. In 2006, Lucent was acquired by Alcatel and in, 2015, by Nokia.

[dd] In my case, Canada's branch plant economy had never enjoyed much corporate research.

[ee] I can still remember the time before the excellence mania, when research was judged by peers on its scientific merit with at least some effort to make this objective. During the 1990s, industrial managers increasingly replaced the peers and substituted their own subjective notions of excellence that they could define and evaluate at will.

[ff] Specific Canadian examples that stick in my memory include a case in which hundreds of millions of dollars in tax relief was given to banks that had "invested" in science and technology simply by upgrading their office equipment. This was during the early 1990s, when the banks were doubling their profits every year. Another example from the same period was of tax relief given to a company "researching" new flavors of beer. This type of accounting trick allowed the government to boast of stable levels of support for science.

[gg] In Canada, for example, the staff of the federal weather office, the Atmospheric Environment Service (now the Meteorological Services of Canada, part of Environment Canada) was cut by nearly 50% during the 1990s, and this included the elimination of a small fund used to support academic research.

without, but every dollar had to be spent exactly as specified in increasingly de-tailed submissions often written years earlier.[hh] Rather than helping their faculty juggle grant deadlines and applications creatively to allow them to focus on the science, university accounting departments administering the grants played an in-creasingly police role, protecting the sponsors' money from being spent irrespon-sibly in advancing science, rather than following the sponsors' dictates. Academic scientists were gradually being transformed into cheap labor.

But technology continued to advance, and rapidly increasing computer sizes and speeds, improved algorithms, and mushrooming quantities of remotely sensed and other data ensured that the 1990s were a golden age for atmospheric science. Fundamental NWP and GCM issues were also being resolved. In partic-ular, advances in data assimilation opened the doors to the widespread "inges-tion" of satellite and other disparate and hitherto under- or unexploited sources of data,[49] so that by the decade's end, weather forecasts had improved significantly. During this decade, ways of using the usual deterministic forecasting procedure to generate an ensemble of slightly perturbed forecasts were being developed thanks to the pioneering work of Eugenia Kalnay and Tim Palmer, the "breeding vector" and "singular perturbation" techniques. This marked the beginning of Ensemble Forecasting Systems (EFS), the first step towards truly stochastic forecasts. It fi-nally allowed weather forecasts to be assigned probabilities of occurrence and hence quantitative uncertainties.

By the 2000s, NWPs and GCMs increasingly appeared to be atmospheric science's way of the future—indeed, the *only* way. In numerical models, scale and scaling could have intruded in the choice of model grid size; but, in prac-tice, this was determined by computer technology. The main choice—the ratio of the number of vertical to horizontal pixels—was made by empirical experience rather than by using theory. The fact that it ended up following the 23/9 D model (Fig. 1.9) was by empirical necessity, rather than by theoretical understanding.[50] Scale and scaling were abstract and seemed unnecessary.[ii]

The divorce between practical atmospheric science and turbulence theory was hardly new. Atmospheric science had always suffered from the gulf between the idealized smooth, calm models concocted by theoreticians[jj] and real-world wild irregularity.[kk] During the 1970s, a popular adage was: "No one believes a

---

[hh] Because the outcome of research is intrinsically unpredictable, the best strategy for scientists was to somehow get ahead of the game and ask for funding for projects that they had already performed, thus guaranteeing they could fulfill the precise terms of the grant.

[ii] By the end of the 1990s, the behavior of the atmosphere at scales less than those resolved by the grids (corresponding to 100 km or so) started to be parameterized stochastically using random num-bers ("stochastic parametrization"). But even this was done in a relatively ad hoc way, without regard to the relevant (cascade) theory.

[jj] The first (potentially realistic) strongly nonsmooth, violently variable models came out of only one of the strands of turbulence theory—the scaling strand: multifractals.

[kk] This was because, unlike scaling approaches, the dominant theories made many conventional smoothness and regularity assumptions that were totally unrealistic.

theory except the person who invented it. Everyone believes the data except the person who took them." In the new ambiance, cynicism about theory was especially strong. When discussing atmospheric scaling in the early 1990s, a prominent colleague commented: "If no progress is made for long enough, the problem is considered solved and we move on."[51]

It was thus not surprising that when the budgetary screws were turned, this ambient negativism hit the nonlinear revolution and conventional turbulence theory alike: Interest was scant, and funding even more so. Theory of any kind was increasingly seen as superfluous; it was either irrelevant or a luxury that could no longer be afforded. Any and all atmospheric questions were answered using the now-standard tools: NWPs and GCMs.

Unfortunately, these models are massive constructs built by teams of scientists spanning generations. They were already "black boxes," and even when they answered questions, they did not deliver *understanding*. Atmospheric science was gradually being transformed from an effort at comprehending the atmosphere to one of imitating it numerically (i.e., into a purely applied field). New areas—such as the climate—were being totally driven by applications and technology: climate change and computers. In this brave new world, few felt the need or had the resources to tackle basic scientific problems.

The sorry state of scientific research in general—and climate science in particular—was strikingly revealed in June 2017. Only weeks after being elected president of France, Emmanuel Macron loudly launched a French initiative to "Make the Planet Great Again," generously extending the ambitions of the emblematic Reagan–Trump slogan from the United States to encompass the globe. The initiative re-earmarked a modest €60 million in funds.[ll] Beyond stealing the climate limelight, the aim was to incite fifty foreign climate scientists to exile themselves to France. The results were extraordinary. Within a month, more than 2,500 scientists from all over the world (a third of these from the United States) as well as another 8,000 students and "climate entrepreneurs" had filled an online form asking for scientific exile ~~to France~~. "March for Science" organizer Olivier Berné gave this appreciation: "[W]e are irritated because we were expecting improvements in the work conditions of French scientists."[mm] In December, when the names of the first cohort of eighteen scientists were announced,[nn] Patrick Monfort, Secretary General, National Union of Scientific Workers, commented:

> There isn't much enthusiasm for the project. We're bringing our colleagues to France and telling them: "Come and share the means that we don't have." In addition, it's very insulting for French scientists: this project gives the

---

[ll] It later became a French–German initiative and the French government downgraded its contribution to €30 million.

[mm] Larousserie, D. & Roger, S. Climate: 2500 chercheurs prêts à un exil en France [Climate: 2500 Scientists ready for an exile in France]. *Le Monde*, July 10, 2017, p. 10. (my translation).

[nn] Not surprisingly, thirteen of these were from US institutions.

impression that French science is so bad that we have to bring 50 people to raise its level and they will be paid twice as much and they will take our resources. How do you think they will be welcomed in our laboratories? It's a terrible welcome.[oo]

## 4.4  The rise of nonlinear geoscience

So it was that the excitement engendered by the nonlinear revolution during the 1980s faded slowly and, with it, general scientific interest in geosystem scales and scaling. Yet the revolution had succeeded in establishing a beachhead, a community of like-minded scientists organized first in the European Geophysical Society's[pp] (EGS) Nonlinear Processes division[qq] (in 1989), around the *Nonlinear Processes in Geophysics* journal (1994), and a little later in the AGU's Nonlinear Geophysics (NG) focus group (1997). Following a 2009 workshop on geocomplexity,[rr] a dozen scientists published a kind of nonlinear manifesto titled "Nonlinear Geophysics: Why We Need It."[52] It proclaimed:

> [T]he disciplines coalescing in the NG movement are united by the fact that many disparate phenomena show similar behaviours when seen in a proper nonlinear prism. This hints at some fundamental laws of self-organization and emergence that describe the real nature instead of linear, reductive paradigms that at best capture only small perturbations to a solved state of a problem.[53]

It was largely in nonlinear geophysics that evidence for wide-range atmospheric scaling slowly accumulated, notably by the study of radar rain reflectivities[54] and satellite cloud radiances[55] (see, for example, Fig. 4.3). But these analyses were generally restricted to scales smaller than 1,000 kilometers and, crucially, they didn't involve the wind field, which could not be sensed reliably by remote means. For the wind field, the only alternative to aircraft data was the analysis of the outputs of numerical models and, at the time, these didn't have a wide-enough range of scales to be able to settle the issue either.[56]

---

[oo] Patrick Monfort, interviewed by Thomas Baïetto. Projet "Make our planet great again": Les chercheurs américains "vont venir partager notre misère," déplore un syndicat. *France Info*, December 12, 2017. Available at https://www.francetvinfo.fr/monde/environnement/retrait-americain-de-l-accord-de-paris/projet-make-our-planet-great-again-les-chercheurs-americains-vont-venir-partager-notre-misere-deplore-un-syndicat_2509589.html. (my translation).

[pp] Since 2002, following a merger with the European Union of Geodesy, it has become the European Geosciences Union.

[qq] This precocious development was greatly helped by the EGS's visionary director Arne Richter (1941–2015).

[rr] Organized by Qiuming Cheng and myself at York University, Toronto, Canada.
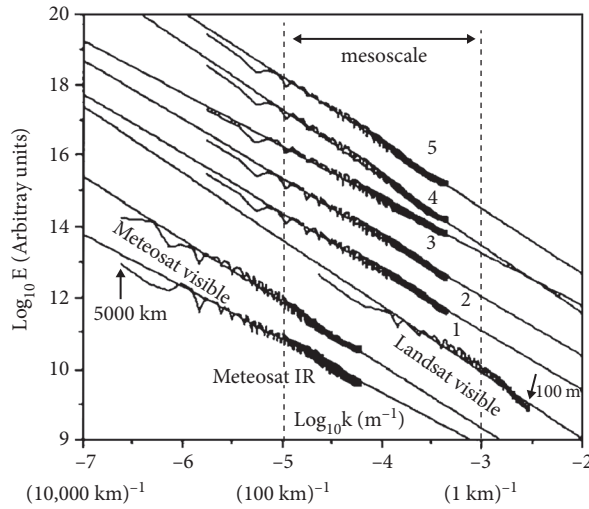
FIGURE 4.3 *Spectra from three different satellites from largely cloudy regions. Meteosat, geostationary, 8-km resolution; Landsat, at 83-m resolution and the numbers 1 through 5 from the NOAA-9 satellite; with channel 1 in the visible; channel 5 in the thermal infrared, and channels 2, 3, and 4 in the in-between wavelengths. Scaling power laws are straight lines on the log-log plot; the range of 1 to 100 km is indicated by "mesoscale" and it shows no signs of a break in the scaling. Overall, the figure covers the range of 166 m to 5,000 km.[57]*

The next major advance appeared during the second half of the 2000s, with the beginning of the widespread availability[ss] of truly global-scale atmospheric data sets—notably, massive archives of satellite data—and they invariably showed excellent wide-range scaling, although, again, not for the hard-to-measure wind field (see Fig 4.4). Also at this time, the numerical models were getting big enough to analyze, and once again, wide-range scaling was found[58] (see the reanalyses, Fig. 4.4, lower right). Regardless of whether meteorologists liked scaling (or were even aware of it!), their models respected the symmetry extremely well![tt] By 2008, it seemed that the only evidence that apparently contradicted the wide-range scaling hypothesis were aircraft wind spectra (see Box 4.1).

[ss] The key advance was not so much the actual existence of large-scale data sets (this was not so new); rather, it was the rapidly increasing Internet speed that allowed them to be disseminated readily and be available online. For example, the single 30 MByte Landsat image analyzed in Figure 4.3 occupied a full twelve-inch-diameter spool of magnetic tape and had to be ordered and paid for months in advance ($120/tape). In comparison, after it was launched in 1998, the data from the TRMM satellite analyzed in Figure 4.4 (top) took nearly ten years to become freely available on the Internet. But, when it did become available, there were hundreds of thousands of gigabyte images that could be downloaded easily and free of charge.

[tt] But this didn't prevent the continued development of all kinds of simplified scalebound models attempting to understand NWP and GCM outputs more fully.
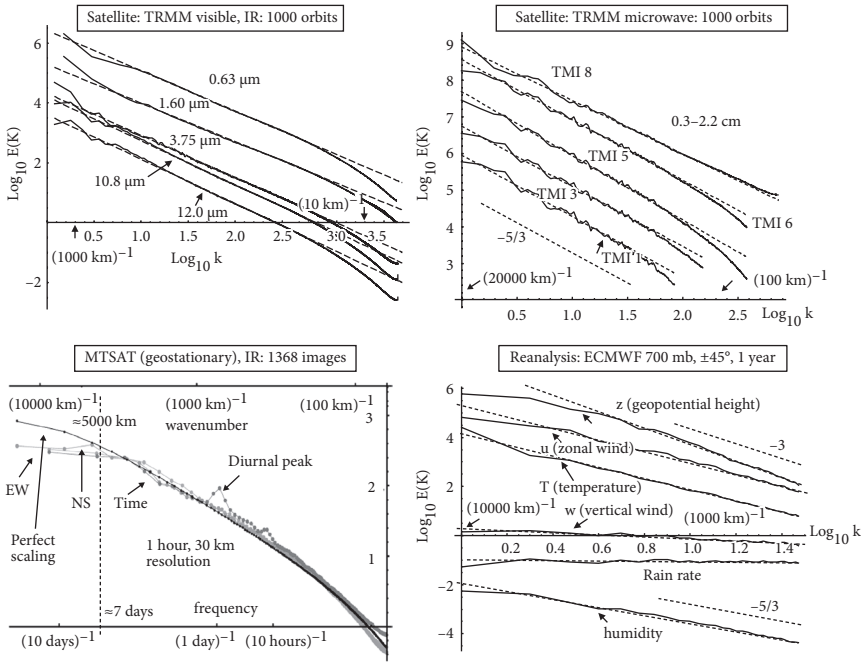
FIGURE 4.4 *(Upper left) Spectra from more than 1,000 orbits made during the Tropical Rainfall Measurement Mission (TRMM) of five channels visible through thermal infrared (IR) wavelengths displaying very accurate scaling, down to scales on the order of the sensor resolution (≈10 km). (Upper right) Spectra from five other (microwave) channels from the same satellite. The data are at a lower resolution and the latter depends on the wavelength. Again, the scaling is accurate up to the resolution. (Lower left) The zonal, meridional, and temporal spectra of 1,386 images [two months of data, September and October 2007] of radiance fields measured by a thermal infrared channel (10.3–11.3 μm) on the geostationary Multifunctional Transport Satellite (MTSAT) over the southwest Pacific at resolutions of 30 km and one hour over latitudes 40°S–30°N and longitudes 80°E–200°E. With the exception of the (small) diurnal peak (and harmonics), the rescaled spectra are nearly identical and are also nearly perfectly scaling. The black line shows exact power law scaling after taking into account various geometrical effects.[59] The vertical axis is $\log_{10}E(k)$ (i.e., the same as on the lower left plot). The bottom horizontal axis is frequency (applicable to the time spectrum), whereas the top horizontal axis is in wavenumbers—spatial frequencies—and is appropriate for the two spatial (east–west and north–south) spectra. (Lower right) Zonal spectra of reanalyses from the European Centre for Medium Range Weather Forecasting (ECMWF), once daily for the year 2008 over the band ±45° latitude.[60]*

BOX 4.1 Aircraft turbulence: It's not just the bumps!

"Fasten your seatbelts. We are expecting turbulence."

On long-haul flights, this is a routine announcement intended for the lay public, yet it conveys a deep-seated misconception about the nature of turbulence. It reinforces the idea that atmospheric air motions are basically smooth (technically,

"laminar"), interspersed occasionally with small, embedded turbulent zones. The subtext is that somehow the "smooth" and turbulent parts of the atmosphere are associated with distinct dynamics. From there, it is a small step to develop separate scientific theories for each.

The imputation of distinct mechanisms to (apparently) distinct phenomena is a species of phenomenological fallacy (Section 3.3.4). Here, the fallacy is making a qualitative distinction between miniscule bumps that are smoothed out by aircraft inertia and stronger ones that are easily discerned. Rather than making a spurious categorization of phenomena according to their spatial extent, it makes a qualitative distinction between phenomena based on their intensities. It has not only contributed to prolonging the life of outdated scalebound approaches to atmospheric dynamics and turbulence, but also is responsible for erroneous interpretations of aircraft measurements on which I elaborate in this box.

Figure 4.5 shows examples of legs of aircraft trajectories over the Pacific.[61] Although this was an instrumented (scientific) aircraft, as with commercial flights, it flew at constant pressure levels: "isobars" (constant here to within 0.1%). One can see that they move up and down, so the isobars are not at all the same as isoheights (levels of constant altitude). To see this more clearly, consider the zoom sequence in Plate 4.1. We see that both the wind speed and the altitude are highly variable at all observed scales. The aircraft trajectory is a fractal, although—fortunately, for our comfort—at scales smaller than a few kilometers, it is primarily smoothed out by the aircraft's inertia.[uu]
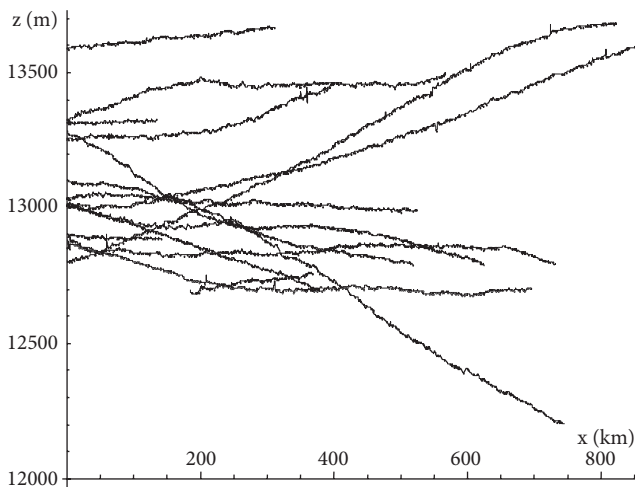


FIGURE 4.5 *Trajectories of an aircraft following lines of constant pressure to with 0.1% near the 200-mb pressure level (height in meters). Data were taken every second (horizontal distance, 280 m). The trajectories are far from constant in altitude (see Plates 4.1 and 4.2).*[62]

This (smoothed) fractality superposed on a gently sloping isobar is by no means the whole picture. From experience, we expect that the aircraft's sudden

---

[uu] If the fractality continued to small scales, it would imply enormous accelerations. Thanks to the inertia of the aircraft, it (usually!) doesn't jiggle up and down very much.

"transitions from quiescence to chaos" (from apparent smoothness to bumpiness) to be episodic (i.e., intermittent). Recall (Fig. 1.6) that an innocuous-looking fractal transect can hide violent spikiness that can be exposed by examining the absolute gradients. Wind is directly related to the turbulent energy flux (i.e., the energy rate density; Section 4.6), and this flux clearly displays its large spatial intermittency (Plate 4.2). From the blowups, we see that, even apparently calm, laminar regions have embedded regions of high activity. As we zoom into smaller and smaller regions, this strong heterogeneity continues in a scaling manner, presumably until we reach the millimetric dissipation scale. This explains why aircraft measurements of wind invariably find roughly Kolmogorov-type (i.e., turbulent) statistics, even in apparently calm regions of the atmosphere. In any event, regions of true laminar atmospheric flow have yet to be documented by actual measurements. It would therefore be a mistake—a phenomenological fallacy—to separate these regions of high and low "turbulent intensities" and associate them with different mechanisms.

Aircraft measurements are often our only direct source of horizontal data about wind, temperature, humidity, and other atmospheric variables, and yet the very turbulence they seek to quantify modifies the trajectories that in turn modify the results of the measurements in ways that we are only now starting to understand. Certainly, we shouldn't be surprised that, by itself, the constant up-and-down intermittent "jiggling" of aircraft taking the wind measurements leads to biases.[63] However, there is an even more important bias caused by the departures of the trajectories from levels of constant altitude. This is because the atmosphere is highly stratified (Chapter 3), so that moving up and down only a little bit can lead to much larger variations in the wind than simply moving along in the horizontal direction.[vv]

To understand the consequences of nonflat aircraft trajectories on measurements and to resolve the apparent contradiction between the scaling 23/9 D dynamics and aircraft observations of broken scaling (e.g., Figs. 4.6 and 4.7A), we need only note that, at a critical scale—which depends on aircraft characteristics as well as the turbulent state of the atmosphere—the aircraft "wanders" sufficiently off a constant altitude level so that the wind it measures changes more ~~from~~ a result of the level change than from its horizontal displacement. It turns out that this effect can easily explain the observations. Rather than a transition from characteristic isotropic three-dimensional to isotropic two-dimensional behavior (spectra with transitions from $k^{-5/3}$ to $k^{-3}$, where $k$ is a wavenumber, an inverse distance), instead, one has a transition from $k^{-5/3}$ (small scales) to $k^{-2.4}$ at larger scales (e.g., Figs. 4.6 and 4.7A), the latter being the typical exponent found in the vertical direction [e.g., by dropsondes (Fig. 4.1[64])].

---

[vv] If the turbulence really was isotropic, then the up-and-down movement of the aircraft would not be anything to worry about. It would not give spurious estimates of the scaling exponents. This explains why it was—and still is—ignored!
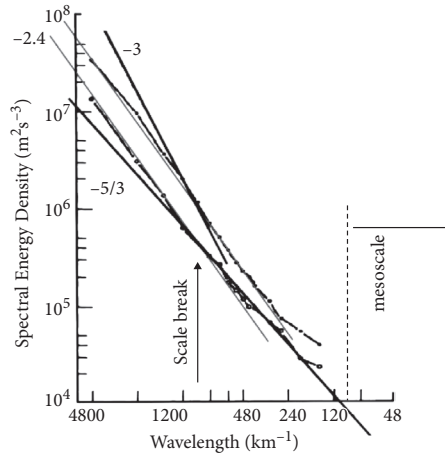
FIGURE 4.6 *The Global Atmospheric Sampling Program (GASP) spectrum of long-haul flights (more than 4,800 km; two data sets are shown), with reference lines corresponding to the horizontal and vertical behavior expected from the two-dimensional/three-dimensional model (slopes indicated). The rough position of the scale break is shown. It is near 1,000 km and is much larger than any possible two-dimensional/three-dimensional transition scale. I have added thin lines with a slope of –2.4 that show the behavior expected if the aircraft spectrum was dominated by the vertical rather than the horizontal displacement of the aircraft. The largest scale in the usual mesoscale range (100 km) is shown at the right.[65]*
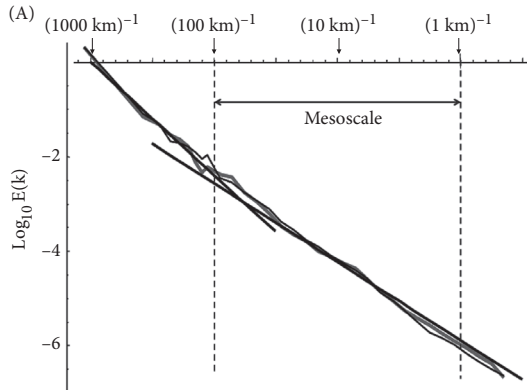


FIGURE 4.7 *(A) Spectra from the Pacific Winter Storm experiment. Included are the averages over twenty-four legs, each 280 m in resolution of length 1,120 km at 200 mb (altitude, ~12 km). The longitudinal and transverse components of the horizontal wind are shown, along with reference lines indicating the horizontal (Kolmogorov) exponent 3 and the vertical exponent 2.4 (close the Bolgiano-Obukhov value 11/5). (B) This figure is the same as (A), except for temperature (T), humidity (h), and log potential temperature[ww] (logθ). A reference line corresponding to the $k^{-2}$ spectrum is shown . The mesoscale (1–100 km) is shown between the dashed lines.[66]*

---

[ww] The log potential temperature is a measure of the entropy of dry air.

(B)

$(1000 \text{ km})^{-1}$    $(100 \text{ km})^{-1}$    $(10 \text{ km})^{-1}$    $(1 \text{ km})^{-1}$

Mesoscale

$-2$

Log$\theta$

h

$-4$

$-6$

T

$E(k) \approx k^{-2}$

Log$_{10}$ E(k)
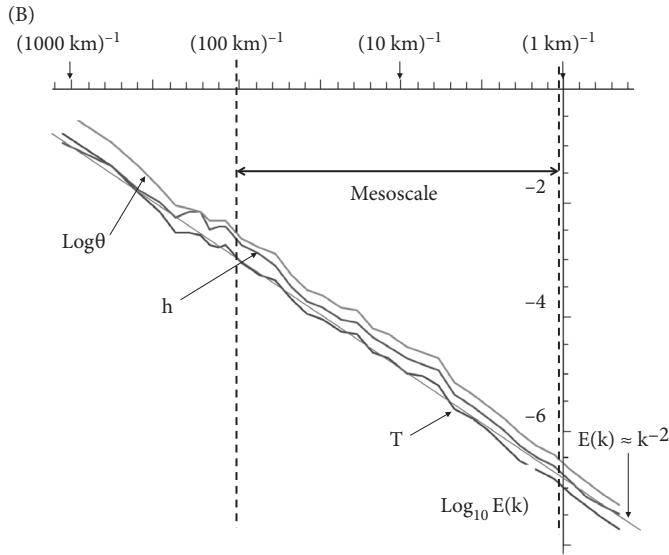
FIGURE 4.7  *Continued*

Even if one accepted the original interpretation of the EOLE experiment in terms of two-dimensional turbulence, it measured the dispersion of balloons, not the wind speeds needed for a ~~direct~~ test of the two-dimensional/three-dimensional model.[67] It was therefore only with the first large-scale aircraft campaigns in the 1980s that the theory could be tested seriously—and this turned out to be the beginning of a multidecadal saga that paralleled the convoluted Richardson's law tale. The first and still most famous of these was the GASP experiment, with data that were analyzed by Nastrom and Gage in a series of works from 1983 to 1986.[68] The dominant interpretation was that the GASP spectrum did indeed show a transition from the Kolmogorov three-dimensional isotropic turbulence to two-dimensional isotropic turbulence (see, for example, Fig. 4.6).[69]

The most glaring problem with the GASP results was that the apparent two-dimensional/three-dimensional transition scale was typically at several hundred kilometers.[70] At such distances, three-dimensional isotropic turbulence would imply that clouds and other structures extended well into outer space! Simply calling the phenomenon "squeezed three-dimensional isotropic turbulence"[71] or "escaped three-dimensional turbulence"[72] explained nothing. In 1999, an update of the GASP experiment—MOZAIC—found essentially the same results, which were strongly interpreted as support for the two-dimensional/three-dimensional theory.[73]

To get to the bottom of this, with the help of Daniel Schertzer and Adrian Tuck,[xx] we reanalyzed[74] the published wind spectra and showed that a key point had been overlooked: The spectra did *not* transition between $k^{-5/3}$ and $k^{-3}$, but rather between $k^{-5/3}$ and $k^{-2.4}$ (see, for example, Figs. 4.5 and 4.7A). In an article published in 2009, titled "Reinterpreting Aircraft Measurements in Anisotropic Scaling Turbulence,"[75] we proposed a simple explanation: The low wavenumber ($k^{-2.4}$) part of the spectrum was *not* simply a poorly discerned $k^{-3}$ signature of isotropic two-dimensional turbulence, but rather the spectrum of the wind in the vertical rather than the horizontal direction! If the turbulence was never isotropic, but rather anisotropic, with different exponents in the horizontal and vertical directions, then a theory was needed to interpret aircraft measurements correctly.[yy] Our new theory easily explained the results by realizing that the aircraft were following gently sloping isobars (rather than isoheights) (Fig. 4.5 and Plate 4.1). The reason was simple: The change in the wind over even quite small vertical displacements could easily exceed its change over large horizontal displacements. On sloping trajectories, at large scales, the vertical statistics could easily become dominant.

But even this didn't satisfy the diehard two-dimensional/three-dimensional theorists—notably, Eric Lindborg. He incited experimentalist colleagues at the National Center for Atmospheric Research, Rod Frehlic and Robert Sharman,[76] to use the big Aircraft Meteorological Data Relay (AMDAR) database to disprove our hypothesis by attempting to demonstrate empirically the statistical equivalence of wind data at constant heights and constant pressures (isoheights vs. isobars). But, the AMDAR technology didn't include global positioning satellite (GPS) altitude determinations, which were needed to distinguish accurately between the two. To prove that our explanation was correct and to close the debate,[77] it was necessary to determine the joint (horizontal–vertical) velocity structure function.[zz] This was finally done with the help of wind data from 14,500 aircraft trajectories that allowed its first direct determination (Fig. 4.8). The aircraft used in our study had Tropical Airborne Meteorological Reporting (TAMDAR) technology, with the necessary metric-scale GPS altitude measurements. The results of this

---

[xx] Tuck was pioneer in aircraft measurements. At the time, he was head of the atmospheric chemistry group at NOAA in Boulder, Colorado. He was responsible for the Antarctic aircraft campaign that, during the late 1980s, established conclusively the existence of the ozone hole and the link to chlorofluorocarbons.

[yy] Until then, turbulence had always been assumed to be isotropic, and isotropy implied there was a unique exponent for each field, so that if aircraft measured scaling spectra, the resulting exponents would be unbiased. A theory of how the aircraft was affected by turbulence and of how the up-and-down aircraft movement might bias the results hadn't seemed necessary.

[zz] This means that we estimated the typical change in the horizontal wind for arbitrary displacements in vertical cross-sections.

massive study[78] showed that the horizontal wind was scaling with an anisotropic in-between "elliptical dimension" of 2.56 ± 0.02, which is close to the theoretical value of 23/9 (=2.555 ... ) discussed in Chapter 3.
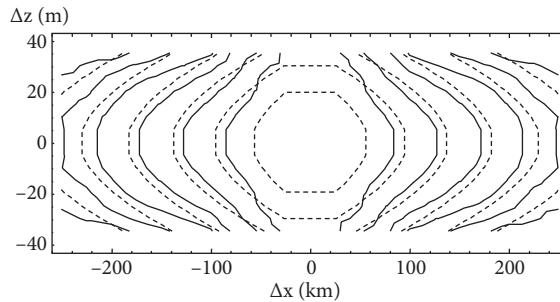


FIGURE 4.8  *This figure shows typical variations in the transverse component of the wind. The solid lines are the measurements and the dashed lines are the theoretical contours for a 23/9 D atmosphere. The contour lines indicate constant values of the typical change of the wind when comparing two points separated by Δx in the horizontal direction and Δz in the vertical direction. Note the very different units (meters and kilometers) in the horizontal and vertical directions; the aspect ratio of the figure is about 1,000:1. If this was plotted using the same units on both axes, the roundish cross-sections marking the spheroscale would be only a few millimeters in size.[79]*

## 4.5  The triumph of scaling stratification

I have summarized a series of attempts to shoehorn the atmosphere into theoretically inspired scalebound frameworks, followed by long struggles against them to exonerate Richardson's wide-range scaling idea. First, there was the mesoscale gap. Had it existed, it would have allowed the small scales to be explained by isotropic three-dimensional turbulence theory and it would have justified Van der Hoven's famously "convenient" division of the atmosphere into large-scale weather and small-scale turbulence. Although the gap had never enjoyed much empirical support, it was marginalized not so much because it contradicted the data, but once again because of convenience—this time of isotropic two-dimensional and three-dimensional theories. Although initial evidence that apparently supported the two-dimensional/three-dimensional theory was eagerly embraced (EOLE, GASP), decades later more careful analyses showed that the data supported scale symmetries rather than direction symmetries: Isotropy had become far too convenient to drop.

Instead of considering the (anisotropic) scaling of the equations themselves, or at least the scaling of model–data hybrids [the reanalyses (Fig. 4.4 lower right)], attention shifted to the numerical models that were called upon for support. The idea was to show that they reproduced "realistic" two-dimensional/three-dimensional transitions (detected in spectra as transitions from $k^{-3}$ to $k^{-5/3}$ behavior).[80] Indeed, the failure of NWPs or GCMs to reproduce such spurious transitions is taken as a sign of model *failure*, rather than of model success![81]

Once again, the trouble is that most numerical weather models are scaling and do *not* display the transition,[82] whereas others may display it but only over very small ranges. A complication is that the majority of models do not display a $k^{-5/3}$ spectrum, but rather a spectrum close to $k^{-2.4}$, which we have seen is, in fact, the large-scale spectrum found on isobars (and equal to the spectrum of the horizontal wind in the vertical direction). The likely explanation is that the models generally use an approximation called the "hydrostatic approximation,"[83] so that the intrinsic horizontal direction in the models is also along isobars, not along isoheights. See, for example, the popular[84] Weather Research and Forecasting (regional) weather model (Fig. 4.9). As for the minority (those numerical models that do seem to show transitions from $k^{-3}$ to $k^{-5/3}$), their spectra are either unrealistic at large scales[85] or, alternatively, in at least one prominent case of a quasi-geostrophic simulation,[86] are spurious.[87]
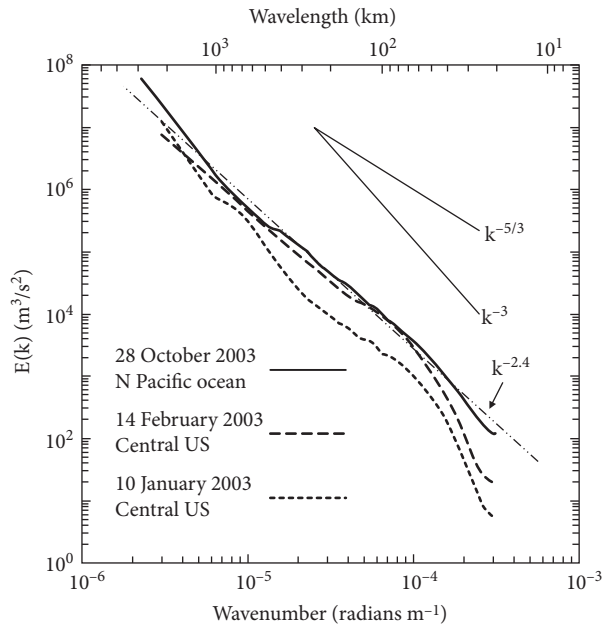


FIGURE 4.9 *This shows sample spectra from Weather Research and Forecasting forecasts of zonal wind averaged over the isobaric surfaces covering, roughly, a range from 3 to 9 km in altitude.[88] Although the authors claimed that this shows a "clear $k^{-3}$ regime" for the solid (oceanic) spectrum, it only spans a range of factor 2 to 3 in scale, and this at relatively unreliable, extremely low wavenumbers. Except for the extremes, the spectra again follow the isobaric predictions $k^{-2.4}$ (thin dashed line) very well over most of the range.[89]*

Sixty years after the mesoscale gap, forty years after the EOLE experiment, and thirty-five years after GASP, Richardson's wide-range scaling has finally been vindicated empirically. The last obstacle preventing closure of the scaling-versus-scalebound debate was a theoretical demonstration that the observed anisotropic

scaling is compatible with the equations. As discussed in Chapter 3, this was indeed done at nearly at the same time,[90] so that the theoretical debate between the two-dimensional/three-dimensional isotropic model and the anisotropic 23/9 D scaling alternative is finally nearing closure.

## 4.6  The atmosphere as a heat engine and the lifetime–size relation

We have examined in considerable detail the history and developments surrounding the issue of wide-range horizontal scaling. Not only is it important in its own right, but also—because of the wind—it implies temporal scaling, at least up to scales of about ten days (the size of the planet divided by the typical large-scale wind speed). Figure 4.4 (lower left) shows that, indeed, up to scales of about 5,000 km and seven days in time, the spatial and temporal spectra are essentially indistinguishable from each other and are also scaling.

To understand this, and in particular to work out the fundamental timescale at which the weather regime breaks down and makes a transition to the lower frequency macroweather regime, we can again go back to the work by Van der Hoven.[91] Aside from the ill-starred spectral gap, his spectrum also showed a more robust feature: a drastic change in statistics at timescales of several days (Fig. 4.2, the "bump" on the left[92]). At first ascribed to "migratory pressure systems," and later termed the "synoptic maximum,"[93] it was eventually theorized to be a result of the unstable nature of atmospheric layers.[94] However, neither its presence in all the atmospheric fields nor its true origin and fundamental implications could be appreciated until the mesoscale gap was eliminated and—with the help of anisotropic, stratified atmospheric scaling—the turbulent laws were extended to planetary scales.

The key feature of anisotropic scaling is that the vertical is controlled by the buoyancy force variance flux (the Bolgiano-Obukhov law) and the horizontal dynamics by the energy flux to smaller scales (in units of watts per kilogram, also known as the "energy rate density"). This is the same dimensional quantity upon which the Kolmogorov law is based, although in the 23/9 D picture, the law only holds in the horizontal, not the vertical. The classic lifetime–size relation is then obtained by using dimensional analysis:

$$(\text{Lifetime}) = (\text{Energy rate density})^{-1/3} (\text{Scale})^{2/3}.$$

This is the space–time relation for structures such as storms; it would be obtained if one followed an individual storm (or other structure, "cell" or "eddy," that could be identified and tracked over time) as it propagated. This is the "Lagrangian" space–time relation,[95] and it explains the slope of the straight line in Figure 2.5. Rather than identifying a structure and following it, it is usually

easier simply to take a series of snapshots—such as those taken by geostationary satellites to produce Figure 4.4 (lower left)—and compare typical sizes in space with typical size in time (durations), which gives the fixed-frame "Eulerian" space–time relation.[96] Using the same infrared imagery in Figure 4.4, one finds that this Eulerian space–time relation is linear.[97] Figure. 4.10 shows the relationship that this spectrum implies: an average wind speed of about 900 km/day ($\approx$11 m/s).
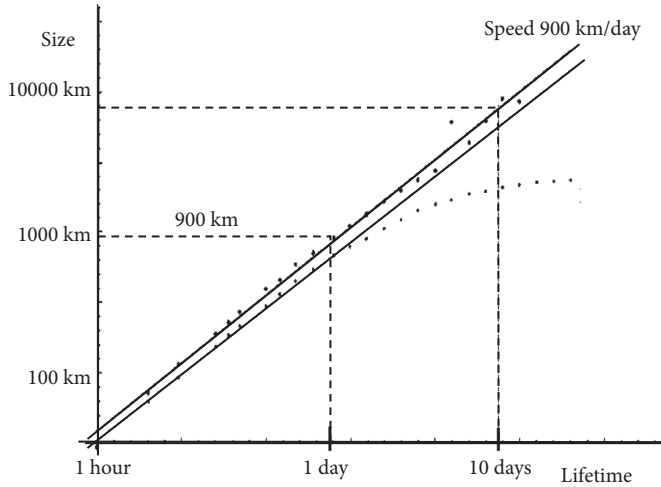


FIGURE 4.10 *The Eulerian (fixed-frame) space–time diagram obtained from the satellite pictures analyzed in Figure 4.4 (lower left). The slopes of the reference lines are nearly the same and correspond to average winds of 900 km/day (i.e., about 10 m/s). The dashed reference lines show the spatial scales corresponding to one day and ten days, respectively.[98]*

If the previous explanation for the space–time relation is correct, then the energy rate density is a fundamental quantity and we should expect that it is linked directly to the main force driving the atmosphere: the sun. Think of the atmosphere as a giant machine—a "heat engine"—for converting solar energy into mechanical energy as wind. Almost all the solar energy is delivered to the atmosphere by solar heating of the surface, and then—thanks to turbulent dynamics (the weather!)—the average power is more or less redistributed uniformly over the atmosphere.[99]

The energy rate density turns out to be easy to estimate. Start with the average solar power delivered to Earth (about 238 W/m$^2$), [100] then divide it by the total atmospheric mass[101] (about 10 t/m$^2$) to yield 0.024 W/kg. This is the average amount of energy per unit time per unit mass that is received from the sun, and almost all of it is at the "short" (visible) wavelengths, and almost all of it is delivered to Earth's surface.[102] From the surface, warm air rises, distributing the energy throughout the troposphere, and north–south winds distribute the energy from the equator to the higher latitudes. Most of this energy ends up being

re-emitted to outer space at "long" (infrared) wavelengths, with only a small fraction doing work (i.e., being transformed into wind)—and this is the part we seek (Box 6.1). This fraction is the thermodynamic efficiency of the atmospheric heat engine.[aaa]

To get an estimate of the efficiency, we can follow Pauluis,[103] who modeled the atmosphere as a heat engine operating between 12°C and 27°C. The theoretical efficiency of a (maximally efficient) Carnot cycle[104] operating between these temperatures is 5%,[105] but Pauluis also considers a "steam cycle," which treats water vapor more realistically and with an efficiency that varies over a of range 1% to 5% (low to high humidity). Using the intermediate value 4% yields an estimate of $0.024 \times 0.04 \approx 0.001$ W/kg. This is essentially the same as the direct estimate of the energy rate density based on global wind data.[106] Although the amount is somewhat variable with latitude and altitude, the global average value is indeed close to 0.001 W/kg.[107]

The thing about the thermodynamic analysis is that it only depends on the succession of thermodynamic states of "packets" of air. It doesn't require assumptions about the detailed dynamics. Because the energy flux density is the basic horizontal scale-invariant quantity, this analysis reveals the physical mechanism behind the thermodynamic cycle: turbulent cascades. If we analyze the data at any scale (above the millimetric dissipation scale), we expect to get—on average—the same value. Alternatively, we have found that by using turbulence theory to estimate the energy flux density from wind gradients, we are able to estimate the atmosphere's thermodynamic efficiency.

## 4.7  The weather–macroweather transition

Now that we have obtained an estimate of 0.001 W/kg, we can use the formula (Lifetime) = (Energy rate density)$^{-1/3}$(Scale)$^{2/3}$ to estimate the lifetime of the largest structures, which is also our estimate of the weather–macroweather transition timescale. Using the largest great circle distance (20,000 km), we obtain a lifetime of approximately five to ten days. Figure 4.11 shows that this simple theory even explains the latitudinal variations: Because the winds (and hence the rate density) are lower near the equator, the transition timescale is a little longer, exactly as the theory predicts.

---

[aaa] For perspective, consider the gasoline-powered internal combustion engine in your car, which is the product of centuries of development and has a thermodynamic efficiency of about 20%, implying that 80% of the fossil fuel energy is wasted. It ends up heating the environment without performing useful work.
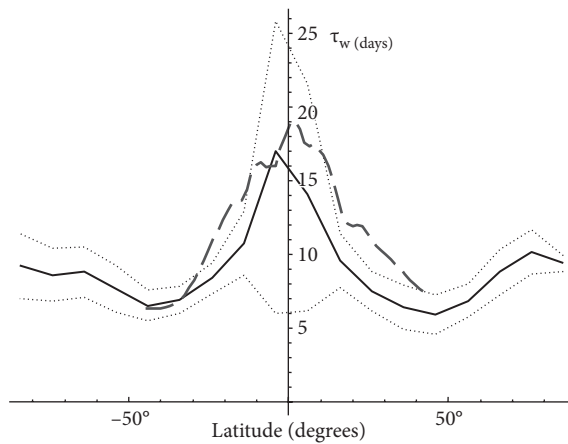
FIGURE 4.11 *The weather–macroweather transition scale ($\tau_w$) estimated directly from break points in the spectra for the temperature as a function of latitude, with the longitudinal variations determining the dashed one-standard deviation limits. The mean over all longitudes is represented by the solid black line; the spread (±34%) is indicated by thin dashed lines.[108] The thick dashed curve is the theoretical $\tau_w$ obtained by estimating the distribution of ε from the European Centre for Medium Range Weather Forecasting reanalyses for the year 2006 (using $\tau_w = \varepsilon^{-1/3} L^{2/3}$, where L is half of Earth's circumference). It agrees very well with the temperature $\tau_w$. $\tau_w$ is particularly high near the equator because the winds tend to be lower, hence a lower ε value. Similarly, $\tau_w$ is particularly low for precipitation because it is usually associated with high turbulence (high ε).[109]*

According to our derivation, structures that are larger and larger live longer and longer, but eventually, because of the finite size of Earth, this must break down. The scale that we just estimated is the breakdown scale—the lifetime of the largest possible structures. At the same time, because structures can only be forecast reliably over their lifetimes, this is also close to the overall deterministic predictability limit.[bbb]

If we go beyond this transition timescale, then we are effectively considering several lifetimes of the largest structures. We should not be surprised to find that the fluctuations over consecutive lifetimes tend to cancel. From the point of view of turbulent laws, the transition from weather to macroweather is a "dimensional transition" because, at longer timescales, the spatial degrees of freedom are essentially "quenched," so that the system's dimension is effectively reduced from (Time + Space) to (Time). Both turbulent cascade models and GCM control runs (i.e., with constant external forcings) reproduce the transition and also produce realistic low-frequency variability.[110] The fact that these weather models reproduce this justifies the term "macroweather" discussed in Chapter 5.

---

[bbb] This conclusion is actually not as obvious as it may seem. One of the consequences of the two-dimensional/three-dimensional model—and one of its attractive features—would have been a potentially much longer predictability limit pertinent to two-dimensional turbulence.

Surprisingly, this energy rate density-based explanation was not discovered[111] until 2010,[112] presumably because no one had dared to suggest that the Kolmogorov law could possibly be relevant at planetary scales! This result is so fundamental that we need to validate it as thoroughly as possible. For example, the ocean is also a turbulent fluid. In many respects, it is similar to the atmosphere, only it is stirred not so much by the sun but by the wind—so that it too should have an "ocean weather"–"ocean macroweather" transition determined by its own typical energy rate density. To test this, we used data from thousands of ocean "drifters" and estimated empirically the near-surface[113] energy rate density. We found that it was about 100,000 times smaller than in the atmosphere: 10 nW/kg. This value yields an ocean weather–ocean macroweather transition time of about one to two years, which is also observed (see Fig. 4.12).[114]
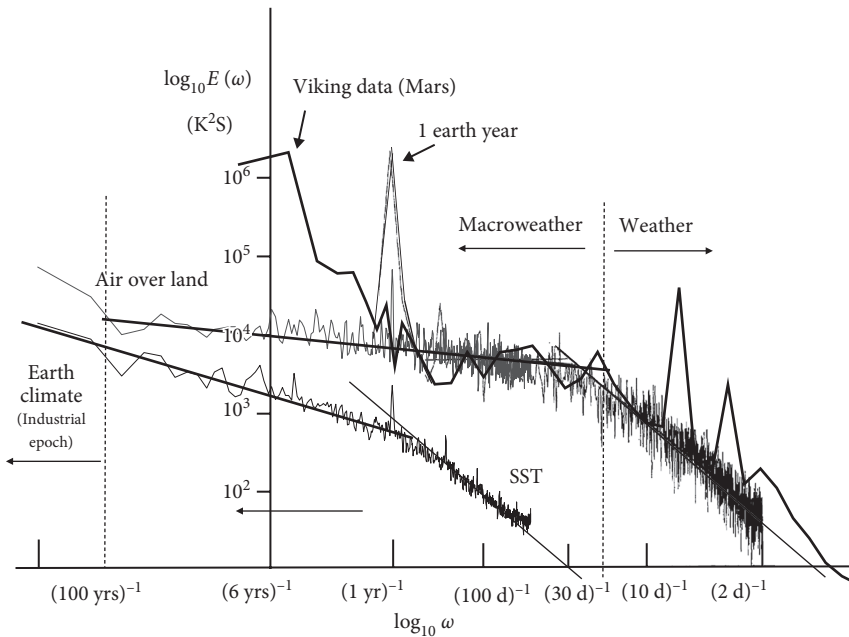


FIGURE 4.12 *The three known weather–macroweather transitions: air over Earth (black and upper left gray lines), the sea surface temperature (SST; ocean) at 5° resolution (lower thin black line), and air over Mars (thick black line). The air-over-Earth curve is from thirty years of daily data from a French station (Macon, black) and from air temperatures for the past one hundred years (5° × 5° resolution, the National Oceanic and Atmospheric Administration's National Climatic Data Center). The spectrum of monthly averaged SSTs is from the same database. The Mars spectra are from Viking lander data. The strong "spikes" at the right are the Martian diurnal cycle and its harmonics.[115]*

We have discussed the fact that these turbulent scaling laws are expected to hold "universally" (i.e., under fairly general conditions in roughly similar atmospheres). But, for the most convincing test of the scaling theory, we need another planet!

## 4.8 The twin planet

"Good afternoon Martians, I bring good tidings from planet Earth: We're twins!"

So began my presentation to a room packed with Mars specialists at a session at the European Geosciences Union in April 2016. Mars may be a sister planet,[ccc] but when it comes to its atmosphere, up until now, scientists had focused on the differences, not the similarities. My introduction caught the audience off guard because I had inverted the usual procedure. I had not gone to Mars to understand the red planet better; rather, I remained on the blue one.

The most important differences between the dynamics of the two atmospheres are the strong control of Martian atmospheric temperature by dust, the larger role of topography, the stronger diurnal and annual cycles, and the larger role of atmospheric tides. But these differences affect mostly the forces driving the system and the nature of the boundaries. If the turbulence approach is correct, at small-enough scales and far enough from the boundaries, then we expect to find the same statistics, the same scaling. The behavior was expected to be independent of the details; it should be "universal."

Using the theory based on the energy rate density, we can easily calculate the Martian weather–Martian macroweather transition timescales. All we need are its distance from the sun (and hence solar insolation per square meter, which is about 55% of ours), the albedo (which is about 53% of ours), and the surface pressure and gravity (and hence the mass per square meter, which is about 1.3% of our value[116]). These values yield an expected 40 mW/kg on Mars, about forty times larger than on Earth. Taking this and the small planetary size into account (about half the diameter of Earth), we predicted a Martian weather–Martian macroweather transition timescale of about 1.8 sols.[ddd]

The problem was then to find data to test the prediction. There had been many Mars landers carrying meteorological instruments, but it turned out that the first and oldest—the two *Viking* landers (1976, 1977)—were still the best for our purpose, with hourly temperature and wind data spanning three Earth years nearly continuously.[118] The spectra obtained from these landers showed that the theory result is, indeed, well respected (Fig. 4.12). Because the spectrum implies that fluctuations longer than 1.8 sols have nearly identical behavior as those on Earth. As proclaimed in the title of our article: "On Mars too expect macroweather."[119]

Our article was published in one of Earth's leading geoscience journals,[120] yet the Martians appeared to be far more interested than the Earthlings. A few days after the article was published, the story was picked up by the amateur *Astronomy Magazine* and was featured prominently on its website.[121]

---

[ccc] Venus is usually called the "sister" planet, and Mars the "red" planet. But now that we're twins . . . .

[ddd] A sol is a Martian day, about 25 Earth hours.

This happened to be at the same time that the first comet lander *Philae* discovered organic molecules on comet 67P/Churyumov–Gerasimenko; the two stories were placed right next to each other on the website. Although Martian macroweather garnered 3,360 likes, poor *Philae* had only 324!

Along with the wind field, the other atmospheric fields were also expected to be universal, and this includes the spatial scaling. It turned out that full reconstructions of the Martian atmosphere existed for nearly three Martian years.[eee] These were calculated using an Earth weather model adapted to Mars, then using an orbiting infrared satellite (*Mars Express*) to estimate the temperature profile. This was enough data to update Martian "reanalyses" constantly, which included the temperature, wind, and pressure. Figures 4.13 and 4.14 show that Earth and Mars do indeed have virtually the same statistics of pressure, wind, and temperature in the east–west and north–south directions. Even the turbulent intermittencies (Fig. 4.14) turned out to be the same.[122] Thanks to Mars, we can be confident about scaling on Earth!
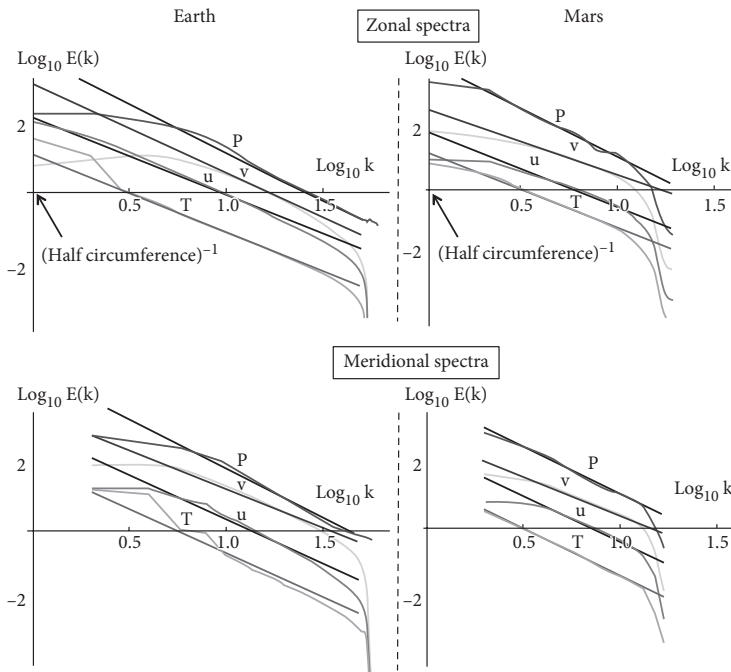


FIGURE 4.13 *A comparison of spectra from terrestrial and Martian reanalyses (left and right columns, respectively) showing the universality of the scaling behavior. The top row shows the zonal (east–west) spectra; the bottom row shows the meridional (north–south) spectra. [The drop-off at high wavenumbers (the extreme right of each curve) is an artifact of the "hyperviscosity" used in the models needed for numerical stability.] P is pressure, v is north–south wind, u is east–west wind, and T is temperature. The value $k = 1$ ($log_{10} k = 0$) corresponds to the half circumference.[123]*

---

[eee] About eight Earth years. After this time, the Mars Express satellite went out of operation. The reanalysis had a resolution of 5° in space and six hours in time.
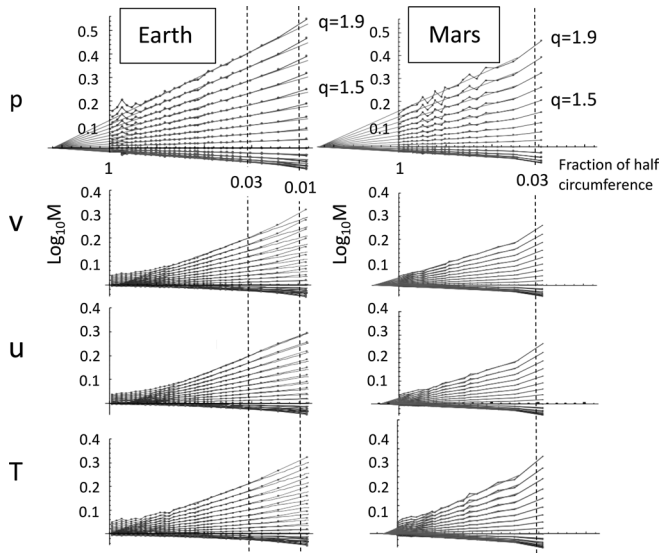
FIGURE 4.14 *A comparison of trace moments (M), terrestrial (left) and Martian (right), for moments q = 0.2, 0.4, 0.6, and 0.8 (downward sloping, at the bottom of each plot); and q = 1.2, 1.4, 1.6, 1.8, and 2 (upward sloping, bottom to top). The q = 1 moment is independent of the averaging scale, it is the flat line along the axis. The horizontal axis (log$_{10}\lambda$) is the horizontal scale expressed as the fraction of the largest available scale [i.e., λ = (Half planet circumference)/(Spatial resolution)]. (λ is the same as the nondimensional wavenumber k used in Figure 4.13.) At a given scale λ, the spread of the lines indicates the variability. By visual inspection, we see that variability increases from the large scales (left) to the small scales (right). The linearity shows that this occurs in a scaling power law manner, as expected for a cascade process. The place where the lines cross is the effective outer scale of the process, the scale from which the variability starts to build up. [In some cases—for example, the pressure (top)—the lines cross at a scale larger than the half circumference. This is because even at this largest scale, there is still variability present resulting from the interactions between the pressure and other fields.] The graphs (top to bottom) are for the surface pressure, north–south (N-S) wind, east–west (E-W) wind, and temperature (the latter three at an altitude of 70% surface pressure). The Martian plots should be compared to the terrestrial ones to the left of the thin black dashed lines on the right of each graph (the points to the right are at scales not represented in the lower resolution Martian reanalysis; the vertical dashed line indicates scales at 4% of the largest distance).*

## Notes

1. Because of the curvature of Earth, beyond about 100 km, ground-based radars start to see over the top of the weather.

2. The ratio is the "Reynolds number."

3. Later, various compressible approximations were made, such as the Boussinesq approximation, but the main theory was based on incompressibility. The atmosphere is

highly compressible and, consequently, is much more dense near the surface than at altitude. However, if the atmosphere were "barotropic"—with constant pressure and constant density surfaces parallel to each other—this wouldn't matter. It is the "baroclinicity" that quantifies the deviations from incompressibility.

4. Taylor, G. I. Statistical theory of turbulence. *Proc. Roy. Soc.* **I–IV** (A151), 421–478 (1935).

5. For example, the idea of "return to isotropy" was interpreted in this way. See Rotta, J. C. Statistische theorie nichtonogenr turbulenz. *Z. Phys.* **129**, 547–572 (1951).

6. I say "on average," because typical experiments would be far from smooth, with energy dissipation occurring very unevenly in both space and in time—even in "bursts." This was the phenomenon of intermittency discussed earlier—the "spottiness" of turbulence—but its full significance was not understood until much later. For a discussion of bursts, see Schertzer, D. & Lovejoy, S. Uncertainty and predictability in geophysics: Chaos and multifractal insights In: *State of the Planet: Frontiers and Challenges in Geophysics* (eds. R. S. J. Sparks & C. J. Hawkesworth), pp. 317–334. (American Geophysical Union, 2004).

7. To obtain such a steady state requires a constant input of energy so that the overall system is very far from thermodynamic equilibrium.

8. Notably in the form of the Karman-Howarth equations (1938).

9. Kolmogorov, A. N. *Grundebegrisse der Wahrscheinlichkeitrechnung.* (Springer 1933). An English translation by N. Morrison appeared under the title *Foundations of the Theory of Probability* (Chelsea) in 1950, with a second edition in 1956.

10. The units of the energy rate density—watts per kilogram—are the same as meters squared per cubic second.

11. Kolmogorov, A. N. Local structure of turbulence in an incompressible liquid for very large Reynolds numbers. *Proc. Acad. Sci. URSS Geochem. Sect.* **30**, 299–303 (1941). [English translation: *Proc. Roy. Soc.* **A434**, 9–17 (1991)].

12. Obukhov, A. M. On the distribution of energy in the spectrum of turbulent flow. *Dokl. Akad. Nauk SSSR* **32**, 22–24 (1941).

13. Onsager, L. The distribution of energy in turbulence. *Phys. Rev.* **68**, 286 (1945). (abstract.).

14. Heisenberg, W. On the theory of statistical and isotropic turbulence. *Proc. Roy. Soc. A* **195**, 402–406 (1948). Veizacker, C. F. Das spektrum der turbulenz bei grossen Reynolds'schen zahlen. *Z. Phys.* **124**, 614 (1948).

15. Corrsin, S. On the spectrum of isotropic temperature fluctuations in an isotropic turbulence. *J. Appl. Phys.* **22**, 469–473 (1951). Obukhov, A. Structure of the temperature field in a turbulent flow. *Izv. Akad. Nauk. SSSR Ser. Geogr. I Geofiz.* **13**, 55–69 (1949). The ~~original and~~ more rigorously justified Corrsin-Obukhov law applied only to "passive scalars" (i.e., to the concentration of dyes or chaff that were pushed around by the flow but that did not themselves affect the flow). However, it is frequently applied to the temperature—as here—even though, in the atmosphere at least, this is not a very good approximation.

16. Bolgiano, R. Turbulent spectra in a stably stratified atmosphere. *J. Geophys. Res.* **64**, 2226 (1959). Obukhov, A. Effect of Archimedean forces on the structure of the temperature field in a turbulent flow. *Dokl. Akad. Nauk SSSR* **125**, 1246 (1959).

17. Batchelor, G. K. *The Theory of Homogeneous Turbulence.* (Cambridge University Press, 1953).

18. Monin, A. S. *Weather Forecasting as a Problem in Physics.* (MIT Press, 1972).

19. Monin, A. S. *Weather Forecasting as a Problem in Physics*. (MIT Press, 1972); quote, p. 5.

20. Monin, A. S. & Yaglom, A. M. *Statistical Fluid Mechanics*. (MIT Press, 1975).

21. Reproduced from Lovejoy, S., Tuck, A. F., Hovde, S. J., & Schertzer, D. Is isotropic turbulence relevant in the atmosphere? *Geophys. Res. Lett.* doi:10.1029/2007GL029359, L14802 (2007).

22. Dewan, E. Saturated-cascade similitude theory of gravity wave spectra. *J. Geophys. Res.* **102**, 29799–29817 (1997).

23. Lovejoy, S., Tuck, A. F., Hovde, S. J., & Schertzer, D. Do stable atmospheric layers exist? *Geophys. Res. Lett.* **35**, L01802, doi:01810.01029/02007GL032122 (2008).

24. We later argued that the results of wave theories could be reinterpreted in terms of strongly nonlinear but scaling dynamics: Lovejoy, S., Schertzer, D., Lilley, M., Strawbridge, K. B., & Radkevitch, A. Scaling turbulent atmospheric stratification, part I: Turbulence and waves. *Q. J. Roy. Meteorol. Soc.* **134**, 277–300 (2008). Pinel, J. & Lovejoy, S. Atmospheric waves as scaling, turbulent phenomena. *Atmos. Chem. Phys.* **14**, 3195–3210 (2014).

25. Panofsky, H. A. & Van der Hoven, I. Spectra and cross-spectra of velocity components in the mesometeorological range. *Q. J. Roy. Meteorol. Soc.* **81**, 603–606 (1955). Van der Hoven, I. Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour. *J. Meteorol.* **14**, 160–164 (1957).

26. Panofsky, H. A. & Van der Hoven, I. Spectra and cross-spectra of velocity components in the mesometeorlogical range. *Q. J. Roy. Meteorol. Soc.* **81**, 603–606 (1955); quote, 606.

27. Panofsky, H. A. & Van der Hoven, I. Spectra and cross-spectra of velocity components in the mesometeorological range. *Q. J. Roy. Meteorol. Soc.* **81**, 603–606 (1955). Van der Hoven, I. Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour. *J. Meteorol.* **14**, 160–164 (1957).

28. This uncommon way of plotting the spectrum was done because the areas under the resulting log ω versus ωE(ω) plot between any two frequencies could be interpreted as the total variance contributed by all the frequencies in the range—an interpretation not possible on usual log-log plots. Adapted from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

29. Robinson, G. D. Some current projects for global meteorological observation and experiment. *Q. J. Roy. Meteorol. Soc.* **93**, 409–418 (1967). Vinnichenko, N. K. The kinetic energy spectrum in the free atmosphere for 1 second to 5 years. *Tellus* **22**, 158 (1969). Goldman, J. L. *The Power Spectrum in the Atmosphere Below Macroscale*. (Institute of Desert Research, University of St. Thomas, 1968).

30. Atkinson, B. W. *Meso-scale Atmospheric Circulations*. (Academic Press, 1981).

31. ~~Fjortoft, R. On the changes in the spectral distribution of kinetic energy in two-dimensional, nondivergent flow. *Tellus* **7**, 168–176 (1953).~~ Kraichnan, R. H. Inertial ranges in two-dimensional turbulence. *Phys. Fluids* **10**, 1417–1423 (1967).

32. This was effectively ~~a derivation~~ of Kraichnan's pure two-dimensional turbulence, starting from a series of nontrivial approximations to the governing equations of the atmosphere.

33. Charney, J. G. Geostrophic turbulence. *J. Atmos. Sci.* **28**, 1087 (1971).

34. Schertzer, D. & Lovejoy, S. The dimension and intermittency of atmospheric dynamics. In: *Turbulent Shear Flow* (eds. L. J. S., Bradbury, F. Durst, B. E. Launder, F. W. Schmidt, & J. H. Whitelaw), pp. 7–33. (Springer-Verlag, 1985).

35. It *also* generally had a $k^{-5/3}$ regime, only this was at very large scales—a fact that was often conveniently forgotten.

36. Morel, P. & Larchevêque, M. Relative dispersion of constant level balloons in the 200 mb general circulation. *J. Atmos. Sci.* **31**, 2189–2196 (1974).

37. More precisely, close to the 200-mb pressure level.

38. The balloons stayed (nearly) on isopycnals (i.e., surfaces of constant density), not on isobars (surfaces of constant pressure)—the key difference being that, although the latter are gradually sloping, the former are highly variable, with large-scale average slopes diminishing at larger and larger scales.

39. Morel, P. & Larchevêque, M. Relative dispersion of constant level balloons in the 200 mb general circulation. *J. Atmos. Sci.* **31**, 2189–2196 (1974).

40. Lacorta, G., Aurell, E., Legras, B., & Vulpiani, A. Evidence for a $k^{-5/3}$ spectrum from the EOLE Lagrangian balloons in the lower stratosphere. *J. Atmos. Sci.* **61**, 2936–2942 (2004).

41. See Appendix 2A in Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades.* (Cambridge University Press, 2013).

42. Bush, V. *Science: The Endless Frontier: A Report to the President on a Program for Postwar Scientific Research.* (National Science Foundation 1960 [1945]); quote, p. 19.

43. Ross, A. *Strange Weather: Culture, Science and Technology in the Age of Limits.* (Verso, 1991).

44. Lederman, L. *Science: The End of the Frontier?* (American Association for the Advancement of Science, 1991).

45. Lederman, L. Science: The end of the frontier? *Science* **21** (Suppl.), 9 (1991); quote, p. 7.

46. Lederman, L. Science: The end of the frontier? *Science* **21** (Suppl.), 9 (1991); quote, p 5.

47. Oreskes, N. & Conway, E. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth from Tobacco Smoke to Global Warming.* (Bloomsbury Press, 2010).

48. Otto, S. L. *The War on Science: Who's Waging It, Why It Matters, What We Can Do About It.* (Milkweed Editions, 2016).

49. The development of revolutionary new data assimilation techniques: (spatial) "3D var" and, later, (space–time) "4D var."

50. Recall that the necessary theory—the 23/9 D model—had been around since the mid 1980s.

51. Quoted from memory from a discussion with Doug Lilly (1929–2008).

52. Lovejoy, S., Agterberg, F., Carsteanu, A., Cheng, Q., Davidsen, J., Gaonac'h, H., Gupta, V., L'Heureux, I., Liu, W., Morris, S. W., Sharma, S., Shcherbakov, R., Tarquis, A., Turcotte, D. & Uritsky, V. Nonlinear geophysics: Why we need it. *EOS* **90**, 456–457 (2009).

53. Lovejoy, S., Agterberg, F. Carsteanu, A. Cheng, Q., Davidsen, J., Gaonac'h, H., Gupta, V., L'Heureux, I., Liu, W., Morris, S. W., Sharma, S., Shcherbakov, R., Tarquis, A., Turcotte, D., & Uritsky, V. Nonlinear geophysics: Why we need it. *EOS* **90**, 456–457 (2009); quote, p. 456.

54. Schertzer, D. & Lovejoy, S. Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes. *J. Geophys. Res.* **92**, 9693–9714 (1987).

55. Gabriel, P., Lovejoy, S., Schertzer, D., & Austin, G. L. Multifractal analysis of resolution dependence in satellite imagery. *Geophys. Res. Lett.* **15**, 1373–1376 (1988). Lovejoy, S., Schertzer, D., & Tsonis, A. A. Functional box-counting and multiple elliptical dimensions in rain. *Science* **235**, 1036–1038 (1987). Lovejoy, S., Schertzer, D., Silas, P., Tessier, Y., & Lavallée, D. The unified scaling model of atmospheric dynamics and systematic analysis in cloud

radiances. *Annal. Geophys.* **11**, 119–127 (1992). Lovejoy, S. & Schertzer, D. Multifractals, universality classes and satellite and radar measurements of cloud and rain fields. *J. Geophys. Res.* **95**, 2021 (1990). Lovejoy, S., Schertzer, D., & Stanway, J. D. Direct evidence of planetary scale atmospheric cascade dynamics. *Phys. Rev. Lett.* **86**, 5200–5203 (2001).

56. A partial exception was seen in an article by Strauss and Ditlevsen that considered reanalyses. A reanalysis is a complex data–model hybrid that, effectively, fills in holes in the data (and partially corrects for errors) by constraining the system using the equations of the atmosphere as embodied in a numerical model. Although these authors strongly criticized the reigning two-dimensional picture, rather than analyzing the scaling of model outputs directly, they instead analyzed more complex, theoretically inspired constructs so that their conclusions were less clear. Strauss, D. M. & Ditlevsen, P. Two-dimensional turbulence properties of the ECMWF reanalyses. *Tellus* **51A**, 749–772 (1999).

57. Adapted from Tessier, Y., Lovejoy, S., & Schertzer, D. Universal multifractals: Theory and observations for rain and clouds. *J. Appl. Meteorol.* **32**, 223–250 (1993).

58. Stolle, J., Lovejoy, S., & Schertzer, D. The stochastic cascade structure of deterministic numerical models of the atmosphere. *Nonlin. Proc. Geophys.* **16**, 1–15 (2009).

59. Normally, a straight line on a log-log plot shows scaling, but in this case there are spurious edge effects related to the geometry of the images with respect to the dominant directions of the structures.

60. These views are adapted from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

61. This is the same experiment that included the data used in Figure 1.6. The fractality of aircraft trajectories (fractal dimension = 1.56) was first discovered in stratospheric aircraft trajectories in Lovejoy, S., Schertzer, D., & Tuck, A. F. Fractal aircraft trajectories and nonclassical turbulent exponents. *Phys. Rev. E* **70**, 036301–036305 (2004).

62. Adapted from Lovejoy, S., Schertzer, D., & Tuck, A. F. Fractal aircraft trajectories and nonclassical turbulent exponents. *Phys. Rev. E* **70**, 036306-036301-036305 (2004).

63. In turbulence jargon, it leads to "intermittency corrections" to the exponents so that, for example, the 5/3 in the spectrum $k^{-5/3}$ would be about 0.03 too big. It would be closer to $k^{-1.63}$.

64. Lovejoy, S., Tuck, A. F., Hovde, S. J., & Schertzer, D. The vertical cascade structure of the atmosphere and multifractal drop sonde outages. *J. Geophys. Res.* **114**, D07111, doi:07110.01029/02008JD010651 (2009).

65. Adapted from Gage, K. S. & Nastrom, G. D. Theoretical interpretation of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft during GASP. *J. Atmos. Sci.* **43**, 729–740 (1986).

66. Both views are reproduced from Lovejoy, S., Tuck, A. F., Schertzer, D., & Hovde, S. J. Reinterpreting aircraft measurements in anisotropic scaling turbulence. *Atmos. Chem. Phys.* **9**, 1–19 (2009).

67. By making several additional assumptions, winds were deduced indirectly, but these assumptions later turned out to be questionable.

68. Nastrom, G. D. & Gage, K. S. A first look at wave number spectra from GASP data. *Tellus* **35**, 383 (1983). Nastrom, G. D., Gage, K. S., & Jasperson, W. H. Kinetic energy spectrum of large and meso-scale atmospheric processes. *Nature* **310**, 36–38 (1984). Nastrom, G. D. & Gage, K. S. A climatology of atmospheric wavenumber spectra of wind and temperature by commercial aircraft. *J. Atmos. Sci.* **42**, 950–960 (1985). Gage, K. S. & Nastrom, G. D.

Theoretical interpretation of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft during GASP. *J. Atmos. Sci.* **43**, 729–740 (1986).

69. Interestingly, Nastrom and Gage themselves interpreted their results as support for yet another theory based on gravity waves.

70. Lilly, D. K. Two-dimensional turbulence generated by energy sources at two scales. *J. Atmos. Sci.* **46**, 2026–2030 (1989). Bacmeister, J. T., Eckermann, S. D., Newman, P. A., Lait, L. Chan, K. R., Loewenstein, M., Proffitt, M. H., & Gary, B. L. Stratospheric horizontal wavenumber spectra of winds, potential temperature, and atmospheric tracers observed by high-altitude aircraft. *J. Geophys. Res.* **101**, 9441–9470 (1996). Gao, X. & Meriwether, J. W. Mesoscale spectral analysis of in situ horizontal and vertical wind measurements at 6 km. *J. Geophys. Res.* **103**, 6397–6404 (1998). Lindborg, E. Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? *J. Fluid Mech.* **388**, 259–288 (1999).

71. Högström, U., Smedman, A. N., & Bergström, H. A case study of two-dimensional stratified turbulence. *J. Atmos. Sci.* **56**, 959–976 (1999).

72. Lilly, D. K. Two-dimensional turbulence generated by energy sources at two scales. *J. Atmos. Sci.* **46**, 2026–2030 (1989).

73. Lindborg, E. & Cho, J. Horizontal velocity structure functions in the upper troposphere and lower stratosphere II: Theoretical considerations. *J. Geophys. Res.* **106**, 10233–10241 (2001). Cho, J. & Lindborg, E. Horizontal velocity structure functions in the upper troposphere and lower stratosphere I: Observations. *J. Geophys. Res.* **106**, 10223–10232 (2001).

74. Lovejoy, S., Tuck, A. F., Schertzer, D., & Hovde, S. J. Reinterpreting aircraft measurements in anisotropic scaling turbulence. *Atmos. Chem. Phys.* **9**, 1–19 (2009).

75. Lovejoy, Tuck, Schertzer, & Hovde. Reinterpreting aircraft measurements.

76. Frehlich, R. G. & Sharman, R. D. Equivalence of velocity statistics at constant pressure or constant altitude. *Geophys. Res. Lett.* **37**, L08801, doi:08810.01029/02010GL042912, 042010 (2010).

77. Frehlich & Sharman, Equivalence of velocity statistics. Lindborg, E., Tung, K. K., Nastrom, G. D., Cho, J. Y. N., & Gage, K. S. Comment on "Reinterpreting aircraft measurements in anisotropic scaling turbulence" by Lovejoy et al. 2009. *Atmos. Chem. Phys. Discuss.* **9**, 22331–22336 (2009).

78. Pinel, J., Lovejoy, S., Schertzer, D., & Tuck, A. F. Joint horizontal–vertical anisotropic scaling, isobaric and isoheight wind statistics from aircraft data. *Geophys. Res. Lett.* **39**, L11803, doi:10.1029/2012GL051698 (2012).

79. To verify this, you can check that the roundish shape in the figure shows that a vertical displacement of $\Delta z \approx 20$ m corresponds to a horizontal displacement of $\Delta x \approx 50$ km. It can be checked that these values verify the relation $(\Delta x/l_s)^{1/3} \approx (\Delta z/l_s)^{3/5}$, with sphero-scale $l_s \approx 1$ mm. Adapted from Pinel, J., Lovejoy, S., Schertzer, D., & Tuck, A. F. Joint horizontal–vertical anisotropic scaling, isobaric and isoheight wind statistics from aircraft data. *Geophys. Res. Lett.* **39**, L11803, doi:10.1029/2012GL051698 (2012).

80. Even from the purely theoretical perspective of isotropic turbulence theory, making a model that stably allows the coexistence of both three-dimensional and two-dimensional regimes is not trivial. This is because of the possibility of "three-dimensionalization" of two-dimensional flows (i.e., the likelihood that, if it existed, isotropic three-dimensional turbulence would destabilize an otherwise two-dimensional flow). Ngan, K., Straub, D. N., & Bartello, P. Three-dimensionalisation of freely-decaying two-dimensional flows. *Phys. Fluids* **16**, 2918–2932 (2004).

81. Berner, J., Shutts, G. J., Leutbecher, M., & Palmer, T. N. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.* **66**, 603–626 (2009).

82. Palmer, T. N. A nonlinear dynamical perspective on model error: A proposal for non-local stochastic–dynamic parameterisation in weather and climate prediction models. *Q. J. Roy. Meteorol. Soc.* **127**, 279–304 (2001).

83. In this approximation, the pressure at any point is exactly equal to the weight per area of the atmosphere above the point. It ignores any vertical accelerations.

84. Skamarock, W. C. Evaluating mesoscale NWP models using kinetic energy spectra. *Monthly Weather Rev.* **132**, 3020 (2004).

85. Takayashi, Y. O., Hamilton, K., & Ohfuchi, W. Explicit global simulation of the meso-scale spectrum of atmospheric motions. *Geophys. Res. Lett.* **33**, doi:10.1029/2006GL026429 (2006). Hamilton, K., Takahashi, Y. O., & Ohfuchi, W. Mesoscale spectrum of atmospheric motions investigated in a very fine resolution global general circulation model. *J. Geophys. Res.* **113**, D18110, doi:18110.11029/12008JD009785 (2008).

86. Tung, K. K. & Orlando, W. W. The $k^{-3}$ and $k^{-5/3}$ energy spectrum of atmospheric turbulence: Quasigeostrophic two-level model simulation. *J. Atmos. Sci.* **60**, 824–835 (2003).

87. In this reference, the carefully crafted high-wavenumber $k^{-5/3}$ regime is a spurious consequence of an energy buildup resulting from unresolved high wavenumbers with respect to an incorrectly "tuned" hyperviscosity.

88. Adapted from Skamarock, W. C. Evaluating Mesoscale NWP models using kinetic energy spectra. *Monthly Weather Rev.* **132**, 3020 (2004).

89. Reproduced from Lovejoy, S., Tuck, A. F., & Schertzer, D. The horizontal cascade structure of atmospheric fields determined from aircraft data. *J. Geophys. Res.* **115**, D13105, doi:doi:10.1029/2009JD013353 (2010).

90. The article cited here actually showed the anisotropic scaling of the slightly simpler but still fundamental vorticity equation. Schertzer, D., Tchiguirinskaia, I., Lovejoy, S., & Tuck, A. F. Quasi-geostrophic turbulence and generalized scale invariance: A theoretical reply. *Atmos. Chem. Phys.* **12**, 327–336 (2012).

91. Van der Hoven, I. Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour. *J. Meteorol.* **14**, 160–164 (1957).

92. The fact that it was displayed as a "bump" rather than simply as a transition between two low- and high-frequency regimes was an artifact in the way that the spectrum was plotted. As usual, the logarithm of the frequency was used to display a wide range of timescales; however, the vertical axis used the logarithm of the product of the frequency with the spectrum rather than just the spectrum itself. This was done so that, in the log-log plot, the total area under a bump would be preserved. However, it had the perverse effect of transforming the smooth transition from one scaling regime to another into a bump that lent itself easily to scalebound interpretations as a simply localized quasi-periodic phenomenon.

93. Kolesnikov, V. N. & Monin, A. S. Spectra of meteorological field fluctuations. *Izvestiya Atmos. Oceanic Phys.* **1**, 653–669 (1965).

94. Specifically, a result of baroclinic instabilities. These occur when the pressure and density gradients in a stratified fluid are strongly misaligned. One of the many problems with this explanation is that it only works in mid latitudes whereas the transition time-scale is not much different in equatorial regions (see Fig. 4.11). Vallis, G. Mechanisms of

climate variability from years to decades. In: *Stochastic Physics and Climate Modelling* (eds. P. Williams & T. Palmer), pp. 1–34. (Cambridge University Press, 2010).

95. Pinel, J., Lovejoy, S., & Schertzer, D. The horizontal space–time scaling and cascade structure of the atmosphere and satellite radiances. *Atmos. Res.* **140–141**, 95–114, (2014). Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

96. There are many ways to determine this. The figure was constructed by considering how the average amplitudes of fluctuations in infrared radiances change as functions of spatial scale and then as functions of temporal scales. In this way, each timescale can be assigned to a spatial scale. For example, at a fixed location, the average fluctuation from one day to the next is the same as the average fluctuation at a fixed instant (image) between two points separated by 900 km. On average, moving 900 km at a single instant changes the radiances just as much as staying in a fixed location and waiting for a day to pass.

97. For a theoretical explanation, see Radkevitch, A., Lovejoy, S., Strawbridge, K. B., Schertzer, D., & Lilley, M. Scaling turbulent atmospheric stratification, part III: Empirical study of space–time stratification of passive scalars using lidar data. *Q. J. Roy. Meteorol. Soc.* doi:10.1002/qj.1203 (2008).

98. Reproduced with permission from Pinel, J. *The Space–Time Cascade Structure of the Atmosphere*. PhD thesis, McGill University (2012).

99. Due to intermittency ε varies hugely from day to day, place to place but when averaged over long periods, stable estimates can be obtained. The systematic variations are with altitude and with location, especially latitude, overall these variations are by the relatively small factors of 2 and 6, respectively.

100. This average takes into account the day-to-night and latitudinal variations, as well as the average reflection coefficient (albedo) of the earth (about 30%).

101. It might be a little more accurate to take only the lower atmosphere, the "troposphere," but this has about 80% of the overall mass, so this distinction isn't very important.

102. Very little is absorbed directly in the atmosphere.

103. Pauluis, O. Water vapor and mechanical work: A comparison of Carnot and steam cycles. *J. Atmos. Sci.* **68**, 91–102 (2011).

104. The (simplest) cycle proposed by Pauluis is an isothermal expansion at 300K (the surface) followed by an adiabatic expansion (resulting from rising warm air), with the temperature dropping from 300K to 285K. This is followed by an isothermal compression at 285K and then an adiabatic compression that warms the air back to 300K again at the surface. See Laliberté, F., Zika, J., Mudryk, L., Kushner, P. J., Kjellsson, J., & Döös, K. Constrained work output of the moist atmospheric heat engine in a warming climate. *Science* **347** (6221), 540–543 (2015).

105. This is determined by the theoretical Carnot efficiency: the temperature difference (27 – 12 = 15°C) divided by the absolute temperature (27 + 273 = 300 K): 15/300 = 0.05.

106. The energy rate density was be estimated from gridded data by the turbulent formula (Energy rate density) = (Change in the horizontal wind across a grid box)$^3$/(Grid size). Lovejoy, S. & Schertzer, D. Towards a new synthesis for atmospheric dynamics: Space–time cascades. *Atmos. Res.* **96**, 1–52 (2010).

The value 1 mW/m$^2$ could be regarded as a modern estimate of the energy rate density ε. The calculation here is quite similar to that of Palmen (1959) who found a value of $\varepsilon = 4$ mW/kg with an efficiency of 2%. Monin (1972) traces the earliest estimate of ε (5 mW/m$^2$) back

to 1926, cited in: Brunt, D. *Physical and Dynamical Meteorology*, 2nd edition. (Cambridge University Press, 1939). Palmen, E. In: *The Atmosphere and the Sea in Motion* (ed. B. Bolin), pp. 212–224. (Rockefeller Institute Press in association with Oxford University Press, 1959). Monin, A. S. *Weather Forecasting as a Problem in Physics*. (MIT Press, 1972).

*Author: please provide the chapter title for "The Atmosphere and the Sea in Motion".*

107. What about the other thermodynamic variable? The entropy? It turns out that the buoyancy force variance that is conserved in the vertical—the quantity analogous to the velocity variance flux (which is the same as the energy rate density)—is proportional to the entropy squared per unit mass per time (i.e., it is not the entropy directly, but rather its average squared that is constrained).

108. The data are from the 20CR [Compo, G. P., et al. The Twentieth Century Reanalysis project. *Q. J. Roy. Meteorol. Soc.* **137**, 1–28 (2011).]. The $\tau_w$ estimates were made by performing bilinear log-log regressions on spectra from 180-day-long segments averaged over 280 segments per grid point.

109. Adapted from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

110. Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

111. Monin came remarkably close. On page 5 of his influential *Weather Forecasting as a Problem in Physics,* he invoked cascade processes proceeding at a constant rate $\varepsilon$ from large to small scales and then used Richardson's 4/3 law to obtain the same lifetime–size relation that we deduced directly from the energy rate density: $\tau = \varepsilon^{-1/3}L^{2/3}$. Then—without further justification—he invoked Obukhov's estimate of the length scale of synoptic scale processes at 3,000 km, obtained as the ratio of the *speed of sound* to the Coriolis parameter (!). He then plugged this value into the lifetime–size relation (using the value $\varepsilon = 5$ mW/kg) to yield an estimate of about three days for the lifetime of large "synoptic-scale" processes. Monin, A. S. *Weather Forecasting as a Problem in Physics*. (MIT Press, 1972).

112. Lovejoy, S. & Schertzer, D. Towards a new synthesis for atmospheric dynamics: Space–time cascades. *Atmos. Res.* **96**, pp. 1–52 (2010).

113. In the atmosphere, the average energy rate density is not very different throughout, so that a single value is already a good approximation. ~~However, in the ocean, it decreases very rapidly, with depth implying—as is well known—that deep currents may live for thousands of years.~~ Although this estimate of the ocean surface energy rate density gives a good estimate of the observed ocean weather–ocean macroweather transition of the ocean surface temperature, the ocean is fundamentally different from the atmosphere inasmuch as the energy rate density is very far from being uniform. As we move from the surface to lower depths, it decreases very rapidly: by factors of thousands or more in the first 100 m, and perhaps by as much as a billion over the first kilometer or so. Because the transition timescale varies as the inverse cubed root, this corresponds to deep currents with millennial-scale lifetimes.

114. Lovejoy, S. & Schertzer, D. Towards a new synthesis for atmospheric dynamics: Space–time cascades. *Atmos. Res.* **96**, pp. 1–52 (2010).

115. Adapted from Lovejoy, S., Muller, J. P., & Boisvert, J. P. On Mars too, expect macroweather. *Geophys. Res. Lett.* **41**, 7694–7700 (2014).

116. The surface atmospheric pressure is about two hundred times less than on Earth and, as a result of seasonal condensation of $CO_2$ from the atmosphere, it is actually somewhat variable. The mass/area estimate takes into account the reduced surface gravity.

117. ~~The value is forty times greater than the terrestrial, and it~~ is primarily the result of the much smaller atmospheric mass.

118. This may be surprising. The reason is that more recent landers such as *Pathfinder* had a much greater frequency response (seconds), but the records were much shorter. The *Viking* data turned out to be optimal.

119. Lovejoy, S., Muller, J. P., & Boisvert, J. P. On Mars too, expect macroweather. *Geophys. Res. Lett.* **41**, 7694–7700 (2014).

120. Lovejoy, Muller, & Boisvert. On Mars too.

121. The basic article can still be found at http://www.astronomy.com/news/2014/11/mars-has-macroweather-too.

122. The hierarchy of multifractal exponents quantifying the intermittency were nearly identical.

123. Adapted from Chen, W., Lovejoy, S., & Muller, J. P. Mars' atmosphere: The sister planet, our statistical twin. *J. Geophys. Res. Atmos.* **121**, doi:10.1002/2016JD025211 (2016).

{ 5 }

# Macroweather, the climate, and beyond

## 5.1 Macroweather planet

"Expect the cold weather to continue for the next ten days, followed by a warm spell."

This might have been the fourteen-day weather forecast for Montreal on December, 31, 2006 (Fig. 5.1, top). But imagine what it might have been if Earth rotated about its axis ten times more slowly, so that the length of the day coincided with the ten-day weather–macroweather transition scale—an alignment of scales almost achieved on Mars.[a] In that case (Fig. 5.1, bottom) we would have heard, "Expect mild weather on Monday, followed by freezing temperatures, until a warm spell on Thursday, followed by a brisk Friday and Saturday, a warming on Sunday and Monday, followed by freezing on Tuesday, then a four-day warm period followed by freezing and then warming." Although long-term trends in weather can persist for up to ten days or so, in macroweather, the upshifts tend to be followed immediately by downshifts (and vice versa) and, although longer term trends exist, they are much more subtle, resulting from imperfect cancelations of successive fluctuations.

---

[a] In Chapter 4, we saw that Mars was nearly such a macroweather world, with the transition at 1.8 sols. If the terrestrial transition time was one day, then within each day, the temperature variations would be like those at the top of Figure 5.1, except they would be superposed on a roughly sinusoidally varying diurnal cycle.
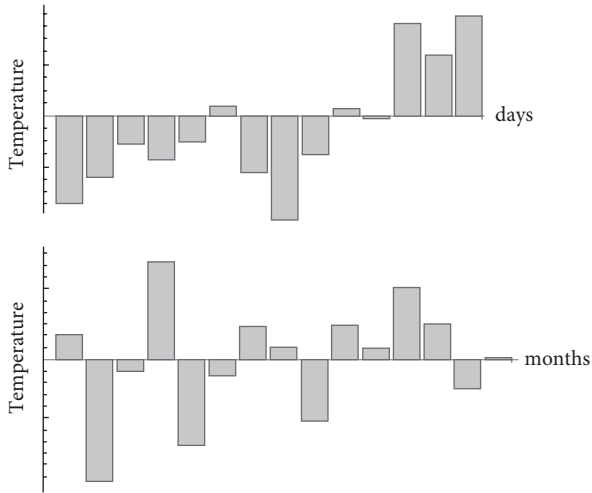
FIGURE 5.1 *(Top) Weather scales: the mean daily temperatures in Montreal, Canada, for January 1 through 14, 2006. (Bottom) Macroweather temperatures for Montreal obtained from fourteen consecutive monthly anomalies from January 2000 through February 2001. To bring out the difference in appearance, the mean (−1°C) was adjusted to be the same as in the top graph and it was scaled so that the spread about the mean (the standard deviation, 4.9°C) was also the same.*

The tendency of macroweather fluctuations to cancel rather than to accumulate is its defining feature, and cancelation is synonymous with stability.[1] Quantitatively, it implies that the temporal fluctuation exponent *H* is negative. In the weather regime with positive *H*, the temperature, wind, and other variables wander up and down with prolonged swings. The weather is a metaphor for instability. If we average macroweather over longer and longer times, its variability is reduced systematically so that it appears to converge to a well-defined value. In that sense, macroweather is what you expect, the weather is what you get.

But what about macroweather's spatial properties? As usual, forecasts can be explained with recourse to maps. For example, Plate 5.1[b] (left) shows the day-to-day evolution of the daily temperatures corresponding to the forecast in Figure 5.1 (top). Focusing on Canada and the United States (within the green ellipses), we would have been told, "A mass of warm air will be displaced gradually by colder Arctic air descending from the northwest and covering the continent

---

[b] In both weather and macroweather, to bring out the temperature changes that are relatively small with respect to the absolute temperatures, I use anomalies—in the first case, with respect to the average for the whole month; in the second, for the previous thirty Januaries (as discussed later).

by Thursday." In the macroweather planet forecast, we might hear (Plate 5.1, right): "The mass of unusually cold air currently over the continent will shrink on Tuesday, spread to the northeast on Wednesday, and, by Thursday, will expand to cover most of North America." Although the sequence of weather maps displays the typical midlatitude west-to-east movement of the weather, the macroweather maps have no such constraint; macroweather "disturbances" can move easily from east to west.[c]

Although the appearance of individual temperature maps for weather and macroweather are not so different, they are both fairly smooth in space, being mainly distinguished by the way they evolve in time: the signs of their temporal $H$ values. But $H$ only characterizes typical, average fluctuations. In Figure 1.6 and Box 4.1, we saw how a fairly innocent-looking aircraft transect hid very strong variability, "spikiness," "intermittency." To bring this out, consider Figure. 5.2, which compares the spikiness of weather, macroweather, and climate both in space (bottom row) and in time (top row). To make the comparisons as fair as possible, I present 360 points for each (corresponding to a spatial resolution of 1° of longitude and, respectively, one hour, four months, and 142 years in time; see Fig. 1.5C). I took the absolute differences (so that the minimum is zero), divided them by their means (so that each series fluctuates around the average value of 1), and used a common vertical scale.[2] By inspection, we can see that the macroweather time series is exceptional, displaying only small, nonintermittent fluctuations. The maximum is quite close to what would be expected if the process were Gaussian. On the contrary, in space (left column), macroweather is highly spiky, as are the weather and climate in both time and in space. Indeed, if any of the latter were produced by a Gaussian process, their maxima would correspond to probabilities of less than one in a million (the top dashed horizontal lines).[d]

---

[c] This may appear to contradict the fact that midlatitude winds are predominantly from west to east. However, over macroweather timescales, weather disturbances can travel around the planet so that macroweather disturbances can move either east or west.

[d] More spike plots can be found in Lovejoy, S. The spectra, intermittency and extremes of weather, macroweather and climate. *Nat. Sci. Rep.* **8,** 1–13 (2018).
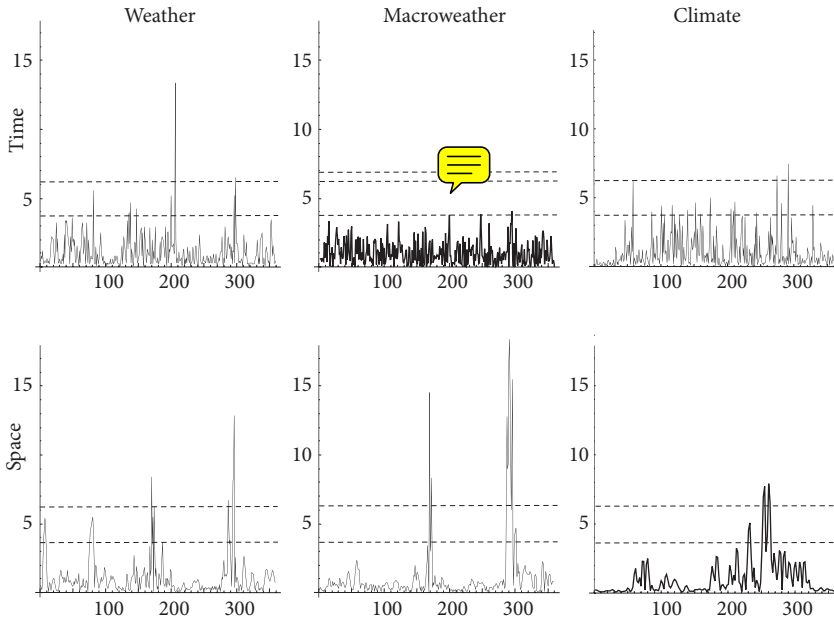
FIGURE 5.2  *A comparison of the "spikiness" (intermittency) in time and space of weather and macroweather series. The graphs show absolute differences of east–west spatial transects [bottom, all along the 45th parallel (north) and time series (top) at weather scales (left), macroweather scales (middle), and climate scales (right)]. The graphs are each from 360 points and show the absolute differences between consecutive values. All series were normalized by their means. Each graph has two parallel dashed lines; the lower one corresponds to a bell curve (Gaussian) probability of 1/360, which is the level of the expected maximum value for a series with 360 points. The upper dashed line corresponds to a (Gaussian) probability of one in a million. Although the spatial intermittencies (bottom) are not too different, the temporal intermittencies are nearly absent from the macroweather series (upper middle). This spikiness is a typical consequence of multifractality (Box 2.2) and it implies that the spectrum will itself display spikes—in this case, misinterpreted easily as spurious periodicities (see Fig. 2.2B). (Upper left) Hourly temperature data from January 1 through 15, 2006, from a station in Lander, Wyoming. (Upper middle): Twentieth-Century Reanalysis from 1891 to 2011. Each point is a four-month average. The data are from a single 2° × 2° grid point over Montreal, Canada (45°N). (Upper right) Greenland Ice Core Project (from Summit, Greenland, at about 75°N) paleotemperature absolute differences degraded to 240-year resolution (present to 86,400 years before present, left to right, not right to left!). [The present (Holocene; discussed in Section 5.6) corresponds to the recent period (up to about the fiftieth time interval). It is not very spiky.] (Lower left) European Centre for Medium Range Weather Forecasting reanalysis for the average temperature of January, 21, 2000, along the 45°N parallel at a resolution of 1° longitude. (Lower middle) The same as at left, but for the temperature averaged over the month of January 2000. (Lower right) The 45°N spatial gradient (the resolution is actually 2° longitude) for the Twentieth-Century Reanalysis, averaged from 1871 to 2012.*

The various definitions and illustrations relevant for this chapter are presented in Table 5.1.

TABLE 5.1 A comparison of different notions used in discussing the atmosphere, weather, macroweather, and the climate.

| Weather | Macroweather | Climate |
|---|---|---|
| **DEFINITIONS AND TIMESCALES** | | |
| The high-frequency $H > 0$ regime. The upper limit is the lifetime of planetary structures, less than five to ten days. | The $H < 0$ regime between the weather and the climate. Five to 10 days to ten to thirty years (Anthropocene), 100 to 3,000 years (preindustrial). | The low-frequency $H > 0$ regime: The lower limit is the scale at which low-frequency forcings increasing with scale begin to dominate[a] the macroweather forcing that decreases with scale: ten to thirty years (Anthropocene), 100 to 3,000 years (preindustrial). |
| **SOME CHARACTERISTICS** | | |
| Fluctuations increasing with scale. High temporal and spatial intermittency (turbulence). | Fluctuations decreasing with scale. Low temporal and high spatial intermittency. | Fluctuations increasing with scale. High temporal[b] and spatial intermittency (climate zones). |
| **STATES** | | |
| Weather states depend on the spatial resolution. Temporal averages should be taken over the typical lifetimes of structures at the scales corresponding to the highest map resolution. | Macroweather states are averages over weather scales and then anomalies are taken with respect to the (seasonally) determined climate states (conventionally, 30 years). | Climate states are averages over weather and macroweather scales, conventionally up to the climate-normal period of thirty years (adjusted seasonally). |
| **GLOBAL CLIMATE MODEL NOTIONS** | | |
| The weather state is the collection of atmospheric fields at a single time step (e.g., ten minutes) at the model spatial resolution (e.g., 10–100 km). The time step must be adjusted to correspond to the lifetime of structures at the highest resolution. | For long- or extended-range weather forecasts (conventionally thirty days or longer), the anomalies are defined as the difference between the monthly average and the most recent (now decadal) thirty-year "climate normal" (adjusted seasonally). | The climate is the average obtained over very long control runs (i.e., runs with fixed external parameters). The convergence is ultraslow (see Fig. 5.1). Because in control runs macroweather continues to infinitely long timescales, the control run climate is a climate state. |
| **EXAMPLES** | | |
| Series: Figures 1.5~~D~~, 2.~~18B~~, 5.1, and 5.2 Transects: Figures 1.6, 4.5, and 5.2; and Plates 4.1 and 4.2 ~~Fig. 4.6D~~ Maps: Plate 5.1 | Series: Figures 1.5C, 2.18B, 5.1B, 5.2, and 5.7A; and Plate 5.3 Transects: Figure 5.2 Maps: Plate 5.1 | Series: Figures 1.5B and 2.~~18B~~ ~~(paleo)~~ Transects: Figure 5.4 Maps: Figure 5.4 and Plate 5.2 |

AQ: Please clarify the reference to figure 5.1B here. The figure does not have A and B in it.

AQ: Please clarify the reference figure 4.6 here.

(*continued*)

TABLE 5.1 Continued

| Weather | Macroweather | Climate |
| --- | --- | --- |
| **ANALYSES** | | |
| Series: Figures 2.3, 2.5, 2.18A, 3.11, 4.11, and 4.12 Maps: Figures 2.6, 4.4, and 4.6 Transects: Figures 3.2; 3.7; 3.11; 4.1; 4.5, 4.7A, B; 4.8; 4.9; 4.13; and 4.14 | Series: Figures 2.3, 2.17, 2.18A, 4.2, 4.11, 4.12, 5.2 and 5.8 Transects: Figure 5.2 | Series: Figures 2.3, 2.17, 2.18B, 9, and 5.2 Maps: Plate 5.2 Transects: Figures 5.2 and 5.4 |

[a]There may also be new, slow internal sources of variability that have yet to be discovered.

[b]Inferred from analyses of paleodata.

## 5.2 Macroweather, climates, and climate states

The strong intermittency of the weather regime is not surprising and is a result of its multifractal turbulent nature (discussed in Chapter 4). However, the averaging over weather scales used to obtain the macroweather series greatly reduces the temporal intermittency, yet it completely fails to reduce the spatial intermittency. The intermittency of the spatial macroweather transect is even a bit stronger than in the weather regime! This turns out to be the statistical consequence of the existence of geographically distinct climate zones. It is the consequence of huge spatial variability that persists over long periods of time characterizing very slowly changing climate states. Of course, at each location, the long-term temperature fluctuations are combined (correlated) with long-term fluctuations of other variables—notably, precipitation—to yield the climate types familiar from physical geography: "Mediterranean," "temperate," "desert," and so on.

To highlight the relatively small month-to-month changes with respect to these longer term atmospheric conditions,[e] the maps in Plate 5.1 show "anomalies," not absolute temperatures. Just as the daily weather maps (Plate 5.1, left) define anomalies as differences of the daily temperatures (the daily "weather state"[f]) with the current one-month average (the current "macroweather state"), the macroweather series and maps (Plate 5.1, right column) are for anomalies obtained as the differences of the actual macroweather temperatures with the standard thirty-year average. This climate normal[3] implicitly defines the current climate state.[4]

There are thus two conventional averaging periods: one month and thirty years, neither of which has been justified scientifically. Although the use of monthly averages is ubiquitous, it is not even an exact length of time! On the rare occasions when

[e] By definition, the atmosphere is stable in the macroweather regime. What is not so obvious is that each distinct region has significantly different characteristics.
[f] In principle, a "weather state" could be a temporal average of the atmosphere over any weather timescale. However, as a result of the link between space and time, low temporal resolution weather states will also have low spatial resolutions (averaging in time filters out small structures that are short-lived).

attempts have been made at justifying it, they have only appealed to convenience. After all, monthly averages are obviously well adapted to national weather service operations![g] Fortunately, a one-month duration is quite close to the weather–macroweather transition. Monthly averages effectively define "macroweather states." Physically, these are averages over a couple of lifetimes of planetary structures. Macroweather states can then be used to define weather[h] anomalies[5] as the differences between the weather and the current macroweather state.

The weather–macroweather transition justifies the use of monthly averages, but what about the thirty-year climate-normal period used to determine monthly macroweather anomalies? Certainly it is also convenient. Although, over the globe, the monthly absolute reference temperature varies from one region to another by 70°C or more, the anomalies with respect to the thirty-year normal typically vary over a range of only a few degrees.[6] Had Plate 5.1 shown the monthly variation of the actual temperatures, we wouldn't have seen much beyond the seasonal temperature variation.[7]

In Chapter 1, we traced the origin of the thirty-year timescale to the rather arbitrary "climate normal" defined by the International Meteorological Organization (IMO) as the period from 1900 to 1930. When it became clear that the climate was changing, the reference period was updated regularly (at first, every thirty years; now, it is every decade), but its thirty-year duration is still with us.

Fortunately, nearly nine decades later[8] an objective ex-post facto justification was finally found, the basic evidence of which was given in Chapter 2 using both spectral and fluctuation analysis (Figs. 2.3A and 2.4). These analyses showed—at least in the instrumental period (roughly, 1850 to present[i])—that there is a new regime that starts at around twenty to thirty years: the true climate regime. At longer times, the fluctuations increase rather than decrease, so this scale marks the end of macroweather and the beginning of climate. Figure 5.3 shows that the transition timescale[9] varies somewhat with latitude. During the industrial epoch, a value of thirty years is a reasonable overall characterization.

---

[g] An informal minisurvey of colleagues revealed that most of them accepted monthly resolutions purely as a convenience. At best, one opined that averaging over a month "reduces noise," but this only makes sense if one assumes there is a "weather noise" that is objectively distinct from a macroweather "signal." So, to be justified objectively, monthly averaging still requires a weather–macroweather scale break.

[h] These are defined by averaging over a much shorter weather scale. Here, I averaged over a day. Because of the weather regime space–time relationship, averaging in time effectively smooths structures in space. For example, Figure 4.10 shows that daily averages correspond to a spatial resolution of about 900 km.

[i] Before this, some measurements exist but are too sparse or unreliable for global climate studies. It was only during the second half of the nineteenth century that they became sufficiently numerous for most climate applications. Modern campaigns to recover and digitize old naval and other historical records usually don't ~~attempt to~~ go much further back than 1850.
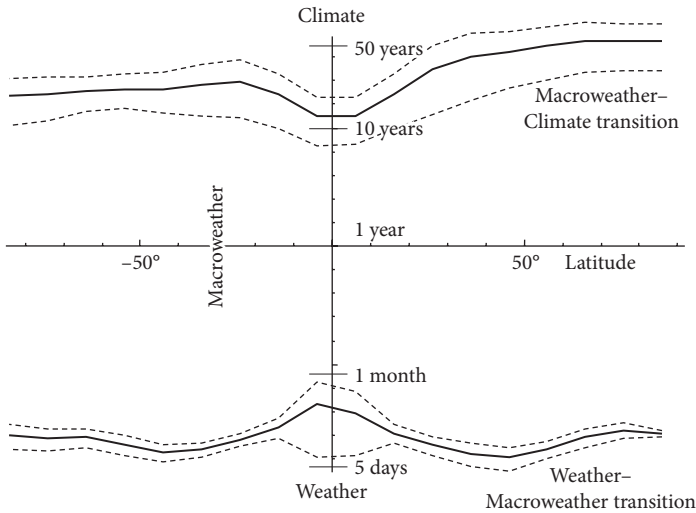
FIGURE 5.3  *The variation of the weather–macroweather transition scale (bottom, an extract of the curves in Fig. 4.11, where it is described) and the macroweather–climate transition scale (this is only valid in the Anthropocene) (top) as functions of latitude (estimated by the position of scale breaks in Haar structure functions from the Twentieth-Century Reanalysis data). The solid curves show the mean over all the longitudes; the dashed curved lines are the longitude-to-longitude variations (the one-standard deviation limits). The macroweather regime is the regime between the top and bottom curves.[10]*

In retrospect, it is hard to escape the conclusion that the choice of thirty years was fortuitous. Using preindustrial "multiproxies" (described later), we will see that, during the 1930s, the destabilizing anthropogenic warming was still too small to detect. For example (Fig. 6.4A), we find that over the original climate-normal period (1900–1930), there has only been a global increase of about 0.1°C as a result of anthropogenic warming.[11] Because this is the same as the typical global scale year-to-year natural variability, it is impossible to detect the anthropogenic "signal" in the midst of the natural "noise." To IMO scientists, the constancy of the first climate-normal period would have been quite plausible.

A straightforward estimate of the macroweather–climate transition scale is given by the time that it takes for the anthropogenic warming "signal" to equal the typical natural variability "noise." During the 1930s, this was about sixty years (see Fig. 6.4A, B).[12] Since then, emissions and other anthropogenic forcings have greatly increased, so that today it takes only about sixteen to eighteen years for the anthropogenic temperature increase to exceed the natural variability. As a consequence, the thirty-year weather–macroweather transition time (Figs. 2.4A and 5.3) is, in fact, only an average over the recent epoch (since roughly 1880).

Now that we have justified the thirty-year averaging period, we can use it to define Anthropocene climate states and see what they look like. Plate 5.2 shows the result using data over the 140-year period from 1871 to 2010. The data were divided into five nonoverlapping twenty-eight-year periods (nearly thirty years) and

the differences (anomalies) are shown with respect to the first of these (1871–1898), taken as the reference period. Not surprisingly, the figure mostly displays warming trends, which are quite variable from place to place, reflecting the corresponding variations in "climate sensitivities" (Chapter 7[j]). Finally, we can consider the intermittency of the climate states. Although there are not enough climate states (four or five at twenty-eight-year scales) to analyze their temporal intermittency,[13] we can study it readily in space using the method of Figure 5.2 to determine the normalized absolute gradients (Figs 5.2 and 5.4). The figure shows that the intermittency is indeed very strong and that it is largely (but not only) associated with coastlines and mountain ranges.
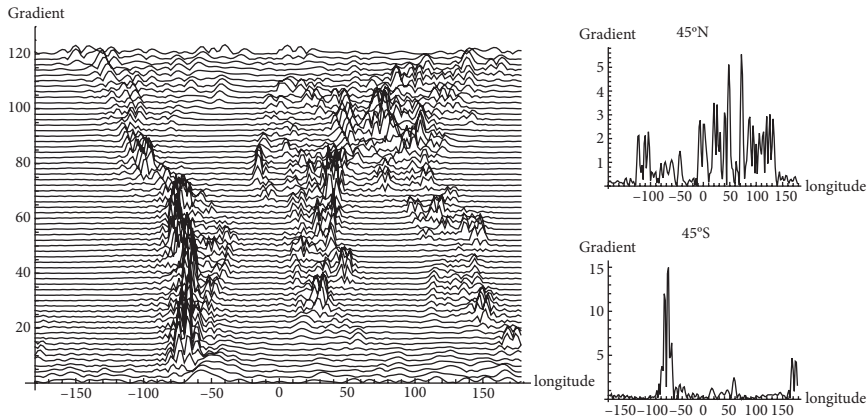


FIGURE 5.4 *The absolute east–west gradients of the temperature climate state obtained by averaging over 140 years, from 1871 to 2010. The data are between 60°S and 60°N and were taken from the Twentieth-Century Reanalysis at 2° spatial resolution. For each latitude, the gradients are normalized by the mean gradient at that latitude. (Left) The gradients from successive latitudes are offset by 2 units in the vertical. We can roughly make out the major mountain ranges and coastlines. (Right) Specific examples at 45°N and 45°S. Note the different vertical scales. They are highly non-Gaussian and intermittent.*

## 5.3  What's macroweather like?

We defined macroweather as the regime where $H < 0$, so we must find out how $H$ varies from place to place. Figure 5.5 shows its spatial distribution. Although as we move from one region to another, its value is variable and it has a remarkable feature: Its range is limited to the interval[14] $-1/2$ to $0$. This range is very significant. First, consider the lower limit, $H = -1/2$. This corresponds to taking a sequence of independent random numbers[15]: "white noise." Time series with larger $H$ values

---

[j] Compare this with Plate 7.8 (upper left), which shows the warming trends over the whole period.

up to $H = 0$ are obtained by increasingly smoothing such white noise. Although white noise obviously cannot be predicted any better than a coin toss—in other words, not at all—the more it is smoothed, the more the neighboring values are likely to be the same, and the more predictable the process becomes (Chapter 7). The upper limit of the empirical range, $H = 0$, is even *infinitely*[k] predictable.[l] It is therefore significant that although the oceans tend to have a typical value of $H \approx -0.1$, land typically has $H \approx -0.3$.[16] Our ability to predict macroweather air temperature over the oceans is therefore much greater than our ability to predict it over land[17] (Fig. 5.5).
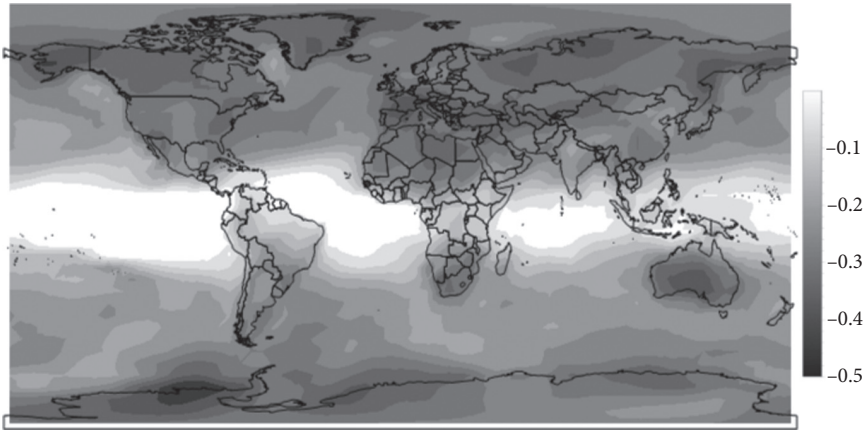


FIGURE 5.5 *The spatial distribution of the fluctuation exponent H estimated from the National Center for Environmental Prediction reanalysis. Monthly anomaly data from 1948 to 2015.*[18]

To get an idea of what typical series with different $H$ values actually look like, Plate 5.3 shows examples of both real macroweather anomalies (right column) as well as simulations using a simple (nonintermittent) mathematical process known as "fractional Gaussian noise" (fGn, left column). At the bottom of Plate 5.3, there is a sequence of independent values with no relation between them ($H = -1/2$, white noise).[19] As one moves from the bottom to the top, the value of $H$ increases. Although fluctuations still tend to cancel as a result of correlations between successive values, the cancelation is less and less perfect, so that for the $H$ values near zero (top), there are coherent undulations present over most of the record. The correlations reflect the long-range memory that can be used for forecasting (see Boxes 7.2 and 7.3).

---

[k] Of course, an infinite amount of past data would be needed to exploit this predictability fully!

[l] Predictability is the limit to which a process can be predicted using the best-possible method with the best-possible past information. As explained in Chapter 7, there are two different predictability limits that are important: the more familiar deterministic predictability limit (resulting from the butterfly effect) and the longer stochastic predictability limits, which are relevant here.

We can now graphically illustrate the sensitivity of the results to the degree of averaging (the temporal resolution), which is a basic consequence of $H < 0$. In Plate 5.3, the original 1,680-point-long series[20] is shown in black. Superposed is the same series, but averaged over a factor of 12 in scale (i.e., at annual resolution; blue) along with the averaged and rescaled series (red). We know from Chapter 2 that when $H < 0$, the effect of the averaging is to reduce the amplitude of the variations by a factor of $12^H$. The rescaling was done by dividing each series by the factor $12^H$ to correct for it statistically. As expected, the averaging without this compensation (blue in Plate 5.3) can strongly reduce the amplitudes.[21] Whereas, the compensated, annually averaged series has comparable variability to the original one-month resolution series.

Although this strong resolution dependence is a basic feature of macroweather, it is neither recognized explicitly nor taken into account either in GCM validation nor in empirical temperature estimates. In the case of monthly and longer GCM series, there is a glaring symptom of resolution dependence: The amplitudes of month-to-month model temperature variations are noticeably smaller than those of the data. But, instead of recognizing this as a resolution issue[22] caused by the nominal GCM resolutions being lower (more averaged) than the data, it is simply treated as yet another model imperfection to be corrected by ad hoc "postprocessing."[23] Not surprisingly, the same resolution issue also turns out to be important in empirical estimates of Earth's temperature: a data–reality mismatch. This is because to make monthly temperature maps on regular grids (e.g., $5° \times 5°$), the raw station and ship data are placed in grid squares, and all the data in the square for that month are then averaged.[m] Some grids have many measurement points whereas some have few, and, typically—since 1880—about half have no data whatsoever![n] Although more than half a dozen of such instrumentally based, centennial-length data sets have been developed, each handles the gridding and missing data issues differently, and this results in different effective resolutions, none of which are exactly the same as the nominal resolution of their gridded products.[o] Because this resolution effect is multiplicative, it is important at all timescales and it ends up dominating the errors in estimates of decadal- and centennial-scale temperature changes that are needed, notably for estimating the amplitude of global warming.[24] Box 5.1 gives more details.[p]

[m] The exact way that this is done depends on the approach, but because of the spatially pointlike nature of the raw in situ measurements, it never corresponds to a true areal average.

[n] This fraction is somewhat variable, depending on the data set in question, but 50% is typical. Data sets with fewer outages typically make various smoothness and regularity assumptions to fill in "holes," and existing hole-filling assumptions are scalebound (e.g., "Kriging"). They are not compatible with the spatial scaling.

[o] The typical magnitude of the effect is about 11% (see Box 5.1).

[p] Global temperature estimates are based on these gridded products so that these multiplicative biases in the gridded data are propagated to the global averages.

BOX 5.1  What's the temperature of Earth?

Estimates of Earth's past and present temperatures are politically charged. The IPCC's 0.85°C estimate of industrial epoch warming is for the long-term change (1880–2012). Its uncertainty of ±0.20°C (with 90% confidence) is a result of the difficulties encountered in separating the centennial-scale change from short-term fluctuations. But, the IPCC numbers are valid only if the basic measurements are accurate enough. So, how accurate are they? In a recent article published in *Climate Dynamics*,[25] I was able to show that, with 90% confidence, the true global monthly temperature lies in the range from –0.109°C to 0.127°C of the reported values, and that the long-term change can be estimated to nearly the same level of accuracy—small enough to justify the IPCC's conclusions. I'll now explain how this was done.

To quantify climate change, long records—preferably at least centennial length—are needed. Over this period, the accuracy debate has focused almost exclusively on potential human bias. For example, thermometers change with technology and, often, so do their exact locations. Such changes are dealt with by numerous comparisons. To start with, absolute biases are eliminated by using station anomalies only. These are differences between the actual temperatures and the long-term averages for the station itself. However, the anomalies may still be biased—for example, when an initially rural site is later urbanized and the average temperature increases as a result of the "heat island effect." In this case, comparisons are made with neighboring rural stations, and potentially biased contributions are weighted accordingly.

Although there is no question that many adjustments are required,[q] there is no absolute truth for validating them. This has allowed climate skeptics to accuse scientists regularly of selectively correcting the data to exaggerate the warming. Indeed, a new breed of "lukewarmers" have accepted that there is some warming, but claim that it is too small to worry about. The problem for the lukewarmers is that other independent data[r] sources—such as satellite measurements—show nearly identical overall trends, so we can be confident that human-induced biases are small. But how small?

Ironically, the spotlight on human biases and errors is misplaced. It turns out,[26] that there are two much more important sources of error that have virtually escaped attention: the unexpected long-term consequences of missing data and biases arising when thousands of sparsely distributed measurements from the oceans and continents are combined to produce a single globally averaged value. Neither of these problems is human in the usual sense. They are consequences of huge variability in atmospheric temperature arising from "whirls" ranging in size from millimeters to the size of the planet, evolving over timescales from milliseconds to the age of Earth. A modern home thermometer can tell us the temperature in our backyard with an accuracy of a tenth of a degree, but how can we estimate the temperature of a city? Of a country? Of the whole planet? We don't have thermometers everywhere, so how do we infer averages over large regions, and how accurate is our result?

---

[q] Even high-quality surface networks, such as the US Historical Climatology Network, "have an average of six discontinuities per century and not a single station is homogeneous for its full period of record." Peterson, T. C. Assessment of urban versus rural in situ surface temperatures in the contiguous United States: No difference found. *J. Climate* **16**, 2941–2959 (2003); quote, p. 2942.

[r] The 20CR series used here and elsewhere in this book is an example. It used only surface pressure and monthly SST measurements. There were no station temperature measurements at all.

To quantify this, six different globally,[27] monthly averaged temperature series were used from 1880 to 2012.[28] Figure 5.6 (top) shows the typical fluctuations averaged over all the series—the same analysis[29] as in Figure 5.14, which shows the slow macroweather decrease (with $H = -0.1$) followed by the increase at scales of twenty years or so, which is the climate regime associated with anthropogenic warming. The middle curve in Figure 5.6 shows the average *differences* of the series taken pairwise (there are fifteen independent pairs from the six series), and the bottom curve displays the fluctuations of the average difference of each curve with the mean of all six curves. We can see that these fluctuations fall off very slowly, indicating long-range statistical dependencies (strong correlations, "memory"). In comparison, the dashed line at the far left (bottom) shows the behavior that would be expected if the differences were the result of measurement errors that were independent or otherwise of short range [such as the autoregressive (AR) or moving-average (MA) processes discussed in Chapter 7].
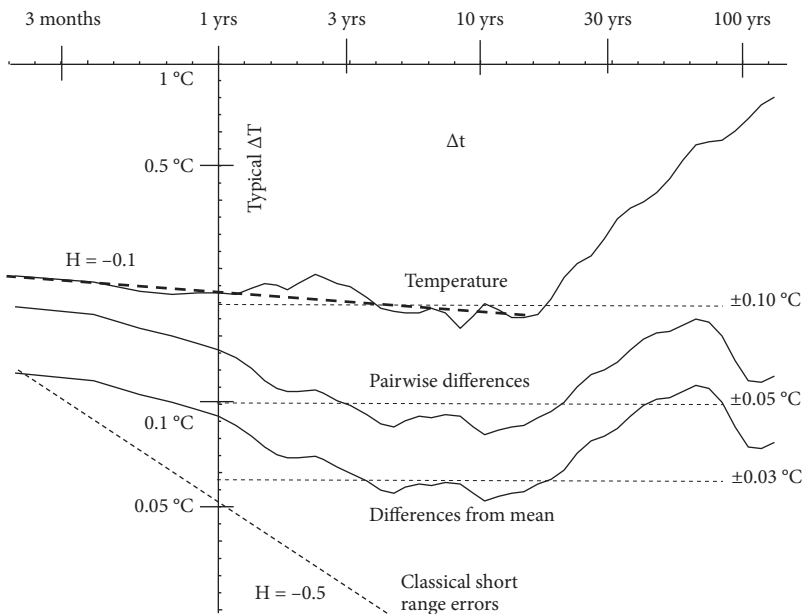


FIGURE 5.6 *The root mean square Haar fluctuations [structure functions, S(Δt)] averaged over the six series (top), averaged over all the fifteen pairs of differences (second from top), averaged over the differences of each with respect with the overall mean of the six series (third from top). Also shown for reference (dashed) is the line that data with independent Gaussian noise would follow.*[30]

The (roughly) scaling, long-range statistical dependence of the differences cannot be explained by standard (short-range) measurement errors, but instead through two scaling effects: missing data and the scale reduction factor resolution effect (Plate 5.3). The missing data problem is easy to understand. Since 1880, for monthly temperatures on 500-km-sized grid boxes, more than half had no data, implying a large uncertainty in the global estimates. Here, the new element is the recognition that this would affect the accuracy of estimates over long periods: years and decades. The resolution effect is a space–time generalization of the temporal resolution effect

in Plate 5.3, which is a consequence of $H < 0$. When pointlike data are "massaged" onto regular grids, there are often insufficient data to average them adequately to their nominal values (e.g., 500 km and one month). Although missing data turned out to be the main source of error for timescales less than a decade or so, I found that the resolution effect dominated the uncertainties at the centennial scales, which are relevant for industrial epoch warming. In comparison, standard human errors were negligible for periods of months and longer.

In conclusion, with 90% confidence, the temperature change since 1880 is correct to within 0.108°C of the true change. This is less than 13% of the IPCC estimated warming. Measurement errors are thus too small to alter the conclusion that we are living through a period of huge warming and that it is occurring over an unprecedentedly short period of time (see Section 6.3).

---

The geographical distribution of $H$ values (Fig. 5.5) is only part of the spatial variability. When we move from one region to another, not only does the $H$ value change, but also the amplitude of the fluctuations changes as well. Figure 5.7A shows a typical space–time plot. Even a visual impression of the density of the lines shows that the amplitudes are far from uniform from one location to another; the exponents themselves are also spatially varying.



(A)

Space
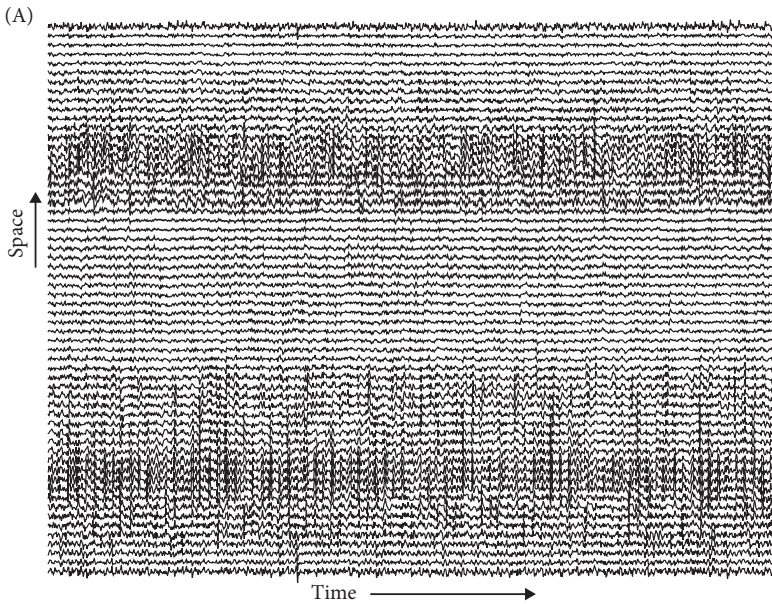
Time

FIGURE 5.7  *(A) Temperature anomalies (time, from left to right) for the 1,000-month period from 1926 (left) to 2010 (right) as a function of longitude (every 6° along the 45°N parallel, from longitudes 0° to 354°, from bottom to top), from the Twentieth-Century Reanalysis. Each series is offset in the vertical. Notice that the fluctuation amplitudes as well as exponents H depend on position. (B) A stochastic macroweather simulation using complex cascades displaying temporal fluctuations (left to right), with both amplitudes and H values that depend on position (bottom to top). In this simulation, both are random. The model is intended to simulate an Earthlike planet with similar space–time macroweather statistics.*

(B)
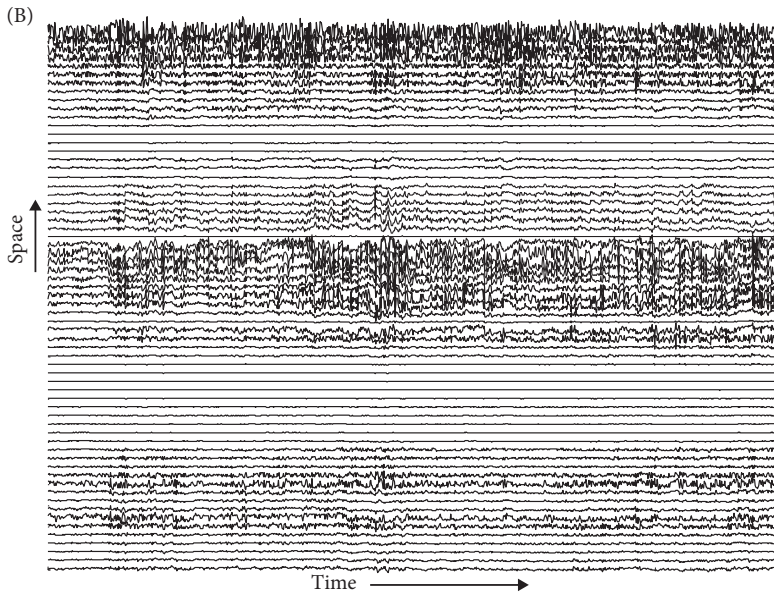


Space

Time

Figure 5.5 and Plate 5.3 underline the fact that a full understanding of macroweather requires knowledge not only of the spatial and temporal variability as a function of scale, but also of the joint space–time variability. Although little is yet known, there are theoretical reasons (from cascade models), numerical reasons (from the analysis of GCM outputs), and empirical reasons (notably from temperature and precipitation data) to suppose that macroweather obeys a symmetry known as "space–time statistical factorization" (STSF[31,s]). Such factorization means that different (spatially distributed) climate zones modulate the local temporal statistics without changing their type (e.g., their temporal scaling), which explains the spatial variability of the amplitudes in Figure 5.7A. However, the only way to combine factorization with variable exponents $H$ turns out to be to make space–time models in which the $H$ values themselves are random. Figure 5.7B shows one such method based on cascades.[32]

This factorization principle is already used implicitly by weather services to "homogenize" data sets[33] or to produce various climate indices that allow regions with different climates to be compared. Let's look at an example. The northern part of Portugal is roughly ten times wetter than its southern region. Let's compare monthly precipitation series to determine whether they are both undergoing similar long-term climate changes. In the dry south, a ten-year rain rate change of 10 mm/month would be highly significant, yet in the wetter north, it would simply correspond to the expected natural variability. A consequence of factorization is that if we divide (normalize) each series[34] by the typical amplitude of its

---

s In Chapter 7, I discuss how STSF greatly simplifies macroweather forecasting.

fluctuations, then the two can then be compared on an equal footing (hence the term "homogenization").

## 5.4 Why not "microclimate?"

"Weather is the state of the atmosphere, to the degree that it is hot or cold, wet or dry, calm or stormy, clear or cloudy."[35] If we combine this definition with the usual view of the climate as average weather—"the climate is what you expect; the weather is what you get"—then there is no qualitative difference between them. The more that one averages the weather, the more climatelike the result becomes. In the long term, one finally obtains *the* climate. But if the weather–climate transition is simply a question of the subjective averaging period, then how can the climate *change*?

Clearly, without objective definitions, our understanding is limited and scientific progress is hindered. In retrospect, Van der Hoven's discovery of the "synoptic maximum"—which he spectacularly misrepresented as a scalebound spectral bump[36] (Fig. 4.2), and its subsequent relegation to a minor role—was an unfortunate lost opportunity to clarify this.

We have argued that to tame atmospheric variability, we need a scaling framework, and Figures 2.3A and 2.4A showed that, out to Ice Age scales, there are three—not two—regimes. There is no question that the high frequencies correspond to our common idea of weather, but what about the other two? One option would be to insist that climate immediately follows the weather. In this case, slow variations over scales of thirty years and longer must be something else: "macro" ("big") climate? This term would lead to some bizarre uses. For example, the "climate" would change from month to month, and "long-range" weather forecasts (such as monthly predictions) would already be "climate forecasts."[t] When referring to periods of thirty or more years in the past, we would talk about "past macroclimates." Ice cores and other paleo proxy data would be "macroclimate indicators," and the fight against global warming would be a struggle to ~~stop~~ "macroclimate ~~change~~."

To be consistent with common parlance, we should probably reserve the term "climate" for the third, longer timescale regime. This leaves us with the middle regime. Why call it "macroweather" rather than "microclimate?" To answer this, we have to consider how climate states can change, and both GCMs and stochastic cascade models help us to understand this. GCMs are based on partial differential equations that embody the physical processes governing the atmosphere.

---

[t] Admittedly, the current term "long-range weather forecasts" is untenable and needs to be replaced by "macroweather forecasts" (Chapter 7).

Similarly, cascade models are based on the corresponding, emergent higher level statistical laws.[37] Both were originally weather models that were later extended to include the oceans, the cryosphere, and the carbon cycle, and both can be used to define the climates of *models* objectively.

Let's focus on the more highly developed of the two approaches—namely, the GCMs. At some initial instant, the state of the atmosphere (and in coupled models, the state of the ocean) is specified everywhere, and then the model is integrated forward to determine its later states. Mathematically, it is an "initial value problem." We know from some theory—going back to Lorenz's inverse cascade of forecast errors[38] (see Chapter 7) during the 1960s and also by direct analysis of GCM outputs—that the errors in predicting large-scale structures double every ten days or so. This error doubling time already provides a kind of operational definition of the weather. At scales longer than this, the results can no longer be interpreted deterministically because any small (microscopic) error doubles and rapidly destroy the forecast.[39]

Unlike the real atmosphere, GCMs can be integrated forward from their initial conditions with fixed solar output, fixed atmospheric composition (e.g., greenhouse gases), no volcanism, and no land-use changes (e.g., no deforestation)—in other words, with constant external conditions. Runs with everything external to the atmosphere fixed are called "control runs." In control runs, the atmosphere displays only "internal variability." It varies in a quasi-steady manner around a long-term state: the model's "climate." For each forcing, control runs thus both define the model's climate and also determine the rate of its convergence to its climate.

Using 500-year-long control runs from eleven GCMs, Figure 5.8 shows the rate at which their average temperature fluctuations[40] converge to their climates. As expected, as a result of the temporal scaling symmetry, one finds that the models have excellent scaling (straight lines in Fig. 5.8), and the convergence rate is quantified by the slope of the line, $H$. From Figure 5.8, we can also see that after 300 simulated years, the global temperature still typically fluctuates by ±0.06°C about[41] its long-term climate.[u] To determine the model climates better, we could extend the simulations to one million years, but because $H$ is near zero ($H \approx -0.15$[42]), this convergence is "ultraslow."[43] Temperature averages over a

---

[u] This ultraslow convergence leads to technical difficulties in determining the model's climate. This is because there is an initial "spin-up" period during which the initial state of the atmosphere—and especially oceans—are "spun up" to adjust to their long-term quasi-steady state. Although spin-up times of several (simulated) centuries are common, even these are not enough to prevent "model drift" [i.e., slow variations such as rising temperatures that are attributed to deep (and slow) ocean currents]. (They cannot be attributed to global warming because all the external conditions are fixed.) In Figure 5.7, overall drifts were approximately removed by linear regression. Only the residuals, after removing these linear drifts, were analyzed.

million simulated years would still typically fluctuate by ±0.02°C from the model's true climate![44]
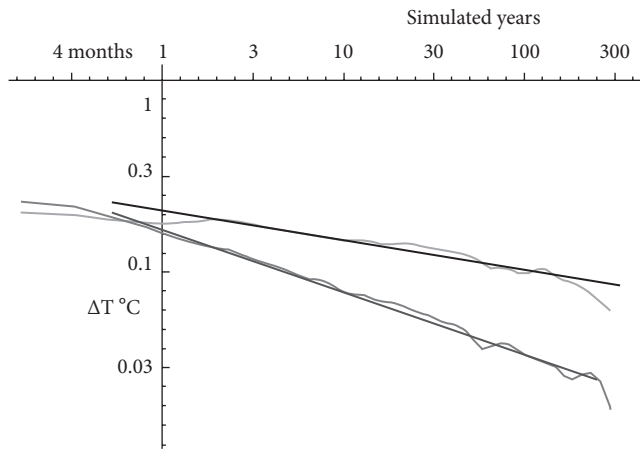


FIGURE 5.8  *(Top black line) The Haar fluctuations for eleven control runs from the Climate Model Intercomparison Project 5.[45] The reference slope is H = –0.15. (Bottom black line) Haar fluctuations from (multifractal) cascade simulations at a resolution of one day for 340 years. The reference slope is H = –0.32.*

In Chapter 4, we saw that GCMs had fairly realistic weather regime statistics. Comparing Figure 5.8 with Figures 5.6, 5.9 and 5.14 shows that the value $H \approx -0.15$ is actually pretty close to the observed global macroweather temperature variations (i.e., up to about twenty to thirty years). GCM control runs thus reproduce the basic weather and macroweather statistics accurately. Similarly, turbulent cascade models that were designed to reproduce weather statistics (including multifractal intermittency) also reproduce weather and macroweather, although with a somewhat more negative[46] $H$ value (Fig. 5.7). When extended to scales beyond the lifetimes of planetary structures, these (essentially) *weather* models[v] thus naturally produce macroweather, so we can finally conclude that the term "macroweather" not "microclimate" is indeed appropriate.

---

[v] For our current purposes, there are no essential differences between NWP models and GCMs.
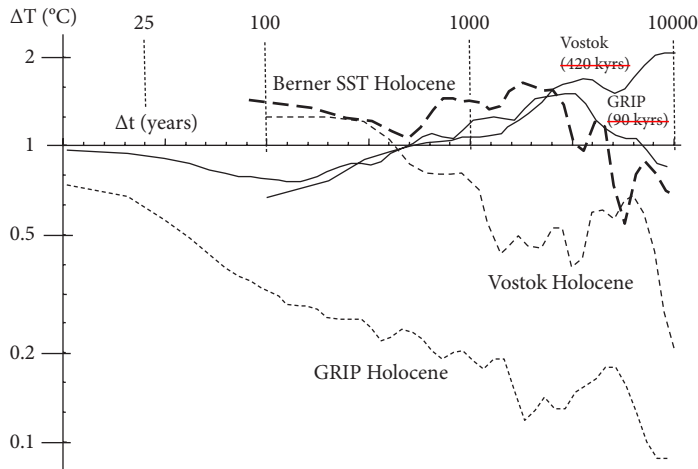
FIGURE 5.9 *A comparison of the root mean square Haar fluctuations for both Vostok and Greenland Ice Core Project (GRIP) cores at resolutions of five years and fifty years, respectively, over the last 90,000 years. These series were broken into 10,000-year sections. The dashed lines show the most recent of these (roughly, the Holocene). The paleo sea-surface temperature (SST) series is the upper dashed black line.[47] The solid lines are of the ensemble of the eight 10,000-year periods from 10,000- to 90,000-years BP from GRIP and Vostok cores.[48]*

## 5.5  Climate states, climates of models, and historical and Last Millennium simulations

We defined the climate as the long-time scale regime beyond macroweather when fluctuations begin to rise again with scale. Climate states are thus averages over weather and macroweather timescales, averaging up to the critical macroweather–climate transition scale. Defined in this way, climate states thus have relatively straightforward definitions and can be determined readily from data. Plate 5.2 displayed the evolution of climate states in the recent period. The closest alternative definitions are theoretical and model based. They are synonymous with the climate of a GCM control run averaged over a very long time. For example, for the GCM climate, Bryson[49] states: "Climate is the thermodynamic/hydrodynamic status of the global boundary conditions that determine the concurrent array of weather patterns." He explains that although "weather forecasting is usually treated as an *initial value* problem . . . climatology deals primarily with a *boundary condition* problem and the patterns and climate devolving therefrom."[50] This could be paraphrased as: "For given boundary conditions, the climate is what you expect." Bryson's definition of climate is essentially the same as a control run definition because for fixed boundary conditions, the climate is only uniquely defined if averaged for long enough. Because in GCM control runs macroweather extends to infinitely long timescales, the climate state of GCM with fixed conditions is the same as its climate.[51]

Although this control-run definition may be fine for models, it is not so helpful in the real world. To start with, even for identical, fixed boundary conditions each different GCM generally has a control-run climate that is quite different from that of other GCMs, and they can't all be realistic. To test, validate, and improve them, we need to be able to compare model climates with the real world. The usual tests involve trying to keep the control-run idea by comparing long-time averages of the data and models. However, there are several reasons why data and model averages might disagree. To start with, the higher frequency weather and macroweather regimes might be different as a result of various model limitations or approximations used in modeling weather-scale processes. Next, for realism, the models would have to include all processes relevant to the climate. Perhaps, because of our limited understanding, the model might have missed some slow decadal-, centennial-, millennial-scale processes. In this case, even if—at the higher macroweather frequencies—model and data agreement are satisfactory, they would nonetheless disagree over the longer term climate scales. In other words, thanks to processes that are too slow to be relevant at annual or decadal scales, model climate states might still be quite different from real-world climate states.

Further problems arise even when all the "internal" (both fast and slow) weather and macroweather processes have been included and are modeled realistically. The trouble is that, to understand and model climate change, we need to change the boundary conditions. Changing "boundaries" means varying the solar output and allowing volcanoes to erupt and humans to intervene. These external changes to the atmosphere force the system to evolve. They are "climate forcings." This terminology can sometimes be a bit confusing because, for example, the standard solar radiation reference level (the long-term average) is *not* a climate forcing. However, the 1,000-times smaller deviations from this level—resulting from sunspots or other solar activity that changes the solar output—*are* climate forcings. Because of these slow forcings, at decades and longer we have a superposition of long-time natural (unforced) macroweather fluctuations superposed on the slow and delayed response of the system to the forcings.[52]

In principle, the "internal" and "forced" responses might interact nonlinearly, making the problem much more difficult.[w] In Box 7.1 we discuss this further with the help of Earth's energy balance and show that, in reality, the responses to the forcings are nearly linear,[x] and also how scaling can be exploited to estimate the climate response. The standard way to quantify the responsiveness of the climate to forcings is through equilibrium climate sensitivity (ECS). ECS is defined as the change of the control-run mean climate temperature following a~~n instantaneous~~ doubling in $CO_2$ concentrations. To determine ECS, one runs the model with fixed $CO_2$ and other boundary conditions for long times, then doubles the $CO_2$ concentrations, and then further integrates the models, waiting for their

---

[w] Nonlinear interactions potentially include tipping points and are already modeled by GCMs. However, in Chapter 7, we will see that over the standard forcing scenarios, the models behave essentially linearly. This allows for alternative "historical" scaling-based methods of climate projection.

[x] This is because the actual forcings are fairly small. Even doubling the $CO_2$ amount of would increase the power (energy per time) in the atmosphere by less than 2% (see Chapter 6).

(ultraslow) convergence to their new climate state. The difference in global mean temperatures of the long-time average before and after the doubling is the estimate of ECS. Because of its definition in terms of control-run temperatures, ECS is a somewhat academic notion.[53]

In view of the problems discussed in validating a control-run definition of climate states empirically, assessing GCM performance is difficult. Naturally, the main scrutiny has come from "historical simulations"—conventionally from 1850 to the present. They focus on reproducing anthropogenic effects. A related, interesting approach is to make long historical-based simulations since the year 1000 AD: "Last Millennium" simulations. Figure 5.10 shows the result when this is done using NASA's Goddard Institute for Space Studies (GISS)[54] model. To make the simulations, one must use appropriate boundary conditions—notably, solar output and volcanic eruptions, and, more recently, land-use changes, greenhouse gases, and aerosols (pollution).[55] For most of the period since year 1000, estimates (called "reconstructions") of these "forcings" are in themselves highly challenging. Direct instrument records don't exist; the forcings must be inferred directly from paleoindicators (see Box 5.2).
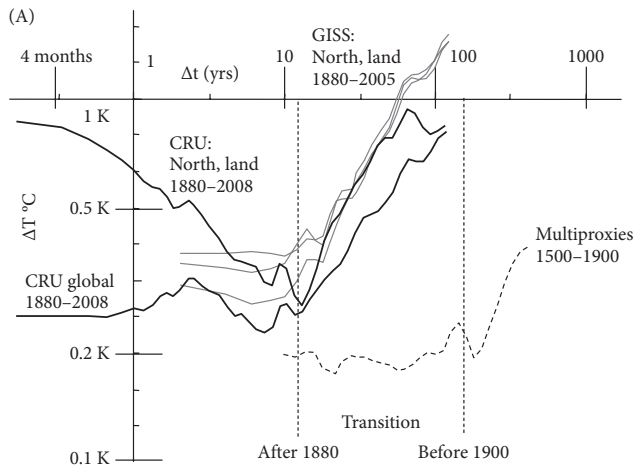


FIGURE 5.10 *(A) Haar fluctuation analysis of temperature fluctuations (black lines)—these data were from the Climate Research Unit Hadley temperature version 3 database: HadCRUtemp3—and globally, annually averaged outputs of past millennium simulations (gray lines) over the same period (1880–2008) using the National Aeronautics and Space Administration's (NASA's) GISS model with various forcing reconstructions. Also shown (dashed line) are the fluctuations of the preindustrial multiproxies, displaying the much smaller centennial- and millennial-scale variability, which holds during the preindustrial epoch. The dashed vertical lines indicate the transition scales for the industrial and preindustrial epochs. (B) Haar fluctuation analysis of globally, annually averaged outputs of Last Millennium simulations over the preindustrial period (1500–1900) using the NASA GISS model with various forcing reconstructions. Also shown (thick black lines) are the fluctuations of the preindustrial multiproxies, indicating they have stronger multicentennial variability. Last (bottom, thin black lines) are the results of the control run (no forcings), showing that macroweather (slope < 0) continues to millennial scales. The reference line has slope –0.2, nearly the same as that in Figure 5.5, which is the average over many different model control runs.[56]*
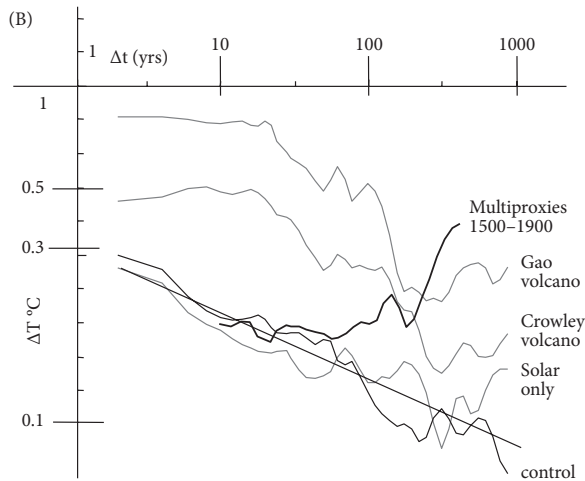
FIGURE 5.10  *Continued*

BOX 5.2  Volcanic and ~~solar forcings~~

Since 1880, the anthropogenic forcing has grown and is currently close to 2 W/m² (see Fig. 6.5B). How does this compare with the main centennial- and millennial-scale natural forcings, both solar and volcanic?[57] To answer this, we need to "reconstruct" these forcings, and this must be done primarily from indirect sources. For example, systematic, direct satellite observation of the solar constant has only been made since the 1980s and, as a result of satellite calibration drift, the critical decadal and lower frequencies are unreliable.[58] One way to reconstruct past solar variations is to use the correlations between sunspots and solar activity observed over the past few decades. Because reliable sunspot records[y] exist only since 1610, this is the limit of sunspot-based reconstructions. Figure 5.11A shows the "background" solar anomalies (with the eleven-year solar cycle removed). We see that the inferred total solar irradiance (TSI) seems to wander. Figure 5.12 confirms the visual impression that that $H > 0$ for this signal (it has $H \approx 0.4$).

[y] Sunspots were discovered by the Chinese as long ago as 800 BC. Unfortunately, systematic, quantitative records had to await the discovery of the telescope and date from 1610 AD.
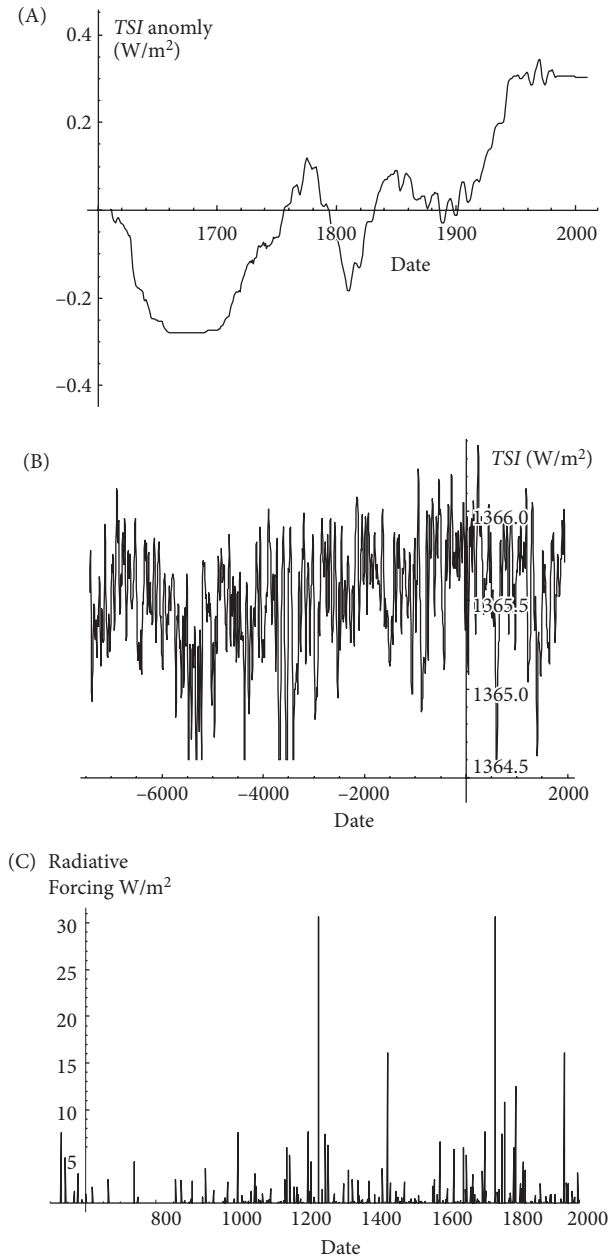
FIGURE 5.11 *The solar forcing (total solar irradiance) measured in watts per square meter reconstructed from sunspot records since 1610.[59] The eleven-year solar cycle has been removed; only the background is shown. Notice the wandering character, H ≈ 0.4 (see Fig. 5.12). (B) The solar forcing.[60] (C) Volcanic forcing reconstruction from the year 500 AD.[61] These forcings are, in fact, negative (cooling), but have been plotted in the positive sense to emphasize the similarities with other highly intermittent processes discussed throughout this book.*
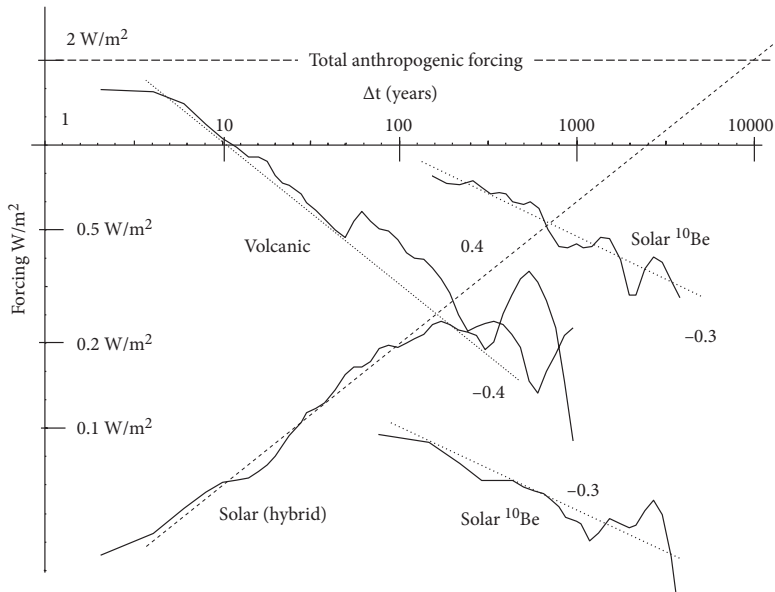
FIGURE 5.12 *A comparison of typical fluctuations (estimated by the root mean square of the fluctuations) of the typical solar and volcanic forcings (measured in watts per square meter) used in Last Millennium simulations (e.g., Fig. 5.10B). The volcanic reconstruction is from Figure 5.11C; the bottom right $^{10}$Be solar is from Figure 5.11C. The top $^{10}$Be curve is based on the same data, but with a different model used to infer solar activity from the same $^{10}$Be fluxes. The solar reconstruction at the lower left is a "hybrid" that uses the sunspot-based reconstruction back to 1610 (Fig. 5.11A), and the lower right $^{10}$Be series is at longer timescales. All the curves have straight, dashed reference lines with the slopes (H values) indicated. The total anthropogenic forcing since 1880 is indicated for reference as the top horizontal dashed line.*[62]

Long reconstructions going back nearly 10,000 years are possible but use extremely indirect techniques, the main ones being based on concentrations of the isotope beryllium 10 ($^{10}$Be) found in ice cores. $^{10}$Be is deposited by the solar wind and its concentration varies depending on the solar magnetic field, which allows either more or less $^{10}$Be to escape from the sun, and the solar magnetic field in turn depends on overall solar activity. Solar models are then used to relate the $^{10}$Be fluxes to the TSI. Figure 5.11B shows a reconstruction with a forty-year resolution. From Figure 5.11B, we can immediately notice the canceling character of the signal, and Figure 5.12 confirms that $H < 0$ ($\approx -0.3$). The problem is that, although the sunspot-based reconstructions indicate that solar forcings increase with scale, on the contrary, the $^{10}$Be reconstructions decrease with scale, so they can't both be right. To make matters worse, the amplitude of the reconstructed forcing depends greatly on the model assumptions. Figure 5.12 shows that the same $^{10}$Be data can lead to a difference in reconstructed forcing amplitudes of nearly a factor of 10!

In practice, Last Millennium simulations are based on a "hybrid" series that splices the sunspot-based TSI estimates back to 1610 with the lower resolution $^{10}$Be

reconstruction back to the year 850. Because the resolution of the $^{10}$Be reconstruction is much lower, it contributes nothing to the variability below about $2 \times 40 = 80$ years. As a consequence, up to nearly centennial scales, the fluctuations of the hybrid series in Figure 5.11 are dominated by the $H > 0$ sunspots, whereas the longer scale scales are dominated by the $H < 0$ $^{10}$Be data. This analysis shows that the centennial-scale fluctuations of the hybrid are maximal at around 0.2 W/m² at about centennial scales (i.e., ±0.1 W/m²). This is about a tenth of the anthropogenic warming indicated by the dashed line at the top of Figure 5.11, and it explains why solar variability has not much affected the climate for the past century. It is interesting to speculate that the sunspot-based reconstruction with $H \approx 0.4$ continues to much lower frequencies. Extrapolation of the solar hybrid curve in Figure 5.12 to long times implies that, at 10,000 years or so (where the dashed lines cross), the typical forcing fluctuations would be about 2 W/m² (±1 W/m²), resulting in something like ±0.3°C of global temperature variation.[63] However, even during the preindustrial epoch, this is probably still too weak to explain the macroweather–climate transition at scales of millennia (see Fig. 5.7 and the discussion in Section 5.6).

Volcanoes are also important. Volcanic ash generally reflects visible sunlight. Big eruptions can throw huge quantities of material into the stratosphere, where their cooling effects can last for several years, and the eruptions appear to be clustered.[z] Figure B5.11C shows a reconstruction based on volcanic dust from ice core records. The radiative forcing has been estimated with the help of radiative transfer models. The reconstruction is incredibly intermittent and the largest eruption[aa] (in 1257) shows a very large (30 W/m²) cooling effect. But does this and other volcanic spikes affect long-term temperatures? The obvious way to answer this is to look, in the usual way, at the fluctuations in volcanic cooling as a function of timescale. Figure 5.12 shows that, not surprisingly, the fluctuations are indeed roughly scaling and decrease rapidly with scale ($H \approx -0.3$),[64] so that much beyond centennial scales, they can be neglected compared to the solar forcing.

---

Given the forcings, the model still needs to be "spun up." This is the technical term for the fact that the model initial state can only be guessed at; the model itself is run for a long period (here, 150 simulated years, from 850 AD), so that the model generates plausible atmospheric and (more importantly) oceanic flows, which take a very long time to reach realistic patterns. This long period is needed simply to start the more trustworthy part of the simulation; it is a symptom of the ultraslow model convergence. Nevertheless, just as with the control runs, even after the spin-up period, the model tends to "drift," with temperatures changing spuriously and slowly over long periods. In Last Millennium simulations, this has been corrected approximately by the modelers using ad hoc "postprocessing."

---

[z] Interestingly, the clustering is not necessarily a consequence of correlations between successive eruptions. It may be simply a consequence of the extreme black swan intermittency statistics.

[aa] Interestingly, the volcano associated with this eruption (Mount Samalas) was only identified in 2013.

AQ: Please verify figure numbers with new numbering scheme.

These Last Millennium runs were divided into industrial and preindustrial periods, using 1880 as the dividing line.[65] Although this is somewhat arbitrary, it corresponds to the beginning of the period of steep increase in fossil fuel use and also to the period when the instrument record begins to be more reliable and complete. Figure 5.10A shows the GISS Haar fluctuations over the industrial period (thin lines) compared with those from an instrument data set, both global and northern hemisphere (thick lines). The three simulations all include the same greenhouse gases and solar reconstructions. They are distinguished by three different natural forcing scenarios: the two different volcanic reconstructions (top: "Gao" and "Crowley," for their authors; Box 5.2) and (at the bottom) solar variability only—no volcanoes.[66]

By comparing the results with the instrument fluctuations (solid lines in Fig. 5.10A), we see that the variability of the simulations is fairly realistic and that it doesn't make much difference which volcanic reconstructions are used—or—whether any are used at all (thin bottom line in Fig. 5.10A). From the figure, we can conclude that, since 1880, greenhouse gases dominate the natural forcings, and that the timescale (twenty years) at which the natural fluctuations become dominated by these anthropogenic forcings—the industrial epoch macroweather–climate transition scale—is reproduced fairly realistically. Also shown for reference is an estimate of the preindustrial variability deduced from temperature proxies (discussed later), which show that the industrial epoch variability is indeed quite different from the preindustrial one.

Now compare this with the same Last Millennium simulations, but over the preindustrial epoch (1000–1880)—notably, with virtually no greenhouse gas forcings (Fig. 5.11B; i.e., the atmospheric composition is fixed to preindustrial levels). We see that at scales up to about a century, the volcanic-induced variability is much too high, and then at scales of several centuries, that strong volcanic forcings are not enough; the long time variability has become too small. We also see that, by itself, the solar forcing is so small that the solar-only run is essentially indistinguishable from the control run (no forcings at all). If the thick black line marked "Multiproxies 1500–1900," representing ground truth, is to be believed, the preindustrial macroweather–climate transition occurs at about centennial scales (the vertical dashed lines in Fig. 5.10A). We see that contrary to the fairly realistic industrial epoch variability, that before 1900, there is a problem of missing multicentennial variability (discussed later).

## 5.6  Is civilization the result of freak macroweather?

"The long, stable Holocene is a unique feature of climate during the past 420,000 years, with possibly profound implications for evolution and the development of civilizations." So concluded an influential analysis by Petit et al.[67] of the 3,300-m-long Vostok (Antarctic) ice core record that spans four complete Ice Age cycles (see Fig. 1.5B). The second half of the sentence assumes there is

a direct link between stable temperatures since the end of the last Ice Age (the Holocene) and the development of civilization. The link is that farming developed less than 1,000 years after the retreat of the ice sheets, and agriculture was crucial for transforming society.

In the article by Petit et al., Holocene stability was inferred by eyeballing the Vostok record over the recent period. Its uniqueness was noted by comparing it with temperature plots over similar periods earlier in the same record. Estimating fluctuations on the same data (Fig. 5.9, top, dashed curve) confirms the stability quantitatively. We see that the stable $H < 0$ regime extends at least to several millennia, implying a preindustrial macroweather–climate transition at multimillennial scales (or more). For comparison, Figure 5.9 also shows the Haar fluctuations averaged over the previous 80,000 years of the same core, indicating an average $H > 0$ (unstable) behavior for scales as short as one hundred years.[bb] Turning to another long ice core, this time from Greenland (the GRIP core at five-year resolution; Fig. 1.5B, middle), it turns out that the Holocene is apparently even more exceptional (Fig. 5.9, bottom, dashed curve), with the pre-Holocene GRIP and Vostok (top, thin solid) curves agreeing remarkably well with each other.[68] The radical Holocene uniqueness can be confirmed simply by visual inspection. Figure 5.13 shows earlier 10,000-year segments (top three segments) each with strong variability, each "wandering" (due to $H > 0$) over a total range of about 5°C, evidently quite different from the fourth (Holocene) series. Using the same data to determine the spikiness (Fig. 5.2, upper right plot), we see that the recent period (at the left, from 0 up to about 50 units = 12,000 years) also has anomalously low intermittency.[cc]

---

[bb] As we go back into the past, we use information from deeper and deeper parts of the core, and these data suffer increasingly from compression and, ultimately, molecular diffusion that spoils the distinct layers that define the temporal resolution. A general feature of the Antarctic is that it is extremely dry, so that snow accumulates very slowly. In general, the resolution of temperature proxies is thus not very high—here, around one hundred years. Dust does not suffer from the same problem and has thus been measured at decadal timescales back 800,000 years in another Antarctic core (EPICA; see Fig. 1.5C).

[cc] The Holocene is an interglacial and it seems that all interglacials—at least in Antarctic cores—are exceptionally stable parts of the Ice Age cycle. This is the conclusion of a study using the dust data (Fig. 1.5A), which is the only data of sufficiently high resolution to permit detailed analysis of each phase of each cycle. Lovejoy, S. & Lambert, F. ~~High resolution EPICA ice core dust fluxes: Intermittency, extremes and Holocene stability~~ *Climate Past* (submitted, ~~August~~ 2018). The analysis shows that the macroweather–climate transition scale is indeed systematically longest in the interglacials—typically about 2,000 years—and that in the Holocene the transition time is the longest of all eight interglacials, with a transition at about 7,900 years. Although this Antarctic dust record has the longest stability duration of any interglacial, it is apparently only a little bit extreme, not an outlier.

Before concluding that the Holocene is miraculously stable, consider a final core, this time from ocean sediments off the coast off Greenland and only 1,500 km from Summit, Greenland, where the GRIP core was taken (bottom segment, Fig. 5.10). The ocean sediment temperature proxy is based on the ${}^{18}$O isotopes in layers of planktonic foraminifera that sank to the ocean bottom after living their lives in near-surface waters, taking their temperature signal with them. Although the temporal resolution of the data is lower (about one hundred years), its character, including its range of temperature variations (Figs. 5.7 and 5.10), is very similar to the pre-Holocene ice cores. The difference in appearance between this SST proxy and the ice proxies—the patently unstable character of the former—led
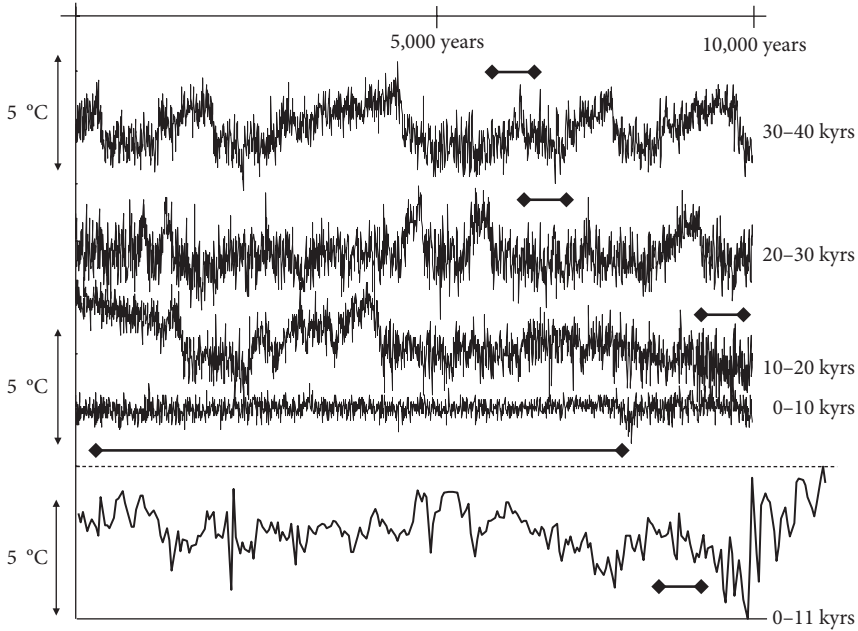
FIGURE 5.13 *The top part of the figure shows four successive 10,000-year sections of the five-year resolution Greenland Ice Core Project (GRIP) data (from Fig. 1.5B, middle), with the most recent to the oldest from bottom to top above the horizontal dashed line. Each series is displaced in the vertical for clarity, and the original data (in parts per thousand of isotope excess) were converted to temperatures using the official calibration. The resulting 5°C spread is shown by double-headed arrows on the left (using a calibration constant of 0.5K per mil). When compared to the previous 10,000-year sections (Fig. 5.9), we see that the bottom Holocene GRIP series is indeed relatively devoid of low-frequency variability. This is quantified by the solid diamond-headed horizontal lines that indicate the estimated duration of the macroweather–climate transition time. Shorter segments of the curves tend to be canceling (H < 0), whereas longer segments tend to wander (H > 0). We can see that in the ~~bottom~~ Holocene record, the transition time is many millennia, whereas for the earlier sections it is only on the order of a few centuries.*
*In contrast, the bottom curve shows a paleo sea-surface temperature record from ocean core LO09-14[69] taken from a location only 1,500 km distant and displaying far larger variability with a multicentennial transition time.[70]*

Both ice core and ocean proxies are generally regarded as robust, trustworthy paleoindicators, so which one is right? Have our species been spoiled by a long and blissful macroweather hiatus, or—on the contrary—did harshly varying climate

Berner et al. to conclude that the Holocene was, on the contrary, unstable. ...er, K. S., Koc, N., Divine, D., Godtliebsen, F., & Moros, M. A decadal-scale Holocene sea surface temperature record from the subpolar North Atlantic constructed using diatoms and statistics and its relation to other climate parameters. *Paleoceanography* **23**, doi:Doi:10.1029/2006pa001339 (2008).

adversity force us to invent new ways of coping? Perhaps the final answer is more mundane: that climate was much less important than other factors, such as rising population levels or technological progress.

Keeping to the climate side of the debate, there may be a simple resolution of this apparent paradox. What if the ice cap Holocene climate is not representative of global conditions? What if the transition from macroweather to the climate varies strongly from place to place? To answer this, we need global-scale, not only regional and high-latitude, temperature proxies. Although the exact timescale of the transition may be poorly discerned, we know that macroweather does indeed eventually give way to the climate. We may be certain of this because, at scales of 50,000 years (half a glacial–interglacial cycle), the temperature varies by ±2 to ±3°C (i.e., a total range of 4–6°C, somewhat less near the equator), so that the typical Holocene fluctuations in Figure 5.9 must increase rapidly at timescales only a little larger than those shown (the glacial–interglacial window of Fig. 2.4A and Fig. 5.14).
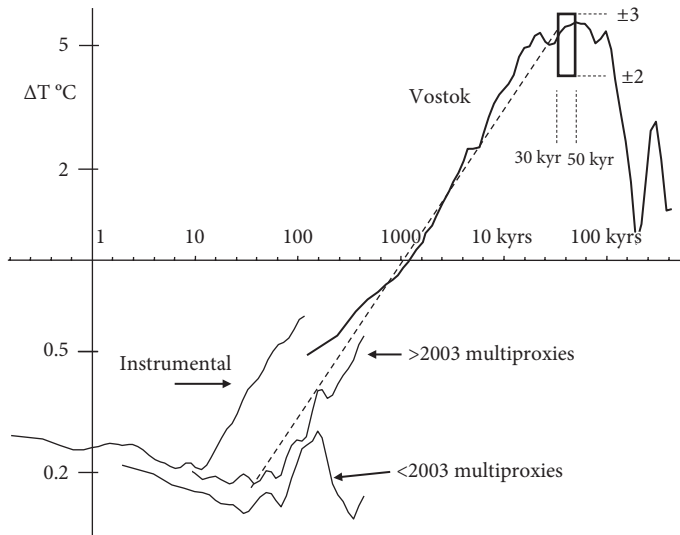


FIGURE 5.14 *A comparison of typical temperature fluctuations from three globally averaged instrument series (upper left, from 1880), from the five pre-2003 multiproxy series (upper curves in Fig. 5.15, covering the preindustrial period, 1500–1900), and three post-2003 multiproxies (the lower curves in Fig. 5.15). Also shown for reference is the Vostok core (proxy data from the past 420,000 years) and the glacial–interglacial window (the upper right rectangle). Finally, a reference line (slope, H = 0.5) is shown. Although the instrument series are too variable (as a result of anthropogenic warming), and the pre-2003 multiproxies seem to have overly weak variability at centuries and longer, the post-2003 multiproxies extrapolate nicely to the Vostok variability at longer timescales. This is as expected if the global Holocene (represented by the multiproxies) was the same statistically as the pre-Holocene Vostok proxies.*

In 2015, the question of Holocene stability and its regional variations were a main theme at a workshop I organized with colleague Anne de Vernal in Jouvence,

Quebec. The interest that it sparked in understanding and quantifying centennial and millennial Holocene variability led to its support by the international organization Past Global Change (PAGES). A series of three PAGES workshops is currently attempting to answer this question more completely by bringing together experts on paleodata with nonlinear geoscience specialists.

---

box 5.3  Hockey stick wars and multiproxy science

---

After the basic technique for producing multiproxies was known, they rapidly proliferated so that, by 2002, there were already five covering the northern hemisphere. Not surprisingly, the hockey stick was scrutinized—and vilified—intensely by climate skeptics (notably, statistician Steven McIntyre and engineer Ross McKitrick[71]), who spent considerable effort attacking it. Even today, skeptics refer ritualistically to multiproxies as "the discredited hockey stick."

To understand the issues, consider the original series[72] (which soon extended back to 1000 AD[73]), which illustrates both the technique and its attendant difficulties. The problems are in the longer timescales and the low frequencies. A basic issue was in getting long series that were both uniform temporally and representative spatially. For example, the original six-century-long multiproxy series presented in Mann et al.[74] had 112 indicators going back to 1820, 74 to 1700, 57-seven to 1600, and only 22 to the year 1400. Because only a few of the series went back more than two or three centuries, the series' "multicentennial" variability depended critically on how the loss of data at longer and longer time intervals was taken into account.

A somewhat different low-frequency issue concerns the EOF technique, which is critical for calibrating the proxies and rejecting noise. Theoretically, it only works properly if $H < 0$, yet global warming means that $H > 0$ at scales of twenty to thirty years and longer.[dd] This means that today's EOFs are effectively a little bit different from those of the preindustrial past, and this can potentially lead to errors in extrapolating current calibrations into the past. The politically charged issue at stake was whether the warming, such as the one we have enjoyed for the past century, was rare or common in the preindustrial record. In particular, skeptics seized on the fact that the hockey stick did not corroborate the "medieval warming event" and claimed that its nonappearance in the multiproxy reconstruction proved that the multiproxies were wrong. In actual fact—being a northern hemisphere, rather than European, average—Mann's hockey stick simply demoted medieval warming from an event of global significance to a regional one.[ee] More generally, the multiproxies allow us to address the question: Does the global temperature record exhibit GNFs at centennial scales? This is the subject of Chapter 6.

Awareness of low-frequency problems meant that increasing attention was paid to the treatment of the centennial and longer scales. One way to do this is to use borehole data, which—when combined with the use of the equation of heat diffusion—has

---

[dd] This point is still not fully appreciated.
[ee] It didn't eliminate a European warming, only a general hemispheric warming.

essentially no calibration issues whatsoever. Huang[75] used 696 boreholes, obtaining significant low-frequency variability (Fig. 5.15).[76] Similarly, to give proper weight to proxies with decadal and lower resolutions, especially lake and ocean sediments, Moburg et al.[77] used wavelets to calibrate separately the low- and high-frequency proxies. Once again, the result was a series with increased low-frequency variability. Ljundqvist[78] used a more up-to-date and more diverse collection of proxies to produce a decadal resolution series going back to 1 AD, also with stronger centennial and millennial variability. Finally, Marcott[79] created the first multiproxy for the entire Holocene, going back 11,000 years at twenty-year resolution.[80]
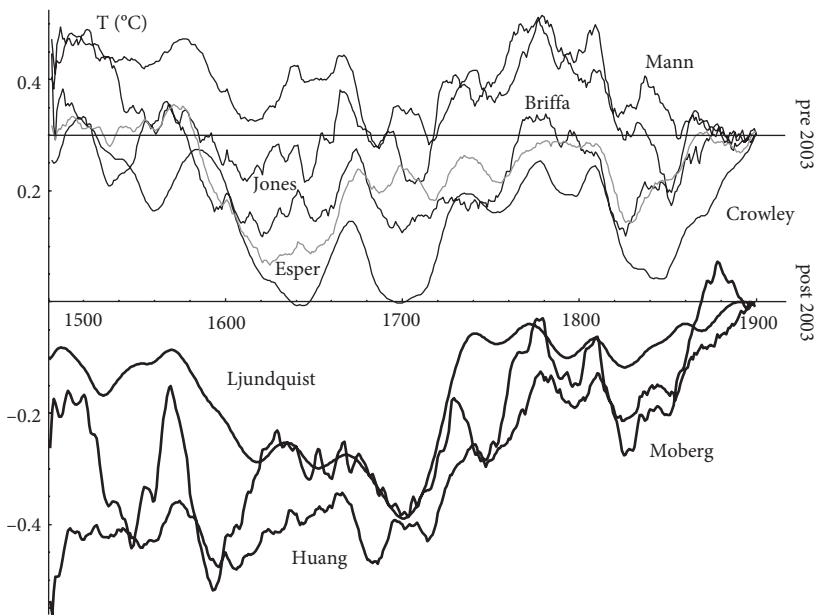


FIGURE 5.15 *Eight multiproxy reconstructions published since 1998, from 1500 to 1900, and smoothed to thirty-year resolution to bring out the low frequencies. This range was chosen to reflect the preindustrial fluctuations; pre-1900 anthropogenic warming is quite small. The multiproxies are divided into two groups: at the top, the pre-2003; at the bottom (shifted by 0.3°C for clarity) are four post-2003 multiproxies.[ff] The names correspond to the first authors of the articles in which the series were described. All the series have been shifted up and down to agree on the temperature in the year 1900 AD. At the left, we can clearly see the wide range of the inferred temperature change since 1500 AD, and the fact that, in general, the post-2003 multiproxies are more widely varying.[81]*

We can compare several of the reconstructions visually. Figure 5.15 shows that more recent reconstructions (those published after 2003; thick lines in the figure) generally have considerably larger overall temperature variations than the earlier

---

[ff] "Pre-" and "post-" 2003 refer to the publication dates of the multiproxies, not to the range of dates covered in the reconstructions.

(pre-2003) multiproxies (top). Quantifying this using fluctuations, we see in Figure 5.14 that, up to scales of 150 years or so, they all agree pretty well with each other, but for scales longer than a century or two, the more recent multiproxies are more variable at long times. There are two arguments in their favor: First, one of them is the borehole-based series and, because it is a direct physical measurement, it avoids all the usual proxy calibration issues. Second, Figure 5.14 also shows that the low-frequency variability seems to be more or less what is needed to explain the lower frequency Vostok paleodata—assuming, of course, that the global-scale (northern hemisphere) Holocene has the same type of variability as the previous epochs. These multiproxies are consistent with the hypothesis that, at global scales, the Holocene is statistically much like the previous periods. Also shown in Figure 5.14 is the industrial epoch instrument-based Haar fluctuations confirming they are too strong to be compatible with the preindustrial multiproxies, so industrial epoch warming is indeed unprecedented (see Chapter 6).

## 5.7 The multiproxy revolution

The key discovery that transformed climate science—and that can potentially answer the question about the Holocene stability—was the development in 1998 by Mann, Bradley, and Hughes of "multiproxy reconstructions."[82] Up until then, numerous individual climate series from tree rings (dendrochronology), varves (lake sediments), palynology (pollen, dust), ice cores, magnesium/carbon variation in shells, [18]O in foraminifera, diatoms, speleothems (stalagmites), biota, and other sources had began to proliferate, leading to numerous local reconstructions of past climates. Not surprisingly, they were all highly variable in time and in space, and their chronologies and calibrations in terms of temperatures were poorly discerned. The wide sweep of global climate change was only qualitative and it was quite Eurocentric, being heavily based on historical evidence[gg] such as records of vines growing in Britain during the Middle Ages, illustrating the "medieval warming event," or Pieter Breugel's sixteenth-century renditions of skaters on frozen Dutch canals, illustrating the "little Ice Age."

Although each individual proxy series has climate information,[hh] up until the advent of multiproxies, they were noisy and the calibration required knowledge of the temperatures at the proxy sites over decades and longer, and these data were rarely available. In addition, although neighboring proxies could potentially back up each other to reject noise and to detect systematic regional changes better, there was no proper way of combining them statistically. The multiproxy breakthrough

---

[gg] At the time, the classic reference was Horace Lamb's monumental "Past Present and Future." Lamb, H. H. *Climate: Past, Present, and Future*. Vol. 1, *Fundamentals and Climate Now*. (Methuen and Co., 1972).

[hh] For simplicity, we restrict our attention to temperature proxies, which are the most reliable; but others—notably, precipitation—do exist.

was to use a technique called "empirical orthogonal functions" (EOFs), which allowed the instrument record to determine the main components (EOFs) of the temperature variability both to reject noise and to determine the temperature at locations where there were no direct measurements: to fill in data "holes."

By using hundreds of proxy series, including (annual resolution) dendrochronologies, Mann et al. thus developed the first quantitative reconstruction of global-scale (northern hemisphere) temperatures—at first, since 1500 AD,[83] but they soon extended it back to the year 1000 AD.[84] The result was the instantly acclaimed "hockey stick"—a version of which is shown in Figure 5.16, a graphic showing a gradual decline of northern hemisphere temperatures since 1000 AD, followed by a (relatively) rapid warming since the end of the 1800s. The period of long gradual decline is the stick handle; the recent rapid warming is the business end. The hockey stick[ii] was the first quantification of industrial epoch warming, and it underlined its rapidity visually. This led to the famous conclusion—showcased in the 2001 IPPC AR3[85]—that the twentieth century was the warmest century of the millennium, that the 1990s was the warmest decade, and that 1998 was the warmest year. It foisted its feisty lead author, Michael Mann, onto center stage in the developing "Climate Wars," which are discussed in more detail in Chapter 6.
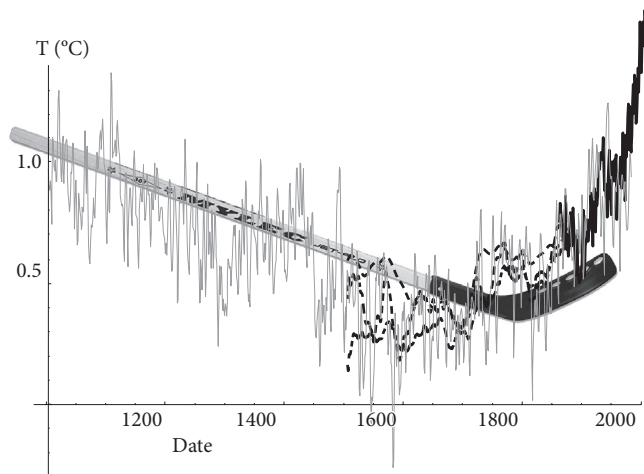


FIGURE 5.16 *An updated version of Mann's hockey stick using the three post-2003 multiproxies in Figure 5.13 (thin solid lines, with a thirty-year smoothing, from 1500–1900 only), and the annual resolution, globally averaged temperatures since 1880 to 2013 (thick, solid, the same National Aeronautics and Space Administration GISS series as in Fig. 6.4). The dashed line is the Moburg multiproxy at annual resolution from 1000 AD [taken from Fig. 1.5D (bottom); it is the same as the one with the thirty-year smoothed data from 1500 to 1900 that was used in Fig. 5.15]. The legendary hockey stick has been superposed, showing slowly cooling preindustrial temperatures followed by a sharp increase during the industrial period.*

---

[ii] At first, this was a climate skeptic term of derision, but it was soon adopted by Mann himself. See his excellent work, Mann, M. E. *The Hockey Stick and the Climate Wars: Dispatches from the Front Lines.* (Columbia University Press, 2012).

## 5.8  Ice Ages and the macro- and megaclimate regimes

Although there is still much uncertainty about the time scales of the transitions between macroweather and climate, we have seen (Figs. 2.4A and 5.14) that if one averages over long ice or ocean cores (i.e., many epochs, glacial cycles), one finds a transition at ~~about one~~ hundred years, and scaling behavior out to 50,000 years or so. But what happens in the vicinity of 100,000 years, and at the really long timescales from millions of years to the age of the planet? Although we have just seen that the external or internal origin of the centennial and multimillennial scaling is still not clear, thanks to astronomical forcing theory (Milhutin Milanković, 1879–1958), climate signals from scales of tens of thousands to hundreds of thousands of years are apparently the result of classic scalebound causes: the precession of the equinoxes (23,000 years), changes to the obliquity of Earth's orbit (41,000 years), and changes in the eccentricity (ellipticity) of Earth's orbit (95,000 years, 125,000 years, and 413,000 years). Box 5.4 gives more detail on this.

---

BOX 5.4  Ice Ages and orbital forcing

I have deliberately focused on the broad sweep of scales from millennia on up, showing how the data—such as they are—broadly support the existence of a megaclimate regime beyond 1,000,000 years and a fairly narrow macroclimate regime with a character that is unclear. In this sweep, I have deliberately underplayed astronomical forcings and responses. The main reason for this choice is that astrophysics provides textbook examples of extreme scalebound—with an even "clockwork"—regularity and, as a consequence, there is a large body of literature on the subject of the corresponding "Milanković" climate forcings.

Although there is no doubt that signatures of Milanković forcings can be found in the climate record, the subject is far from settled. To start with, over the range of timescales between 20,000 years and 400,000 years, there are seven or eight astronomical forcing frequencies, and they can potentially interact with each other. Consider the precession of the equinoxes, the fact that the axis of Earth's rotation slowly changes direction so that the dates of the equinoxes shift. There are actually two precessions. The first is the "axial" precession, with a period of 26,000 years. This is the time that it takes for the axis of Earth to rotate once about a direction perpendicular to the ecliptic (the orbital plane) with respect to the fixed stars.[jj] It is caused by the solar and lunar tidal forces exerted on the solid Earth. The second, the "apsidal" precession, refers to the precession of the (elliptical) orbit of Earth about the sun, with a period of 112,000 years relative to the fixed stars. The apsidal and axial precessions combine to vary the position in the year that Earth reaches perihelion (maximum distance from the sun). On average, apsidal precession shortens this period to 23,000 years (varying between 20,800 years and 29,000 years).

---

[jj] Hence, for example, Polaris would no longer be the north pole star.

Because Earth's orbit around the sun is slightly elliptical,[86] the precession of the equinoxes implies that the northern hemisphere summer is sometimes at the point of closest approach to the sun whereas, half a cycle later, the closest point would occur in the northern hemisphere winter. Because the opposite relationships hold for the southern hemisphere, a north–south symmetrical planet would display no overall increase in incoming solar radiation/solar forcing. Although Earth is not north–south symmetrical, having most of its landmass in the northern hemisphere, the precession of the equinoxes has a small effect and, not surprisingly, any paleotemperature response is undetectable.

However, at 41,000 years, Earth's axis itself (the obliquity) varies in direction periodically, from a minimum of 22.1° to a maximum[87] of 24.5°, with respect to the normal to the orbital plane (it is currently at 23.5°). Because the tilt is responsible for the seasons, greater tilts mean more extreme seasons, effectively varying the amount of summer insolation in northern high latitudes, where there is a large landmass that can accumulate ice. This periodicity was the one that Milanković originally singled out as dominant, and it turns out that he was right—at least for times between 2,500,000 years ago and 800,000 years ago! Indeed, in Figure 1.5A, it can be seen as a fairly uniform oscillation, before disappearing below the background around 800,000 years ago (spectral analysis shows it is still there, but that in the recent epoch it has been much, much weaker than the 100,000-year cycle). The replacement of the then-dominant 41,000-year cycle with a dominant 100,000-year cycle is called the "100,000 year problem."

Moving to longer astronomical timescales, which are the result of changes in the orbit of Earth around the sun, there are "wobbles" caused largely by interactions with Jupiter. These cause the eccentricity[kk] of Earth's orbit to change. The main periodicities in the wobble are at 95,000 years, 125,000 years, and 405,000 years, yet it is the weakest at 95,000 years and apparently has been dominant over the others during the past 800,000 years.[88] This is sometimes called the "unsplit peak problem." Indeed, spectral analysis of a new EPICA (Antarctic core) dust data set, with a remarkable 25-year temporal resolution (Fig. 1.5C), shows that, over the range of frequencies corresponding to 50 years to 800,000 years, 15% of all the variability can be attributed to this single 95,000-year spectral spike, with only a percent or two attributable to all the other astronomical forcing frequencies. The rest—the majority—are associated with the scaling background.[89]

Naturally, these findings have spawned numerous models, especially of the deterministic chaos variety, but the problem is still unsolved.[90]

Milanković proposed his theory just before the first world war, but it took more than fifty years before it was widely recognized, and sixty years before early paleodata gave it convincing support, first by the Climate: Long-Range

---

[kk] The eccentricity is a measure of how noncircular the orbit is, with a value of 0 corresponding to a perfect circle. The eccentricity of Earth's orbit varies from a nearly circular value of 0.000055 to a mildly elliptical maximum value of 0.0679. Currently, it is 0.017 and decreasing.

Investigation, Mapping and Production project in 1972 and then by an influential article by Hayes, Imbrie, and Nicholas Shackleton (1937–2006) titled "Variations in the Earth's Orbit: Pacemaker of the Ice Ages."[91]

By the time (1976) that Mitchell proposed his wide-scale-range spectral composite of climate variability (Fig. 2.3A), in accordance with Milanković the Ice Ages were already represented as a cluster of scalebound spectral "bumps" superposed on his otherwise flat (white noise) background. In 1990, on the basis of early ocean cores, Shackleton and Imbrie[92] made an important advance, noting that the spectrum of benthic $\delta^{18}O$ was "approximately" scaling from about 1,000 years to 100,000,000 years, so that the Ice Ages were instead a scalebound spectral bump on an otherwise *scaling* background.[ll] Looking at the paleodata composite spectrum in Figure 2.3A, we see that this interpretation is tempting. It is easy to imagine a single line through the broad spectral bump at around 100,000-year scales. If this interpretation was correct, it would imply—interrupted only by a short astronomically forced bump—that a single scaling climate regime would be valid from centuries to hundreds of millions of years.

Although this "scaling with a bump" picture is certainly a much better characterization of the variability than Mitchell's "white noise with a bump," it is worth revisiting this scale range with the benefit of both improved data and the use of fluctuation analysis. Figure 5.17 shows the result, focusing on the longer series presented in Figures 1.5A and 2.4A, B (those covering the timescales from 1,000 to 550,000,000 years). With the exception of a single ice core (EPICA, dashed line), the data used in the figure are "benthic"; they come from $\delta^{18}O$ analyses of $CaCO_3$ exoskeletons of ancient planktonic creatures taken from ocean sediment cores. In the benthic series (see Fig. 1.5A), $\delta^{18}O$ varies both as a result of temperature as well as of the mean $\delta^{18}O$ of the water in which the foraminifera lived. The basic relationship between $\delta^{18}O$ and temperature is an inverse one: Increasing $\delta^{18}O$ is associated with decreasing temperatures, although this relationship is complicated by the change in the $\delta^{18}O$ composition of the seawater resulting from the preferential sequestering of light seawater in ice caps (see Box 1.2). There are thus both direct and indirect links to the temperature via the ice sheets.[93] In addition, temperature variations are latitude dependent, so that high-latitude temperature variations are amplified by roughly 50% above global variations.[94] In Figure 5.17, I've given a simplified presentation.[95] A full discussion is provided elsewhere.[96]

---

[ll] A few years earlier, I had focused on scaling up to Ice Age scales. The single benthic series in the scaling composite (see Fig. 2.18) was used only "to get a feel for the temperature variations over the very long timescales" and, retrospectively, it was spuriously flat because it used difference fluctuations: Lovejoy, S. & Schertzer, D. Scale invariance in climatological temperatures and the local spectral plateau. *Annal. Geophys.* **4B**, 401–410 (1986).
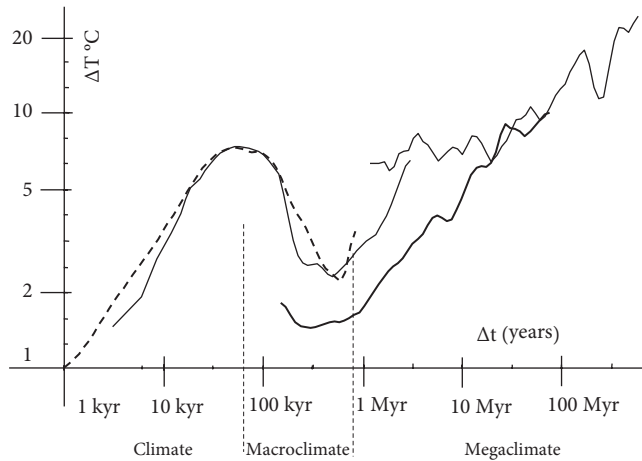
FIGURE 5.17 *A composite focusing on the macroclimate and megaclimate regions using part of the Antarctica ice core analyses (only the part Δt > 1,000 years is shown; EPICA, dashed line, left) with the benthic series (thick line, Zachos; thin line left, Huybers). The Veizer series is the thin line at the upper right. The proxies all agree quite well for Δt below about 200,000 years and for Δt greater than about 20,000,000 years.*[97]

At the shorter scales, from roughly 1,000 to 100,000 years, we see that both the ice cores and the benthic data agree very well with each other. They both show quite accurate scaling, with $H \approx 0.4$, and give significant support for the robustness of our conclusions about the overall scaling in the climate regime. Similarly, at very long timescales, beyond 10,000,000 years, two different benthic "stacks" (in Fig. 5.16, the thick line using more than 14,000 measurements and the thin line, upper right, with nearly 3,000 points) give agreement on a fairly linear (scaling) regime at very long times, right up to the "Cambrian revolution," when the first hard-shelled creatures appeared—the limit of this method. Although the behavior in the range 100,000 to 10,000,000 years is not so clear, the analysis does show convincingly that, despite the near (probably fortuitous) agreement of the fluctuation exponents (*H* values) for the climate and the long-time "megaclimate" regime,[mm] the latter is *not* simply a continuation of the former with a "Milanković bump" as in the early Shackleton and Imbrie picture. The climate and megaclimate regimes are clearly distinct so that, for example, the fluctuations at 100,000 years are of about the same magnitude (8°C; i.e., ±4°C) as those at 20,000,000 years. Physically, megaclimate variability is presumably mostly the result of the rearrangement of landmasses by plate tectonics. An example of this is the current Quaternary period (the past 2,500,000 years), during which glacial–interglacial cycles occur. Before that, the distribution of land in the polar regions couldn't

---

[mm] The new lower frequency regime was dubbed "megaclimate" because "mega" means "large" and the term evokes the megafauna of these ancient climes.

support the buildup of significant ice sheets and there were apparently no such cycles.[nn]

Viewed in this way, perhaps the most significant feature of the scale range between 100,000 years and 1,000,000 years is the significant *drop* in the variability by a factor of about 5 over this fairly narrow range of scales. In Figure 5.17, this appears as a kind of "macroclimate gap." Focusing on the dashed line (ice core analysis), the decline in variability actually seems to be quite linear—scaling—with an exponent $H = -0.8$. Whether this scale range is best characterized as a "gap" or as a short "macroclimate" scaling regime is thus a matter for debate.

## 5.9  The death of Gaia

There is no question that the evolution of life on Earth has transformed the atmosphere, and that life continues to influence the climate. This is the starting point of James Lovelock's (b. 1920) Gaia hypothesis that Earth itself is a giant "superorganism." For this to be more than just a metaphor, Lovelock claims that, just like a living organism, the entire planet displays homeostasis. Homeostasis denotes the fundamental self-regulating attribute of living beings—their ability to maintain their internal conditions close to a point optimal for their development. If the Gaia hypothesis is true, it would mean that, thanks to negative feedbacks, Earth maintains conditions (including temperatures) in a range fit for life (i.e., itself). A major Gaian example is the apparent stability of the concentration of atmospheric oxygen. Over the past hundreds of millions of years, it seems to have remained within the bounds of 15% to 35%. Although data are limited, it is plausible that it is indeed stable, with canceling fluctuations favoring a long-term mean value not far from the current concentration (21%). If true, in our language, the megaclimate $H$ value for oxygen concentration would be negative.

However, Lovelock has had more difficulty in justifying homeostasis for Earth's temperature, producing essentially speculative mechanisms rather than hard evidence. For example, in 1983, Lovelock and colleague Andrew Watson pioneered the use of low-dimensional deterministic chaos models of the climate: "Daisyworld."[98] Daisyworld is a model of a planet covered with two types of daisies: black and white. White daisies have a high albedo but they like warmer temperatures whereas black daisies have low albedos but they prefer cooler temperatures. The resulting Daisyworld model exhibits chaotic temperature variations around a well-defined, long-term fixed temperature. Daisyworld demonstrates that life can *potentially* regulate temperature in a nontrivial (chaotic, randomlike) manner. Yet

---

[nn] Tens of millions of years earlier still, there were likely to have been other epochs with distributions of continents favorable to the operation of glacial cycles. In the much more distant past before the Cambrian about 600 million years ago, there was even believed to be a state called "snowball Earth" in which virtually the entire planet was covered with ice.

one is struck by the artificiality—and given the assumptions—the predictability of the Daisyworld result.

In an attempt to provide more realism, Lovelock later developed another mechanism, postulated to have operated over megaclimate scales, baptized the "CLAW hypothesis,"[99] after the authors' initials.[oo] The CLAW hypothesis is a negative temperature feedback mechanism based on planktonic production of dimethyl sulfide (DMS). It is hypothesized that increasing temperatures (a positive fluctuation) leads to increasing planktonic production of DMS. DMS acts as cloud condensation nuclei that tend to increase cloud cover; increasing cloud cover tends to decrease the temperature (a negative fluctuation). Conversely, negative temperature fluctuations lead to (at least partially canceling) positive fluctuations later on, hence there is overall stability— homeostasis.

But whatever its exact mechanism, by its very definition, negative feedbacks imply that successive temperature fluctuations tend to cancel each other out so that $H < 0$. However, the proxy evidence from the megaclimate regime (Fig. 1.5A, top series, Fig 5.14) is that, on the contrary, $H > 0$, so that the temperature changes are actually more like a drunkard's walk, tending to grow rather than diminish with timescale. There is no evidence that the temperature tends to return to a well-defined value. Even its visual appearance in Figure 1.5A shows a "wandering" rather than a canceling character. Quantitative fluctuation analysis (Fig. 5.17) confirms that, at least during the past 550,000,000 years, $H > 0$. Because temperature is obviously fundamental for life, this is strong evidence against the Gaia hypothesis.

This criticism based on the analysis of megaclimate temperature proxies adds to the already widespread scientific skepticism toward Gaia (see Tyrrell,[100] for example, for a critical survey). Unfortunately, inasmuch as the hypothesis has attained nearly religious status, with Gaia replacing the sun god, these analyses are unlikely to influence its cult followers.[pp]

Beyond its overall quasi-periodicity, we may ask whether these astronomical cycles might help us understand our current interglacial epoch, the Holocene (Section 5.6). Because this is the epoch of agriculture and civilization, for humanity, the key issues are the duration of warm, largely ice-free Holocene conditions and the stability of such conditions. If we compare the Holocene length—already about 11,700 years—to the lengths of earlier interglacials, we find that the last time that an interglacial was longer than about 3,200 years was the 24,000-year-long interglacial that ended 395,000 years ago. The 400,000-year timing between these long interglacials is at least consistent with the astronomical forcing theory. A possible link is provided by the 405,000-year eccentricity cycle

---

[oo] Robert Jay Charlson, James Lovelock, Meinrat Andreae, and Stephen G. Warren. More recently, Lovelock proposed the "anti-CLAW" hypothesis—a positive feedback mechanism that should apply under conditions of anthropogenic warming. In such scenarios, the mechanism effectively amplifies anthropogenic warming. Lovelock, J. E. *The Revenge of Gaia*. (Penguin, 2007).

[pp] Lovelock seems quite happy to cultivate some of the more esoteric interpretations.

that modulates the precession of Earth's axis.[101,qq] Yet when we consider the entire set of eight interglacials, we find that the even earlier interglacials were all around 10,000 years long, and the average over the eight is 9,900 years, with a typical spread of ±7,500 years. From this perspective, the Holocene does not seem so long.

Although not astronomical, I could mention another theory that has been put forward to explain a long Holocene: the Early Anthropocene Hypothesis (EAH) by William Ruddiman.[102] The EAH posits that the current interglacial has been prolonged artificially by the development of agriculture. According to the EAH, starting around 5,000 years ago, early farmers had sufficiently increased atmospheric $CO_2$ and $CH_4$ concentrations that the resulting climate forcing artificially prolonged the Holocene that presumably would otherwise have already reverted to a glacial state. The main criticism of this theory is that the forcings were much too small. Today, natural processes are invoked more commonly to explain the weakly anomalous $CO_2$ and $CH_4$ levels. Section 5.6 addresses the issue of Holocene stability about which astronomical forcing plays little role.[rr]

## Notes

1. Stability is usually defined with respect to a specific theory or model. An example is convective stability. This is a property of a smooth atmospheric layer when it is heated from below. When the temperature gradient exceeds a critical threshold, tiny perturbations are amplified, rapidly destroying the initial smooth state. The scaling $H$-based definition of stability is instead a statistical, phenomenological definition that characterizes the real, observed propensity of fluctuations to amplify or diminish as functions of scale.

2. Figure 1.6 was normalized by the standard deviation, rather than the absolute difference, but this is a minor distinction.

3. This definition of climate state (as a "climate normal") is directly accessible by empirical measurement. It is somewhat different from the climate notion used in GCMs (as discussed later), which can be determined only from models, by running them for a very long time with completely fixed external conditions (in "control runs").

4. Beyond simply defining the macroweather anomaly as the difference of the current macroweather state and the climate state, there is a slight additional complication. For

---

qq Earth's precession affects the amount of energy that it absorbs because sea ice can grow more readily in the southern hemisphere, and this reflects more energy away from Earth. However, precession matters only when eccentricity is large (the orbit is more elongated), hence the potential influence of the 400,000-year cycle. These effects, however, are quite small so that the astronomical theory does not (yet) explain our long interglacial.

rr The anomalously high $CO_2$ and $CH_4$ concentrations invoked by the EAH are on the order of 7% and 15%, respectively, over 5,000 years compared to a 50% increase of these concentrations since the industrial revolution. For natural explanations of these slow increases, see Studer, A. S., Sigman, D. M., Martínez-García, A., Thöle, L. M., Michel, E., Jaccard, S. L., Lippold, J. A., Mazaud, A., Wang, X. T., Robinson, L. F., Adkins, J. F., & Haug, G. H. Increased nutrient supply to the Southern Ocean during the Holocene and its implications for the pre-industrial atmospheric $CO_2$ rise. *Nat. Geo. Sci.* **11** (10), 756–760 (2018).

monthly resolution anomalies, one must remove the annual cycle. For example, the officially ordained procedure for calculating the January anomalies used in Plate 5.1 (right column) includes the differences between the January monthly averaged temperature and the average of all the Januaries over the previous thirty-year reference period.

5. As far as I know, such weather anomalies are not used, probably because the evolution of the atmosphere depends on the actual state of the atmosphere and not on its difference with respect to the current monthly (macroweather) average. Although these weather anomalies are useful for highlighting the day-to-day evolution of the atmospheric state, they do not help to predict it.

6. This statement is true of regions several hundred kilometers across, not of individual station anomalies, which can vary much more strongly.

7. Since $H < 0$, the thirty-year (360-month) average fluctuations are much smaller than the monthly ones. Using the typical value $H = -0.2$, we find that averaging reduces the amplitudes of the fluctuations by a factor of $360^H \approx 0.30$ (see Plate 5.3). (Although, in the recent period, as a result of global warming, this variability corresponds to fluctuations around a slightly higher average value.) Because of this amplitude reduction, most of the anomaly variability is a result of the variations at monthly scales.

8. Lovejoy, S. What is climate? *EOS* **94** (1), 1–2 (2013).

9. Estimated directly from the 1871 to 2010 20CR data at 2° resolution.

10. Adapted from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

11. From 1750 to 1930, the total warming was only about 0.3°C. Today, it is about 1°C (Fig 6.4A).

12. Although the global-scale year-to-year natural variability is close to 0.1°C, over long-enough timescales it is double that value. So anthropogenic warming must exceed about 0.2°C to be detected. This was the total anthropogenic warming over 1880 to 1940 (i.e., sixty years).

13. The upper right of Figure 5.2 uses paleodata.

14. To within statistical error. This range of $H$ values and the macroweather predictability properties discussed here and later are thus only strictly valid for low-intermittency processes. They are thus approximately valid in macroweather.

15. Because of the "central limit theorem" in statistics, it doesn't make too much difference how one determines the distribution of the random numbers. The key point is that the way in which successive numbers are chosen should be identical and they should each be chosen independently.

16. It turns out that spatial averages over large regions leads to the subregions, with $H$ closest to zero dominating the others, so that global averages—dominated by oceans—have $H \approx -0.1$ (see Fig. 5.1e and Box 5.1).

17. This is not surprising because we saw (Fig. 4.12) that the ocean weather–ocean macroweather transition occurs at about one year, which is the lifetime of planetary scale "gyres," so that the SST is within the usual deterministic predictability limit. The air macroweather temperatures over the ocean are thus much more predictable than the corresponding temperatures over land.

18. Reproduced with permission from Del Rio Amador, *The Stochastic Seasonal to Interannual Prediction System*, PhD dissertation, McGill University (2018).

19. This statement is exactly true for the simulation. For the data, it is only approximately true, because it is based on an estimate of $H$. The bottom series, right column, has $H$ closer to −0.45; it is not completely uncorrelated.

20. Corresponding to 140 years, from 1871 to 2010.

21. For the series at the bottom, with $H = -1/2$, the amplitude is reduced by the factor $12^{1/2} = 3.46$. The factors ($12^{-H}$) are 1.28, 1.64, 2.11, 2.70, and 3.46, from top to bottom.

22. The model and data resolutions may have a mismatch in both space and in time. The *effective* spatial resolution of GCMs is not so obvious because of the use of artificial "hyperviscous" dampening/smoothing needed to keep the model stable numerically. It is typically a factor of 2 to 3 less than the nominal resolution (raw grid scale), and the raw GCM outputs are corrected by these ad hoc factors.

23. This now goes under the sophisticated term "quantile matching." Quantile matching is a general method of forcing the probability distribution of the data and simulations to match. For Gaussian processes (a good approximation here), it simply consists of the multiplication of the anomalies by the numerical factor needed to match the amplitudes of the simulations and data. See, for example, Zhao, T., Bennett, J. C., Wang, Q. J., Schepen, A., Wood, A. W., Robertson, D. E., & Ramos, M. How suitable is quantile mapping for postprocessing GCM precipitation forecasts? *J. Clim.* **30**, 3185–3196 (2017).

24. Lovejoy, S. How accurately do we know the temperature of the surface of the earth? *Climate Dynam.* doi:doi:10.1007/s00382-017-3561-9 (2017).

25. Lovejoy, S. How accurately do we know.

26. Lovejoy, S. How accurately do we know.

27. Each of the six data sets has its particular strengths and weaknesses. Those from NOAA and NASA used essentially the same land and marine data, but use different methods to fill (some) of the data holes. In contrast, the Hadley model, version 4 (HAD4), series (from the Climate Research Unit) made no attempt in this direction, thus making fewer assumptions about the spatial statistical properties (especially smoothness, regularity properties). Another series (by Kevin Cowtan and Robert Way) takes the contrary view. It uses the HAD4 data but makes strong spatial statistical assumptions (Kriging) to fill in data holes. This is especially significant in the data-poor high-latitude regions. The 20CR series is of particular interest because it uses no temperature station data whatsoever. Instead, it uses surface pressure station data and monthly SST data (the same as HAD4) combined with a numerical model (a reanalysis). It is the only series that gives actual temperatures rather than changes with respect to a reference period (the "anomalies"). The fact that the 20CR agrees well with the other (station-based temperature) estimates is strong support for all the series [Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruger, A. C., Kruk, M., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., & Worley, S.J. Independent confirmation of global land warming without the use of station temperatures. *Geophys. Res. Lett.* **40**, 3170–3174 (2013)]. A final series [Berkley—shorthand for the series produced by this group: Rohde, R., Muller, R. A., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., Wurtele, J., Groom, D., & Wickham, C. A new estimate of the average Earth surface land tesperature Spanning 1753 to 2011. *Geoinfo. Geostat.: An Overview* **1** (1), 1–7 (2013).] uses the same SST data as both HAD4 and Cowtan and Way, but it uses data from many more stations (≈37,000 compared to only 4,500 for HAD4 and 7,300 for the NOAA series, for example), and it uses a number of statistical improvements in the handling of data homogenization and coverage. This diversity is significant so that the degree of agreement or disagreement between the various series is itself important.

28. This was their common period.

29. The analysis in Figure 5.14 was performed a little earlier and involved only three series.

30. Adapted from Figure 2 in Lovejoy, S. How accurately do we know the temperature of the surface of the earth? *Climate Dynam.* doi:doi:10.1007/s00382-017-3561-9 (2017).

31. Lovejoy, S. & de Lima, M. I. P. The joint space–time statistics of macroweather precipitation, space–time statistical factorization and macroweather models. *Chaos* **25**, 075410, doi:10.1063/1.4927223 (2015).

32. Technically, Figure 5.5B shows a "complex cascade"—one involving the multiplication of complex numbers. Complex cascades are similar to the usual (real-value) cascades discussed in Box 2.2, except that they are based on numbers with both real and imaginary parts. The model in the figure is an extension of models presented in Schertzer, D. & Lovejoy, S. From scalar cascades to Lie cascades: Joint multifractal analysis of rain and cloud processes In *Space/Time Variability and Interdependence for Various Hydrological Processes* (ed. R. A. Feddes), pp. 153–173. (Cambridge University Press, 1995).

33. For example, data from individual climate stations are made commensurate ("homogenized") by using the station standard deviations or probability distributions to adjust them.

34. These statements are true of the precipitation anomalies (i.e., after the annual cycle and long-term average have been removed).

35. https://en.wikipedia.org/wiki/Weather

36. Unfortunately, as we saw in Chapter 4, the fact that it was graphically represented as a *maximum* in the spectrum and not simply a transition of *spectral type* was purely an artifact of the way the graphs were plotted. This allowed scientists to think in scalebound terms. It was relegated to the status of a disturbance over a narrow range of timescales rather than representing a fundamental change in character.

37. In the case of the Fractionally Integrated Flux (stochastic cascade) model, various extensions have been made, such as to couple them with an ocean model as well as to include climate zones, but these models have not yet been much developed: Lovejoy, S. & de Lima, M. I. P. The joint space–time statistics of macroweather precipitation, space–time statistical factorization and macroweather models. *Chaos* **25**, 075410, doi:10.1063/1.4927223 (2015). And Appendix 10D in Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

38. Lorenz, E. N. The predictability of a flow which possesses many scales of motion. *Tellus* **21**, 289–307 (1969).

39. In Lorenz's error doubling model, structures were predictable over their lifetimes. It is ironic that the connection was not made that the ten-day predictability limit implied a ten-day lifetime of planetary structures. The failure to make the connection may be a consequence of the academic two-dimensional/three-dimensional paradigm discussed in Chapter 4. This is because a unique feature of two-dimensional turbulence is its predictability limit, which is much longer than the lifetime of the large structures.

40. Figure 5.8 uses Haar fluctuations, and because $H < 0$, these fluctuations are essentially averages of the temperature anomalies with respect to the model's long-term climate.

41. This means that rerunning the same GCM for 300 simulated years with only tiny differences in the starting conditions would lead to a 300-year average of about $2 \times 0.06 = 0.12°C$ different from the first.

42. This is the average over the eleven models used in Figure 5.5, but there is some variation from model to model. We return to this in Chapter 7, where we forecast these control runs using emergent scaling laws.

43. Lovejoy, S., Schertzer, D., & Varon, D. Do GCM's predict the climate . . . or macroweather? *Earth Syst. Dynam.* **4**, 1–16 (2013).

44. The convergence can be sped up by combining statistical with temporal averaging on an ensemble of model runs, each starting with slightly different initial conditions. This is used to project the climate to the year 2100 (see Chapter 7).

45. These were selected for their long length and for the absence of overactive El Niño dynamics (three- to five-year variability).

46. More sophisticated models can reproduce different *H* values. See appendix 10D in Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. (Cambridge University Press, 2013).

47. Berner, K. S., Koc, N., Divine, D., Godtliebsen, F., & Moros, M. A decadal-scale Holocene sea surface temperature record from the subpolar North Atlantic constructed using diatoms and statistics and its relation to other climate parameters. *Paleoceanography* **23**, doi:Doi:10.1029/2006pa001339 (2008).

48. Adapted from Lovejoy, S. & Schertzer, D. Low frequency weather and the emergence of the climate. In: *Extreme Events and Natural Hazards: The Complexity Perspective* (eds. A. S. Sharma, A. Bunde, D. N. Baker, & V. P. Dimri), pp. 231–254 (2012).

49. Bryson, R. A. The paradigm of climatology: An essay. *Bull. Amer. Meteorol. Soc.* **78**, 450–456 (1997); quote, p. 451.

50. Pielke has criticized this on the grounds that many of the boundaries, such as atmosphere–land, are not just passive, but involve exchanges of energy and other fluxes. Pielke, R. Climate prediction as an initial value problem. *Bull. Amer. Meteorol. Soc.* **79**, 2743–2746 (1998).

51. In dynamical systems theory, GCM control runs are mathematically autonomous systems and the control-run climate is a state-averaged over the system's "strange attractor." The analogous mathematical concept for (nonautonomous) systems that have time-varying forcing is the "pullback attractor." Both concepts are defined in abstract high-dimensional phase space and are not empirically accessible. For a review, see, for example, Dijkstra, H. *Nonlinear Climate Dynamics*. (Cambridge University Press, 2013).

52. Volcanic and solar variability are themselves highly variable over decadal, centennial, and millennial scales. They are also scaling (see Box 5.2)! In any case, it is the way that their variability increases or decreases as a function of timescale that is important (depending on the signs of their own *H* values). As a result of these continual perturbations, the control run-defined climate may never even be approached, so that its relevance to the real world is questionable. For example, Box 5.2 shows that at least volcanic forcings decrease in amplitude at longer and longer times, so although they might not prevent convergence to a climate (i.e., the *H* value of the volcanically forced system might still be less than zero), they would likely yield a model climate different from the real one.

53. For this reason, it is often replaced by ~~the slightly more realistic characterization,~~ the transient climate response (TCR), which replaces the ECS "step function" (instantaneous) increase in the forcing by a seventy-year linear "ramp" doubling of the forcing. Because the TCR is defined by the change of the global temperature only over the ramp period, it does not take into account the longer term changes. It specifically avoids trying to define the new climate.

54. The GISS simulations are all from version E2-R of the model.

55. By astronomical standards, 1,000 years is a short time, so that the orbital parameters of Earth have changed little. They are taken as fixed.

56. Both views adapted from Lovejoy, S., Schertzer, D., & Varon, D. Do GCM's predict the climate . . . or macroweather? *Earth Syst. Dynam.* **4**, 1–16 (2013).

57. Orbital forcings (Box 5.3) were ignored because they are only important at scales on the order of 10,000 years and longer.

58. Since 2004, new detectors on the Total Irradiance Monitor Satellite have solved the problem, even finding a large (5 W/m$^2$) absolute calibration error in the previous estimates. Fortunately, the solar *anomalies* (defined as long-term *changes* with respect to a reference value) were not affected.

59. Reproduced from Lovejoy, S. & Varotsos, C. Scaling regimes and linear/nonlinear responses of Last Millennium climate to volcanic and solar forcings. *Earth Syst. Dynam.* **7**, 1–18 (2016). The data were from Wang, Y.- M., Lean, J. L., & Sheeley, N. R. J. Modeling the Sun's magnetic field and irradiance since 1713. *Astrophys J.* **625**, 522–538 (2005).

60. The figure was reproduced from Lovejoy, S. & Varotsos, C. Scaling regimes and linear/nonlinear responses of Last Millennium climate to volcanic and solar forcings. *Earth Syst. Dynam.* **7**, 1–18 (2016). As deduced from a record of $^{10}$Be plotted from data in Steinhilber, F., Beer, J., & Frohlich, C. Total solar irradiance during the Holocene. *Geophys. Res. Lett.* **36**, L19704 (2009).

61. The figure was reproduced from Lovejoy, S. & Varotsos, C. Scaling regimes and linear/nonlinear responses of Last Millennium climate to volcanic and solar forcings. *Earth Syst. Dynam.* **7**, 1–18 (2016). The data were from Gao, C.G., Robock, A., & Ammann, C. Volcanic forcing of climate over the past 1500 years: And improved ice core-based index for climate models. *J. Geophys. Res.* **113**, D23111 (2008).

62. The figure was reproduced from Lovejoy, S. & Varotsos, C. Scaling regimes and linear/nonlinear responses of last millennium climate to volcanic and solar forcings. *Earth Syst. Dynam.* **7**, 1–18 (2016).

63. Note that the anthropogenic forcing is fairly smooth and is constantly increasing, whereas the sunspot-based solar variability is a random, wandering-type forcing because it has $H > 0$. The 10,000-year timescale at which the extrapolation of the hybrid solar forcing crosses the dashed anthropogenic forcing line is the scale at which the typical change in the wandering solar signal would typically start to dominate the anthropogenic forcing that has occurred to date.

64. Figure 5.12 plots the root mean square (RMS) fluctuations that have an exponent slightly less than $H$ (which is instead defined with respect to the mean absolute fluctuations); the difference between the two (–0.3 instead of –0.4) is the result of the fairly large volcanic intermittency.

65. Lovejoy, S., Schertzer, D., & Varon, D. Do GCM's predict the climate . . . or macroweather? *Earth Syst. Dynam.* **4**, 1–16 (2013).

66. Actually, for each of the three shown, there were subvariants with different land-use scenarios, but these made very little difference.

67. Petit, J. R., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J. M., Basile, I., Bender, M., Chappellaz, J., Davis, J., Delaygue, G., Delmotte, M., Kotlyakov, V. M., Legrand, M., Lipenkov, V., Lorius, C., Pépin, L., Ritz, C., Saltzman, E., & Stievenard, M. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 429–436 (1999); quote, p. 436.

68. Other shorter Greenland cores confirm that the GRIP results are typical of the Greenland Holocene, and scaling spectral analyses agree with the Haar analysis presented here. See, for example, Blender, R., Fraedrich, K., & Hunt, B. Millennial climate variability: GCM simulation and Greenland ice cores. *Geophys. Res. Lett.* **33**, L04710, doi:10.1029/2005GL024919 (2006).

69. Berner, K. S., Koc, N., Divine, D., Godtliebsen, F., & Moros, M. A decadal-scale Holocene sea surface temperature record from the subpolar North Atlantic constructed using diatoms and statistics and its relation to other climate parameters. *Paleoceanography* **23**, doi:Doi:10.1029/2006pa001339 (2008).

70. Adapted from Lovejoy, S. & Schertzer, D. In: *Extreme Events and Natural Hazards: The Complexity Perspective* (eds. A. S. Sharma, A. Bunde, D. N. Baker, & V. P. Dimri), pp. 231–254 (2012).

71. McIntyre, S. & McKitrick, R. Corrections to the Mann et al. (1998) proxy data base and northern hemispheric average temperature series. *Energy Environ.* **14**, 751–771 (2003). McIntyre, S. & McKitrick, R. Hockey sticks: Principal components and spurious significance. *Geophys. Res. Lett.* **32**, L03710–L03714 (2005).

72. Mann, M. E., Bradley, R. S., & Hughes, M. K. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392**, 779–787 (1998).

73. Mann, M. E., Bradley, R. S., & Hughes, M. Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophys. Res. Lett.* **26**, 759–762 (1999).

74. Mann, M. E., Bradley, R. S., & Hughes, M. K. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392**, 779–787 (1998).

75. Huang, S. Merging information from different resources for new insights into climate change in the past and future. *Geophys. Res. Lett.* **31**, L13205, doi:doi: 10.1029/2004 GL019781 (2004).

76. Huang's series only went back to 1500 AD. To obtain annual resolution, he also used dendrochronology, but the critical decadal and longer scales were from the boreholes.

77. Moberg, A., Sonnechkin, D. M., Holmgren, K., Datsenko, N. M., & Karlén, W. Highly variable northern hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature* **433**, 613–617 (2005).

78. Ljundqvist, F. C. A new reconstruction of temperature variability in the extra-tropical northern hemisphere during the last two millennia. *Geograf. Annal. Physical Geogr.* **92A**, 339–351 (2010).

79. Marcott, S. A., Shakun, J. D., Clark, P. U., & Mix, A. C. A Reconstruction of regional and global temperature for the past 11,300 years. *Science* **339** (1198), doi: 10.1126/science.1228026 (2013).

80. This has various problems. See Lovejoy, S., Varotsos, C., & Lambert, F. Atmospheric scaling and climate variability across scales. *Earth Space Sci.* (submitted, September 2018).

81. Adapted from Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades.* (Cambridge University Press, 2013).

82. Mann, M. E., Bradley, R. S., & Hughes, M. K. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392**, 779–787 (1998). Mann, M. E., Bradley, R. S., & Hughes, M. Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophys. Res. Lett.* **26**, 759–762 (1999).

83. Mann, M. E., Bradley, R. S., & Hughes, M. K. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392**, 779–787 (1998).

84. Mann, M. E., Bradley, R. S., & Hughes, M. Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophys. Res. Lett.* **26**, 759–762 (1999).

85. Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., & Johnson, C.A. (eds.) (Cambridge University Press, 2001).

86. In the current epoch, the difference between the maximum and minimum distance is about 1.4%.

87. The current tilt is 23.44°, roughly halfway between its extreme values. The tilt last reached its maximum in 8700 BC. It is now in the decreasing phase of its cycle, and will reach its minimum around 11,800 years from now.

88. There is a final, roughly 100,000-year forcing resulting from changes in the inclination of Earth's orbit around the sun, but it is a very small effect.

89. These percentages are the extra—above background—signal in these peaks. Lovejoy, S. & Lambert, F., High resolution EPICA ice core dust fluxes: Intermittency, extremes and Holocene stability. *Climate Past* (submitted, August 2018).

90. For a technical overview, see, for example, Dijkstra, H. *Nonlinear Climate Dynamics.* (Cambridge University Press, 2013).

91. Hays, J. D., Imbrie, J., & Shackleton, N. J. Variations in the earth's orbit: Pacemaker of the Ice Ages. *Science* **194**, 1121–1132 (1976).

92. Shackleton, N. J. & Imbrie, J. The $\delta^{18}$O spectrum of oceanic deep water over a five-decade band. *Climatic Change* **16**, 217–230 (1990).

93. For example, it has been suggested by Veizer et al. [Veizer, J., Godderis, Y., & Francois, L. M. Evidence for decoupling of atmospheric CO2 and global climate during the Phanerozoic eon. *Nature* **408**, 698–701 (2000).] that for the tropical oceans, as much as two thirds of the variation in $\delta^{18}$O is a result of the sequestering effect. For the early Oligocene (33,000,000 years to the present), this is close to the estimates of Zachos et al. [Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686–693 (2001).].

94. We use the "canonical" calibration coefficient $-4.5$K/$\delta^{18}$O on the basis of laboratory experiments and near the earlier calibration of $-4$K/$\delta^{18}$O. See Shaviv, N. J. & Veizer, J. Celestial driver of Phanerozoic climate? *GSA Today* **July**, 4–10 (2003). And recommended by Barras, C., Duplessy, J.- C., Geslin, E., Michel, E., & Jorissen, F. J. Calibration of $\delta18$ cultured benthic foraminiferal calcite as a function of temperature. *Biogeosciences* **7**, 1349–1356 (2010). Shackleton, N. J. & Imbrie, J. The $\delta18$O spectrum of oceanic deep water over a five-decade band. *Climatic Change* **16**, 217–230 (1990).

95. In any event, because glaciation is ultimately temperature dependent, the $\delta^{18}$O variations over the low frequencies are still, presumably, largely temperature driven—even if modulated by geology.

96. Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

97. The figure is adapted with permission and the calibrations are discussed in Lovejoy, S. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynam.* **44**, 3187–3210 (2015).

AQ: Please provide title of book.

98. Watson, A. J. & Lovelock, J. E. Biological homeostasis of the global environment: The parable of Daisyworld. *Tellus* **35B**, 284–289 (1983).

99. Charlson, R. J., Lovelock, J. E., Andreae, M. O., & Warren, S. G. Oceanic phyto-plankton, atmospheric sulphur, cloud albedo and climate. *Nature* **326**, 655–661 (1987).

100. Tyrrell, T. *On Gaia: A Critical Investigation of the Relationship between Life and Earth.* (Princeton University Press, 2013).

101. Berger and Loutre (1991).

102. Ruddiman and Thomson (2001).

AQ: Please provide complete reference details for endnotes 101 and 102.

{ 6 }

# What have we done?

## 6.1  Into the fray

It was April 11, 2014, and the McGill University press release went online at 1:30 in the afternoon. Although I'd published many articles, they were on fundamental geoscience; the release summarized the first one that had significant social and political consequences. Its title, "Scaling Fluctuation Analysis and Statistical Hypothesis Testing of Anthropogenic Warming,"[1,a] was arcane, but the release was clear enough: "Statistical analysis rules out natural-warming hypothesis with more than 99% certainty" (the article, published in *Climate Dynamics*, is hereafter referred to as CD).

It had been fifteen months since the original submission went to peer review, but now the pace picked up dramatically. Within hours, the tone was set by the skeptic majordomo Viscount Christopher Monckton of Brenchely, who displayed his Oxbridge classics erudition by deliciously qualifying the paper as a "mephitically ectoplasmic emanation from the Forces of Darkness." [b] Three days later, with the release getting 12,000 hits per day, the "Friends of Science" sent an aggressive missive to the McGill chancellor asking that it be removed from McGill's site.[c] The Calgary-based group with its Orwellian name was set up in 2002 to promote the theory that "The sun is the driver of climate change. Not you. Not $CO_2$." (Fig. 6.1). One could understand their thunder. Rather than trying to prove that the warming was anthropogenic—something that is impossible to do "beyond reasonable doubt"—the new paper closed the debate[2] by doing something far simpler: by disproving the "Friends GNF hypothesis. If we exclude either divine or extraterrestrial intervention, then the warming is natural or it is

---

[a] For the press release, see http://www.physics.mcgill.ca/~gang/Society/McGill.Press.release.27.4.14. pdf. By the end of the year, in terms of media attention, it had attained the status of the most "mentioned" of more than 750 articles in 2014 in the journal *Climate Dynamics*.

[b] On the "The Watts Up with That" (or, WUWT) website run by Anthony Watts: https://wattsupwiththat.wordpress.com/2014/04/11/lovejoys-99-confidence-vs-measurement-uncertainty/ . The site touts itself as "the world's most viewed site on global warming and climate change."

[c] McGill ignored the request, but this was only the beginning of hostilities with this group.

human; there is no third alternative. The skeptics were stuck. To add insult to in-
jury, their prepackaged sermons on the inadequacies of computer models or their
speculations about solar variability were irrelevant.



FIGURE 6.1 *The billboard war on science waged by the "Friends of Science" (upper and bottom
left), and defended (upper right) by Quebec's association of scientific communicators Association
des Communicateurs Scientifiques, November 2014. (Top) Montreal billboards. (Bottom left)
Toronto billboard. (Bottom right) Ottawa billboard. The English version of the top left is: "The
sun is the main driver of climate change. Not you. Not $CO_2$." The translation of the upper right
billboard is "What Science really says: The climate is changing. Because of us." In March 2015,
as a result of many complaints—including one that I made—Advertising Standards Canada
determined that the Friends of Science billboards contravened articles 1(e) and 8 of their code.
(Article 1e states: "Both in principle and practice, all advertising claims and representations must be
supportable . . . ." Article 8 states: "Advertising claims must not imply that they have a scientific basis
that they do not truly possess . . . .") To protest the "Friends" billboard campaign, I posted a rebuttal
on my McGill website. In February 2015, in a long missive addressed to the McGill Chancellor, the
"Friends" used scarcely veiled threats of legal action to try to force McGill to take the posting down,
something McGill refused to do. Unfortunately, in Canada, there are no legal consequences for
tenacious advertising.*

    Provoked by the media attention and several Op-Eds[d] in the hours, days, and
weeks that followed, in email, blogs, and Twitter, I was treated to a deluge of

---

    [d] *Live Science*, "Is Global Warming a Giant Natural Fluctuation?" April 18, 2014, http://www.
livescience.com/44950-global-warming-natural-fluctuation.html. And in *The Gazette*, "Research
Shows That Global Warming Isn't Natural," June 10, p. A17, http://www.physics.mcgill.ca/~gang/pop-
ular.articles/Gazette.7.14/Gazette.2.op.ed.orignial.10.6.14.jpg.

abuse: "atheist," "Marxist," "hippy name," and so on—everything, it seemed, short of death threats. Reminiscing later with my colleague Gavin Schmidt—who had been on the firing line for years—I realized that I had only received the standard treatment from the well-funded,[e] well-organized climate skeptic[f] community. Schmidt summed up his own experience: "You're roadkill. It's like having a spotlight trained on you until they move on to their next victim."

But Monckton was mistaken not only about the science substance, but also about the science process. In the very first sentence of his blog, he railed: "It is time to be angry at the gruesome failure of peer review that allows publication of papers, such as the recent effusion of Professor Lovejoy of McGill University . . . ."[3]

The comment was ironic, not least because the manuscript had been rejected successively from three different journals—each time on spurious grounds—and that very morning, a follow-up manuscript[g] had been dismissed by *Nature Climate Change* without even being sent to review.[h] One might have thought that the climate skeptics would be pleased with this 75% rejection rate. No system is perfect and this one was apparently doing its best to keep out the new approach!

To understand this unnatural convergence of skeptic and mainstream views, recall that, for decades, the primary approach to climate science had been so dominated by GCMs that scientists and skeptics alike found it difficult to imagine solving fundamental climate problems without them. Caricaturing this blinkered mindset, one irate reviewer had even claimed that the statistical testing of natural warming could only be achieved with GCMs and told me to "go get your own GCM!"[i] The reception by both scientists and skeptics thus revealed as much about the sociology of science as its content—in particular, the disjunction that had grown between climate science and nonlinear geophysics.

---

[e] According to one study, in the United States alone there are currently *ninety-one* think tanks, advocacy groups, and trade associations, with a combined funding of more $900,000,000 that promote climate change denial. Brulle, R. J. Institutionalizing delay: Foundation funding and the creation of U.S. climate change counter-movement organizations. *Climatic Change* **122**, 681–694 (2014).

[f] I use the term "climate skeptic" to denote those who dispute the theory of anthropogenic global warming. But today, the term "climate denier" is more accurate. Yet the term "denier" has an unnecessary emotional charge, so that, for example, "denialist" has been proposed instead. I find this new proposal awkward and will stick with "climate skeptic." See, however, Gillis, J. Verbal warming: Labels in the climate debate, *New York Times*, p. D1, February 17, 2015.

[g] When it was published a few months later in a more specialist journal, it was front-page news: Believe it: Global warming is real, Global warming slowdown just a "pause." *Montreal Gazette*, July 24, 2014 (see Section 6.6). Lovejoy, S. Return periods of global climate fluctuations and the pause. *Geophys. Res. Lett.* **41**, 4704–4710 (2014).

[h] Rejection without review is a common practice for high-end publications such as *Science* or *Nature,* affecting as much as 90% of the submissions. It sent a clear signal that their scientifically conservative editors found my approach to climate variability to be of limited interest.

[i] Unfortunately for the reviewer, currently, the natural warming hypothesis can *only* be disproved by using empirical data with the help of some nonlinear science. As discussed in Chapter 5, GCMs *cannot* do this because their relevant preindustrial centennial-scale natural variability isn't sufficiently realistic. See also Lovejoy, S., Schertzer, D., & Varon, D. Do GCM's predict the climate . . . or macroweather? *Earth Syst. Dynam.* **4**, 1–16 (2013).

I was not much surprised by these initial rejections. When it comes to genuinely new ideas, science is no different than other areas of life: Acceptance requires protracted struggle. In the real world, one *expects* that the unfamiliar will be rejected. This is nothing like the skeptics' fantasy world, in which climate scientists conspire to foist their pet theories on an unsuspecting public. To make this fairy tale plausible, scientists are portrayed as money-grubbing businesspeople seeking fat research grants. When the skeptics accused me of "wasting tax payer money," it was no more than a ritualistic incantation. The disclaimer[j] at the bottom of the article ("this work was unfunded; there were no conflicts of interest")[k] was irrelevant.

The accusation of being financially motivated with insinuations of misconduct was particularly rich. Back in 1998, to demonstrate their commitment publicly to the newly ratified Kyoto Accords, the Canadian government re-earmarked $150,000,000 of environment and other funds to create the three-year Canadian Climate Action Fund (CCAF).[l] In March 2000, taking advantage of this opportunity, I submitted a proposal to perform a hypothesis testing study virtually identical to the one finally published fourteen years later.[m] Curiously, the proposal was not rejected on scientific grounds (it was admitted that the proposal would indeed fill "a knowledge gap"); rather, it was rejected because the agency doubted that "the deliverables" (techno-speak for "the results") would be ready on time (one year later). This reasoning was bizarre because, depending on the result of the research, filling the "knowledge gap" could potentially save the government $150,000,000, yet it was not ready to spend the requested $28,000 for an assistant to help me find out if the warming was no more than a GNF.

At the time, I only had a few climate publications—the rest concerned atmospheric variability at shorter weather timescales—and this bureaucratically motivated rejection only confirmed my suspicions that climate research was too political to be worth pursuing. A few years later, following the 2006 election of the climate-skeptical Conservative party in the federal elections, the

---

[j] See Oreskes, N., Carlat, D., Mann, M. E., Thacker, P. D., & vom Saal, F. S. Viewpoint: Why disclosure matters. *Environ. Sci. Technol.* **49**, 7527–7528 (2015).

[k] At the time, I had a small National Sciences and Engineering Research Council grant, but it was earmarked for a different project ("Multifractal Geophysics") and, in any case, it barely covered my publication costs.

[l] The fund was set up to "to help develop a national implementation strategy and support early actions to respond to climate change . . . to make possible Canada's ratification of the Kyoto Protocol on December 17, 2002." *Building on Success Climate Change Action Fund (CCAF), 2002–2003 Annual Report.* http://publications.gc.ca/collections/Collection/M171-2-2003E.pdf; quote, p. 2. Ninety-three percent of the fund was in the form of tax breaks to corporations which claimed to be mitigating and adapting to climate change; only 7% was for research. But even this meager offering could only be used as matching funds: A minimum of 50% of the research had to be funded from other sources. In my case, I could only offer an "in kind" contribution, a donation of a percentage of my time (and hence salary).

[m] The proposal, submitted in March 2000 to the CCAF, had two deliverables: (1) a new analysis of climate extremes ["especially with regard to the possible existence of long-range statistical correlations, which give rise to spurious trends, and to fat-tailed probability distributions which, give rise to spurious transitions" (i.e., to "black swans," although the term had yet to be coined)] and (2) the development of "appropriate statistical tests capable of rejecting the (nonclassical) null hypothesis at various confidence levels." The CD paper addressed both issues.

Canadian Fund for Climate and Atmospheric Science (CFCAS; the successor organization to the CCAF) was shut down.[n] Because this was the only public body that specifically funded academic research into the environment and climate, when the CD research was finally performed in 2013, there were no longer any funding possibilities in Canada. I had no choice but to do the research in my spare time.

## 6.2 The damage so far

We have seen that the atmosphere is a turbulent fluid with temperature, humidity, wind, and other fields that vary from tiny millimeter-sized eddies to huge planetary weather systems, from milliseconds to the age of Earth. In Chapters 1 and 2, we used proxy temperatures helped by fluctuation analysis to analyze the variability quantitatively out to megaclimate scales. Although the interpretation of the proxies is not always straightforward, qualitatively different lines of evidence often confirm the conclusions. For example, 65,000,000 years ago, we know that dinosaurs roamed an ice-free South Pole, and this is consistent with proxies that indicate global temperatures 5°C or even 10°C warmer than today. We also know that as little as 14,000 years ago, Earth was still in the throes of an Ice Age, with continent-size ice sheets several kilometers thick, and the proxies show that the global temperatures were indeed 2° to 4° cooler than today (the larger value being relevant at higher latitudes).

We can therefore be quite confident that, over geological epochs, the temperature varies quite considerably. Although the quantitative amounts of warming and cooling may be debated, there is no doubt that over sufficiently long periods, the temperature of Earth can change readily by several degrees. But what about this: Since the end of the nineteenth century, instrument records show that Earth has warmed by about 1°C. The evidence of a warming is all around us—from the melting of polar sea ice (including the summer opening of the Northwest Passage) to rising sea levels to deadly heat waves. But what is the cause? Is it simply a routine natural fluctuation or is it something different, something artificial, something only *we* could have done? More precisely, is a 1° warming of the whole planet[4] *in only a single century* an ordinary—even common—event in the history of Earth? Or is it so exceptional, so unlikely, as to demand a non-natural explanation?

[n] The CFCAS was replaced by the Climate Change and Atmospheric Research Programme (CCAR), but this only funded a handful of megaprojects aimed at specific applications, not fundamental research or small-scale projects. In January 2018, under the new Liberal party government, the CCAR was, in its turn, terminated, implying—among other consequences—the shutdown of the unique polar observatory The Polar Environment Atmospheric Research Laboratory. Ironically, it was Kirsty Duncan in her new role as Liberal Party Minister of Science who announced the shutdown. As an opposition MP, Duncan had been a vocal opponent of an earlier Conservative party government attempt to do exactly the same thing. This new attempted shutdown provoked international outcry—a letter signed by 250 climate scientists worldwide. This resulted in the station being again saved—this time until 2019.

The modern answer to this question emerged well before human emissions had changed the composition of the atmosphere significantly and before the warming itself was noticed. In 1896, in an attempt to understand the causes of the Ice Ages, Swedish chemist Svante Arrhenius (1859–1927) estimated that if the concentration of atmospheric $CO_2$ was doubled, global temperatures would rise by 5 to 6°C, a result revised later by Guy Stewart Callander[5] (1897–1964) to 2°C.[6] From a scientific point of view, the basic action of $CO_2$ is straightforward. $CO_2$ is a "greenhouse gas." It lets visible light from the sun through to the surface while absorbing part of Earth's outgoing heat radiation (Fig. 6.2).[o] Box 6.1 discusses the atmospheric greenhouse effect, the operation of which differs in important respects from that of the glass-enclosed structure in your backyard.
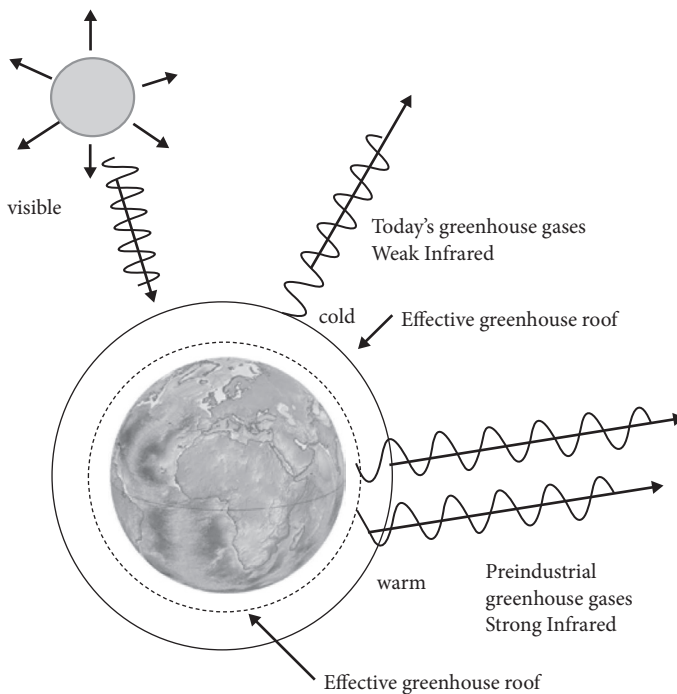


FIGURE 6.2 *The atmospheric greenhouse effect for wavelength bands that absorb greenhouse gases, showing how increasing greenhouse gases effectively increase the altitude of the greenhouse "roof." (The other way that increasing carbon dioxide increases warming is via an easier to understand mechanism that applies to infrared wavelengths that start off essentially transparent to carbon dioxide but start to become opaque to them at greater concentrations.) Because the atmosphere is colder at higher altitudes, a higher roof emits less infrared (heat) radiation, but this does not alter the incoming visible radiation that warms the surface. Therefore, the atmosphere must warm up to regain thermodynamic equilibrium. The dashed circle shows the lower preindustrial "roof"; the other circle shows today's higher, colder roof.*

[o] Although this explanation is more or less correct for Earth in thermodynamic equilibrium, the consequences of *increasing* $CO_2$ are not so obvious (see Box 6.1).

BOX 6.1 The atmospheric greenhouse effect

The overall temperature of Earth is determined by the balance of solar energy it absorbs at visible wavelengths and then re-emits into outer space at infrared wavelengths (see also Box 7.1). This basic physics was discovered by Joseph Fourier (1768–1830); a bit later, the role of $CO_2$ was discovered by John Tyndall (1820–1893). Tyndall added the idea that visible radiation from the sun can come in, but heat radiation is trapped by the $CO_2$ and can't escape. This trapping is the mechanism behind what was later[p] called the "greenhouse effect." To avoid confusion later, let's call this the *radiative* greenhouse effect.

Although the radiative greenhouse effect works both in real greenhouses and in the atmosphere, in neither case does it give ~~good~~ explanations of the phenomena. First consider standard, infrared-absorbing glass enclosures, which are transparent only to visible light. It turns out that when this glass is replaced by a material transparent at both visible and infrared wavelengths, the resulting greenhouse works almost as well![q] Modern research shows that there is, indeed, a radiative greenhouse effect, but it is much, much smaller than the dominant "convective" greenhouse effect. Both effects work by trapping, but the latter traps *air* rather than radiation. Both inside and outside the greenhouse, the surface absorbs sunlight, which warms a layer of near-surface air. This warmed air then rises (convection). The key inside–outside difference is that, inside, the walls and roof prevent air from rising farther than the roof whereas, outside, the air continues to rise, expand, cool, and transfer its energy to the higher reaches of the atmosphere. In greenhouses, the warm air trapped by the roof is only very slightly augmented by trapped radiation.

In the atmosphere, many gases contribute to the radiative greenhouse effect and, overall, it explains why Earth is about 33°C warmer than it would have been in the absence of an atmosphere. However, without going further, this only explains a fixed amount of warming. Pioneers Arrhenius and Callender faced a different problem: What is the *change* in the equilibrium temperature of the atmosphere when the concentration of greenhouse gases is *increased* from its current value? Even when restricted to the radiative effect, the greenhouse analogy explains only part of the story. There are two basic cases to consider: (1) infrared wavelengths for which there is currently only a small amount of absorption and (2) those for which there is already substantial absorption. The former case is indeed like a greenhouse, but with a roof that is so thin, that increasing its thickness traps more radiation. At these wavelengths, the atmosphere is currently ineffective at trapping, but at greater $CO_2$ concentrations it may become absorbing. At these weakly absorbing wavelengths, increasing $CO_2$ concentrations directly increases the basic radiative greenhouse

---

[p] It's not clear who first used the term "greenhouse effect" for atmospheric warming, but it was around 1909. See a historical discussion in Steve Easterbrook's Serendipity blog. *Who First Coined the Term "Greenhouse Effect?"* August 18, 2015. http://www.easterbrook.ca/steve/2015/08/who-first-coined-the-term-greenhouse-effect/.

[q] This was demonstrated by Robert W. Wood in 1909. He built two identical greenhouses but covered one of them with transparent rock salt rather than with ordinary glass, and he found similar amounts of warming: Wood, R. W. Note on the theory of the greenhouse. *Philosophical Magazine,* 6th series, **17**, 319–320 (1909).

effect. However, understanding what happens at wavelengths where $CO_2$ is already strongly absorbed is trickier. This case is more like a greenhouse with glass thick enough to trap virtually all of the heat radiation. For these wavelengths, doubling $CO_2$ concentrations—doubling the roof's thickness—would hardly make any difference.

To understand how increased $CO_2$ concentrations lead to warmer temperatures in this case, recall that matter of any kind, including Earth's surface, emits "black-body" radiation, which carries away energy: The higher the temperature, the higher the overall rate of energy emission. Also, most of the energy of a black body is concentrated in a particular range of wavelengths. For the sun at 5,250°C, this is in the visible range (wavelengths of around 0.5 μm) whereas for Earth, with an average at around 16°C, most of the emission is in the "thermal infrared," at around 10 to 12 μm.

If there was no atmosphere—or for the wavelengths at which there is virtually no absorption—surface photons would travel unhindered to outer space, and the rate at which they removed energy would be relatively high—characteristic of the high surface temperature. However, for $CO_2$-sensitive bands, surface photons will typically be absorbed by $CO_2$ higher up in the atmosphere, where temperatures are cooler. These photons are subsequently re-emitted and either escape directly to outer space or are reabsorbed, then re-emitted, reabsorbed, re-emitted, and so on. Eventually, they reach such high altitudes that they have very little $CO_2$ above them, so they can finally escape to space without being reabsorbed. This final escape altitude is the "effective" height of the greenhouse roof. Because this high-altitude roof is colder than the surface, it emits energy at a lesser rate. Although the altitude of the roof is different for different wavelengths (depending on how strongly the radiation at that wavelength interacts with $CO_2$), the roof at any given wavelength is still colder than the surface and still emits less.

Now imagine Earth without an atmosphere. The temperature would be determined by the condition that the overall incoming visible radiation absorbed at the surface would equal the outgoing infrared radiation from the same surface. In the case of Earth, this would be about –17°C. Now, add an infrared-absorbing and -emitting atmosphere. The surface heating by incoming visible light is the same, but now the outgoing infrared radiation is from the high-altitude cold roof. In order for Earth to be in thermodynamic equilibrium with its atmosphere, it must increase its outgoing radiation by warming up. The cold roof thus explains the basic atmospheric greenhouse effect.

Using the roof idea, we can also understand what happens when we increase $CO_2$ concentrations. First, consider wavelengths that are initially nonabsorbing. At these wavelengths, there is effectively no roof, so they emit at surface temperatures. However, at greater $CO_2$ concentrations, they might start absorbing, effectively adding a cool roof where before there was none. In the other case, in which the atmosphere already absorbs $CO_2$ so that there is initially a cool roof, increasing $CO_2$ concentrations raises the altitude at which photons escape to outer space, effectively raising the roof to a higher altitude where it is cooler[r] (Fig. 6.2). At these lower roof

---

[r] This modern explanation was first clearly explained by Ekholm: Ekholm, N. On the variations of the climate of the geological and historical past and their causes. *Q. J. Roy. Meteorol. Soc.* **27** (117), 1–62 (1901).

temperatures,[s] Earth's outgoing infrared radiation is diminished in intensity whereas the incoming solar radiation is unaffected.[t] In all cases—either bands that have been recently promoted to absorbing status acquiring a cool roof for the first time, or bands with a roof that moves to higher altitudes where it is cooler—Earth will heat up to return to thermodynamic equilibrium. It heats up because the hotter it gets, the more it emits, until eventually (over time), an overall energy balance is restored.[7] Although we have couched this discussion in terms of $CO_2$, it is valid for the other infrared-absorbing and -emitting gases.[8]

Let's follow convention and continue to use the term "greenhouse effect," yet when applied to the atmosphere, it further stretches the greenhouse analogy. Skeptics like to point out—as though scientists had missed it—that water vapor is a much more potent greenhouse gas than $CO_2$ and that, in addition, water in the form of clouds also has a big effect on Earth's radiative equilibrium, so why worry about a much smaller $CO_2$ contribution?

To understand this, we must make a distinction between the long-lived greenhouse gases that stay in the atmosphere for decades ($CH_4$) or millennia ($CO_2$), and water vapor and clouds that are highly variable in space and in time, with a typical water molecule staying in the atmosphere only over weather scales, not climate scales. Although the turbulent distribution of water (vapor and clouds) must be modeled explicitly in GCMs (and is one of their most uncertain components), in contrast, greenhouse gases are mixed fairly uniformly by the weather and are quite rapidly (over a year or two) spread around the whole atmosphere. When greenhouse gas concentrations are increased, the result is a kind of uniform "background" increase in surface heating. But, in other respects, greenhouse gases don't affect atmospheric dynamics. This is why they can all be lumped together into a "$CO_2$ equivalent": the concentration of $CO_2$ that gives the same overall radiative effects. Although the extra energy flux is distributed fairly uniformly over Earth, the complication—important at climate scales (and discussed in Chapter 7)—is that different parts of Earth react more or less strongly to an increase in the heating: Their "climate sensitivities" are different so that some regions warm more than others (discussed later).

Arrhenius's theory signaled the beginning of modern attempts to prove the anthropogenic provenance of a warming that only became strongly apparent during the 1980s. From a purely scientific point of view, the main difficulty is that there are complicated feedbacks between $CO_2$, water vapor, and clouds: Increasing $CO_2$ increases the temperature, and this increases evaporation. The extra water vapor increases the greenhouse effect and warming, but it also increases cloud cover, and this could have either warming or cooling effects depending on the altitude and other complex-to-analyze effects. Indeed, it is the difficulty in modeling this water

---

[s] In the atmosphere, up until the tropopause, the average temperature falls off with increasing altitude.

[t] For a more detailed but still accessible review, see Pierrehumbert, R. T. Infrared radiation and planetary temperature. *Phys. Today*, <SL: Please provide month.> 33–38 (2011). See also the (modern) introduction to the reprint of Arhennius's manuscript in Archer, D. & Pierrehumbert, R. <SL: Please provide title of book.> (Wiley-Blackwell, 2011).

feedback cycle that is largely responsible for the uncertainty in estimates of the warming produced by increases in $CO_2$ concentrations. Arrhenius spent the best part of a year with pencil and paper grappling with these complications; today, they are handled by supercomputers.[u]

Is the warming mostly human-made, through changes in land use, the emission of greenhouse gases, and aerosols (pollution)? Or is it just normal, natural variability? Today, the theory of anthropogenic warming is entering a mature phase in which continued efforts to prove it more convincingly are suffering from diminishing returns. Take, for example, the IPCC's Fifth Assessment Report (AR5; 2013).[9] Not withstanding massive improvements in computers and algorithms, and resulting largely from real scientific difficulties, it cited exactly the same range of temperature increase for a doubling of $CO_2$ as did the US National Academy of Science report in 1979: 1.5 to 4.5°C. Whereas the fourth report (AR4, 2007) stated that it is "likely that human influence has been the dominant cause of the observed warming since the mid-20th century," six years later, the AR5 only upgraded this to "extremely likely."[v]

Despite the strong evidence in favor of the anthropogenic theory, it still faces a chorus of denial backed by entire organizations—think tanks, advocacy groups, and trade associations—that collectively comprise a "climate change countermovement."[10] Significantly, these groups spend no money on scientific research that might prove their theories are correct; instead, they only *invoke* solar, volcanic, and internal climate system variability as *plausible* alternatives to the anthropogenic theory. How can we break through the impasse to "close" the debate?

At this point, it is helpful to recall that science progresses not only by attempting to prove certain theories to be true, but also by rejecting theories that are false. In this, it benefits from a fundamental asymmetry in scientific methodology: Although no theory can ever be proved true "beyond reasonable doubt," even a single decisive experiment can disprove one that is otherwise highly seductive. In their day-to-day work, scientists constantly reject ideas and theories that are incompatible either with observations or with more powerful theories in whose truth they have greater confidence.

To appreciate fully the ability of the scientific method to reject false theories, consider a common medical situation. The human body is a highly complex system, and to come up with a new drug or treatment, pharmacologists focus on only one or two fundamental components ("pathways"). They develop a theory

---

[u] The difficulty of modeling clouds—especially those associated with atmospheric convection—is a major argument used by meteorological services for justifying the acquisition of bigger computers. It is claimed that if GCM resolutions could be decreased to kilometric scales, the resulting high-resolution GCMs would be "cloud resolving," yielding large improvements. By now, however, it should be clear that it is sufficient that the small-scale *statistics* be realistic, and this requires judicious use of the turbulent laws, something that could be done with current—or even lower—GCM resolutions. What is the point of resolving features in the model as small as 1 km that will live for fifteen minutes and will then simply be "averaged out" in a forecast?

[v] In IPCC parlance, "extremely likely" refers to a 95% to 100% probability level.

explaining how a particular drug might be effective. But no matter how beautiful or promising the idea, no one will take the drug without clinical trials. One then performs experiments capable of rejecting the drug if it fails. It is important to appreciate that ineffective treatments can be rejected *even when there is no understanding of the underlying biology*. The converse is also true: Treatments may turn out to be effective even in the absence of a theory. Until recent decades, almost all new drugs were, in fact, discovered by accident, and many of the mechanisms of established drugs are still not understood.

Sometimes theories such as "young Earth"[w] can be rejected confidently because they would contradict of a huge body of evidence. In other cases, the degree of confidence of the rejection must itself be quantified using statistics. Medicine again provides a typical example. Properly designed experiments test both new treatments and placebos on an ensemble (collection, population) of patients with similar afflictions. To avoid bias, both the patients and the scientists are unaware of who gets what; the experiment is "double blind." As a result of various difficult-to-control or simply unknown factors (notably including the placebo effect!), even with an effective drug, some of the patients getting the placebo typically show a positive response whereas some of the medicated patients show none. The outcome of drug trials therefore requires statistical comparisons of the results between two populations. A possible experimental outcome might be a statement such as, "The hypothesis that the medication is clinically effective can be rejected at the 99% level." In this case, we are effectively admitting that there is a small residual chance (1%) that the medication is, in fact, effective; but, common sense and tradition have it that when the probability that the medication is effective is very low, the medication should not by approved for use in the general population.

This example illustrates yet another feature of hypothesis testing: Subjective elements can never be totally eliminated. Few people would take a medication if it only had a 1% chance of working.[x] But what if the experiment found that the medication worked in 30% rather than 1% of the patients? In this case, the conventional conclusion is that the confidence is not high enough and the issue would be have to be resolved by changing the experimental protocols and/or by increasing the size of the population being tested. The traditional threshold for rejecting a hypothesis is 95% confidence, but there are situations—such as global warming—in which the implications of rejecting a hypothesis are so important that the bar would be set much higher.[11] If scientists could only reject the hypothesis that an asteroid was on a collision course with earth at the 95% level, one would not feel serene about the future and one would strive for higher levels of certitude. Obviously, when the stakes are high enough—and with global warming it is potentially the fate of ecosystems and civilizations—then exceptionally high levels of confidence are required.

---

[w] Bishop Usher's interpretation of the Bible that Earth was created in 4004 BC is currently embraced by many religious fundamentalists, including a large fraction of US congressmen.

[x] One could easily imagine exceptions!

The human body and the climate are both complex systems with many strongly interacting components. Why not apply standard scientific methodology to attempt to reject false theories of warming? Although in most cases of statistical testing, the rejection of one theory still leaves several possibilities open, in the case of global warming, the elimination of one virtually forces us to accept the other. In the next section, we describe how this was done.

## 6.3 Testing the GNF hypothesis

### 6.3.1 untangling forced and unforced variability

During the preindustrial epoch, at least up to scales of centuries, Earth's temperature was stable ($H < 0$) and we can conclude that the rate at which solar energy was absorbed and the rate at which it was re-emitted were roughly (albeit not exactly) equal. The planet never strayed far from thermodynamic equilibrium (Section 5.6, Box 7.1). However, during the industrial epoch, anthropogenic climate forcings have increased the rate that energy has been absorbed. To reestablish equilibrium, more infrared energy must be radiated to outer space, and this implies an increase in the average surface temperature.[12] But the temperature has lagged; it hasn't increased fast enough. This means that if tomorrow, $CO_2$, $CH_4$, and other human emissions stopped, so that climate forcings were fixed at their current levels,[13] the temperature would continue to rise (see Box 7.1 for energy storage and relaxation to equilibrium).

Physically, the origin of this "heating in the pipe" memory is easy to understand. Oceans cover 70% of Earth's surface and they are also dark. The overall effect is that more than 90% of the extra heating goes into the oceans, where it is stored and gradually redistributed by a (scaling) hierarchy of currents, some of which are very deep and very slow (Box 7.1). It takes decades, centuries, even millennia, for this increased heat to warm the atmosphere and Earth's surface, which will only then emit enough extra radiation into outer space for the system to reattain equilibrium at a higher temperature.

This memory means that Earth's temperature is still responding to very old climate forcings and will continue to increase even when the forcings are stabilized, and even for a while after they are reduced (Section 7.3). However, even at decadal scales, the delayed responses to old and current forcings are only part of the story. We saw that in control runs (no forcings) there was, nevertheless, a whole hierarchy of macroweather fluctuations, including decadal, centennial, and longer ones. How, then, do we disentangle the forced responses from them? We need to know this not only to determine how much warming we have caused, but also how much we may expect in the future when emissions are reduced, stopped, and possibly reversed.[y]

---

[y] Such "negative emissions" are possible, but it is not at all obvious that they can be deployed at a large-enough scale to remove much $CO_2$. The most plausible negative emission technology is Bio-Energy Carbon Capture and Storage (BECCS). Typical scenarios involve planting trees or other biomass over an area one or two times the size of India, regularly harvesting and burning it in special stations where the $CO_2$ is removed and then buried for millennia. See, for example, Anderson, K. &

The reason we can separate the forced and natural sources of variability is that they don't interact very much with each other. They seem to superpose nearly linearly. This means that if we double the forcings, the consequences for the forced response are also doubled (and the doubled responses are delayed in the same way as the undoubled responses). The superposed macroweather natural fluctuations are, on average, unaffected. Because the weather is highly nonlinear, this may seem surprising. The point is, we are only considering the effect of small changes to Earth's boundary conditions on the long-term climate state (i.e., after the weather variability has been averaged out). In the jargon introduced in Chapter 5, these changes are "climate forcings"[z] and they are relatively small. For example throughout the industrial epoch, the anthropogenic forcing is only (IPCC AR5, 2013) about 2 W/m$^2$ (Fig. 6.5B)—in other words, about 1% of the average 238 W/m$^2$ delivered by the sun (see Section 4.6). From this point of view, anthropogenic forcings are only a relatively small perturbation and, mathematically, the effect of small perturbations on the boundary conditions is expected to be linear.[aa] The response linearity means that a forcing of 2 W/m$^2$ will push us twice as far from our original climate state as a forcing of 1 W/m$^2$. Inasmuch as this approximation is valid, it is tantamount to assuming that nothing fundamentally new will happen, that there are no strongly nonlinear consequences of the increased forcing, and that there are no "tipping points"[bb] (Box 3.1). The possibility—even likelihood—that such highly nonlinear processes might push the system into a radically different climate state is indeed one of the dangers inherent in humanity's ongoing global warming experiment. Unfortunately—precisely because it is outside the realm of linear responses—it is not something that can as yet be predicted with any confidence.[cc]

It should also be clear that if we consider the global temperature response to greenhouse gas forcings, any notion of a linear response can only be meaningful at timescales long enough for the atmosphere and ocean to act as a single,

---

Peters, G. The trouble with negative emissions: Reliance on negative-emission concepts locks in humankind's carbon addiction. *Science* **354** (6309), 182–183 (2016).

[z] Remember that, in the wide sense used here, the atmospheric composition, including $CO_2$ concentration, is considered to be part of the GCM boundary conditions.

[aa] This is confusing because the atmosphere is still sensitively dependent on initial conditions (the butterfly effect), but this only implies that the detailed weather will change when the initial conditions are slightly altered. The questions here are: How does the small change in forcing change the long-term climate state? And, are macroweather variations about the new climate state of the same statistical type? Not surprisingly, there are theoretical, mathematical results that support the idea that the perturbations will indeed lead to linear changes in the climate state. See, for example, Ruelle, D. A review of linear response theory for general differentiable dynamical systems. *Nonlinearity* **22**, 855–870 (2009).

[bb] A frequently cited tipping point is the possibility that a small rise in Arctic temperatures would cause a massive release of $CH_4$ trapped in permafrost. This could initiate a catastrophic feedback cycle in which increases in warming are followed by increases in temperatures that provoke further increases in $CH_4$.

[cc] Indeed, in Chapter 7, we will see that the GCMs are all very linear in their responses to the proposed future emission scenarios. None predict tipping points when run with "plausible" anthropogenic forcings. However, as a result of GCM limitations—especially the differences between the climates of the different GCMs and the real climate—we should not be complacent!

unified (coupled) system. Analysis shows[14] that for timescales below the ocean weather–ocean macroweather transition at one- to two-year scales, the correlation between atmospheric and oceanic temperature fluctuations is very small, whereas at scales larger than this, they become very large.[15] The low amount of coupling at monthly scales means the atmosphere as whole may increase its temperature while, during the same month, the average ocean (surface) temperature might easily decrease. However, for fluctuations over timescales longer than the coupling scale, the land–ocean interaction is so strong, they both tend to increase or decrease together.

### 6.3.2 SIMPLISTIC OR SIMPLE?

It is plausible that climate states exhibit linear responses to changing forcings. In Chapter 7, we will confirm that GCMs respond nearly linearly to industrial epoch forcings. But what about the real world? If we are to reject the GNF hypothesis without using GCMs, we need to investigate the linearity from the empirical point of view using the temperature and forcing data for the industrial epoch. Figure 6.3 compares the globally averaged temperatures at annual resolution, and the $CO_2$ concentration[16] measured in terms of the number of concentration doublings when compared to a conventional reference value of 277 ppm (the rough value that pertained at the beginning of the industrial epoch).[dd] Recall that as a result of feedbacks with clouds and water vapor, although the exact efficacy of $CO_2$ in increasing the heating is uncertain, it nevertheless will depend on the number of $CO_2$ doublings[17]—not on the $CO_2$ concentration directly. A cursory glance at Figure 6.3 shows that the shape of the number of $CO_2$ doublings and of the temperature with time are very similar in appearance, with both notably rising much faster during the second half of the twentieth century.

---

[dd] Extrapolating Figure 6.4A back to zero doublings, we can see that a concentration of 277 ppm implicitly defines the beginning of the industrial epoch to be at about 1750, halfway between Newcomen's "atmospheric engine" (1713) and Watt's rotative steam engines (the 1780s). Although this date is somewhat sensitive to the reference value 277 ppm estimated for the preceding millennium, it can be used to give an objective definition to "Anthropocene." This fact is the basis of the suggestion that the alternative "Capitalocene" would be more accurate: Moore, J. W. *Anthropoce or Capitalocene? Nature, History and the Crisis of Capitalism*. (PM Press, 2016).
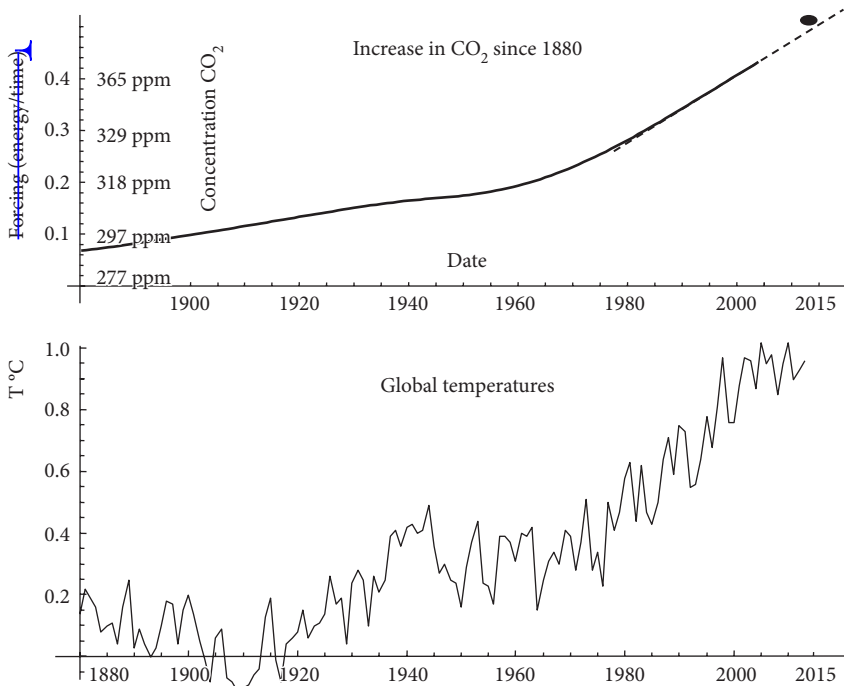
FIGURE 6.3 *(Top) The fraction of a carbon dioxide doubling that has occurred between 1880 and 2013, displayed using the preindustrial reference of 277 ppm. A value of 0.5 (half a doubling, a factor of 1.41) was exceeded in 2014 (the dot at the upper right). According to radiation physics (Box 6.2, Fig. 6.2), the extra energy reaching Earth's surface (the "radiative forcing") is proportional to the number of doublings. (Bottom) The globally averaged temperature (National Aeronautics and Space Administration-GISS), over the same period. Notice that the two graphs have nearly the same shape.*

This suggests that, rather than plotting temperature as a function of date, we should plot it directly as a function of the number of $CO_2$ doublings. The result is the impressively linear graph shown in Figure 6.4A. Its slope[18] (2.33 ± 0.36°C per $CO_2$ doubling) is called the "effective climate sensitivity" (EffCS), and it is the actual sensitivity of the climate to the historical increase in $CO_2$. We see that it is not far from the somewhat different ECS discussed earlier (1.5–4.5°C per doubling). The interpretation that the straight regression represents the total anthropogenic forcing is reinforced in Figure 6.4B, C. In Figure 6.4B, we simply overlaid the straight relationship of Figure 6.4A onto the curve that it implies for temperature as a function of calendar date since 1880. We see that it does an excellent job of reproducing the overall trend. Finally, the difference between the anthropogenic (climate) change and the observed temperature is the macroweather (internal) variability (Fig. 6.4C). This interpretation is supported by the fact that the amplitude of the fluctuations is reasonably bounded to within about 0.2°C of the average and shows no obvious systematic difference between the beginning and end of the series.
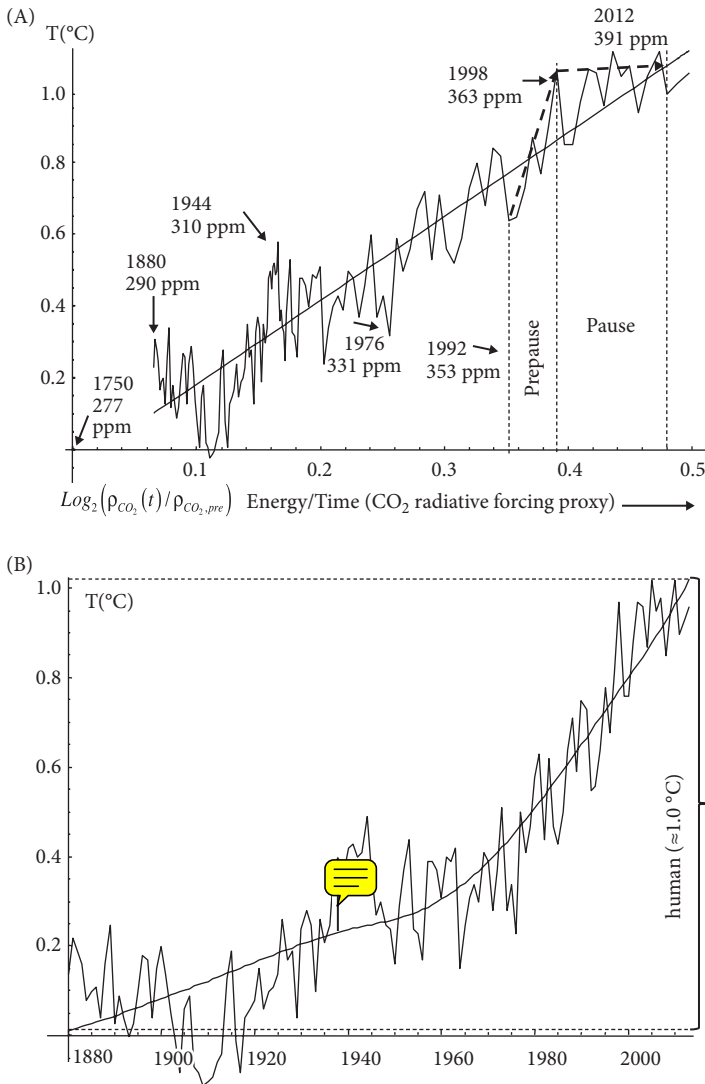
**FIGURE 6.4** *(A) The globally averaged annual temperatures plotted as a function of the number of $CO_2$ doublings instead of time (Fig. 6.3). The slope is the effective climate sensitivity—2.33°C per $CO_2$ doubling. (B) The temperature from Figure 6.3, but with the $CO_2$–temperature relationship determined by the straight line in (A) superposed (smooth line). The arrow shows a typical residual, which is the difference between the (anthropogenic) smooth curve and the actual temperature. The right-hand side parentheses show the total anthropogenic contribution to the warming since 1880. The smooth line represents the climate; the residuals are the macroweather anomalies. (C) The macroweather (residuals) deduced from (B) by removing the anthropogenic climate signal. Also shown (by the dashed lines and arrows) are several macroweather events discussed in the text.[19]*
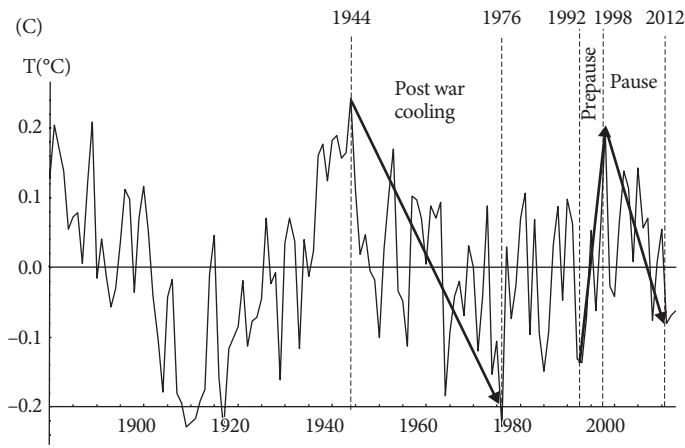
FIGURE 6.4 *Continued*

Yet, at first sight, the interpretation that the anthropogenic temperature change is simply 2.33 times the number of $CO_2$ doublings is too good to be true. It would seem to imply that the warming depends only on $CO_2$ concentrations, whereas there are other anthropogenic factors that are known to be important—with the main ones being changing land use, $CH_4$ emissions, and aerosols (particulate pollution). Indeed, Figure 6.4A had not been published before 2014, presumably because it is simplistic to attribute all warming to $CO_2$. For example, Judith Lean[20] had also empirically investigated the linearity of the temperature response to forcings, including $CO_2$, but—for greater realism—she simultaneously examined the temperature response to several additional forcings, including solar and volcanic eruptions.[21] Another criticism of Figure 6.4A is that it relates the temperature at a given year to the $CO_2$ over the previous five years (a five-year smoothing was used for $CO_2$). Although this is (appropriately) longer than the atmosphere–ocean coupling time,[22] we would still expect that there is a longer time lag/memory in the system.[23] A simple way to take into account the lag between the forcing and the response is to compare the temperature at a given time with the forcing at an earlier time. It turns out that between zero lag (as in Fig. 6.4A) and twenty years, the relationship remains linear,[24] but the EffCS does change, leading to an overall estimate of 3.08 ± 0.95°C for a $CO_2$ doubling (90% confidence), which is close to—but less uncertain—than the IPCC ECS estimate of 3 ± 1.5°C. I won't pursue this issue further here. A better, but more complicated, approach uses a scaling response function and is discussed in Box 7.8.

With this caveat about possible lags, what has happened to the other anthropogenic forcings? Why is Figure 6.4A so straight when we know there are other

factors at work?[ee] Before answering this, consider the problem of reconstructing the historical record of changes in land use and of $CO_2$, $CH_4$, and aerosol emissions. Since 1880, the global mean $CO_2$ concentrations have been fairly well estimated, first from the analysis of air bubbles trapped in ice and, since 1957, from the famous Keeling curve [named after Charles Keeling (1928–2005)], which for the first time clearly showed the systematically increasing levels of atmospheric $CO_2$ taken from the pristine Mauna Loa and Antarctic observatories.[ff] However, the other main forcings—$CH_4$ and aerosols (particulate atmospheric pollution)—are much more poorly estimated.

Even with today's technology and understanding, the overall radiative effect of aerosols is particularly problematic. To start with, unless launched high into the stratosphere by volcanic eruptions or by high-flying military aircraft, their residence time—the typical time they stay in the atmosphere before falling to Earth or being washed out by rain—is on the order of weather scales, not even macroweather scales. They therefore tend to stay near their points of emission, so that their impacts are only regional. Some are famous for being responsible for the brownish atmospheric tint over India and China that is visible from outer space. However, even if the emissions are well estimated, they have nontrivial impacts on the radiative balance. For example, in the "direct" effect, black carbon aerosols absorb sunlight and cause heating, whereas the more prevalent (brownish) sulfate particles reflect light back into space and have a cooling influence (a negative forcing). However, aerosols also have an indirect effect: By seeding clouds, they increase cloudiness and this, in turn, reflects more solar radiation—and traps heat radiation—introducing difficult-to-quantify and conflicting cooling and heating effects. Even today, quantifying all of this is hard enough,[25] but trying to "reconstruct" its effects since 1880—even as only a global average—is very difficult (Box 5.2).

What saves Figure 6.4A is the crucial role of $CO_2$ in our society, in the economy. Its increase is hardly fortuitous. It is the result of the burning of fossil fuels that, today, accounts for 80% of all energy used. Human emissions of $CO_2$ are thus excellent *surrogates* for energy consumption and hence for economic activity. To a good approximation, double the economic activity, double the $CO_2$ forcing, double the $CH_4$ forcing, double the aerosol forcing, double the land-use forcing.[26] Figure 6.5A shows that this is indeed compatible with our knowledge of the global economy and with historical sulfate production, itself a surrogate for sulfate aerosols. Figure 6.5B shows that all greenhouse gases (especially $CH_4$), and even all anthropogenic forcings (including admittedly uncertain estimates of aerosol forcings), are fairly linear in $CO_2$ doublings.

---

[ee] Here we follow the standard procedure of reducing all the forcings to radiative equivalents. This is not always straightforward, especially for aerosols and volcanoes (see the discussion in Chapter 7).

[ff] It is often thought that, before 1957, there had been few direct atmospheric $CO_2$ measurements, but this is not true; more then 100,000 exist since the early nineteenth century. The major difficulty was that, before Keeling's spectroscopic method, measuring $CO_2$ concentrations required time-consuming chemical-based laboratory procedures. These measurements were typically near the ground and near urban areas, so they were badly corrupted by local $CO_2$ sources and by highly variable turbulence.
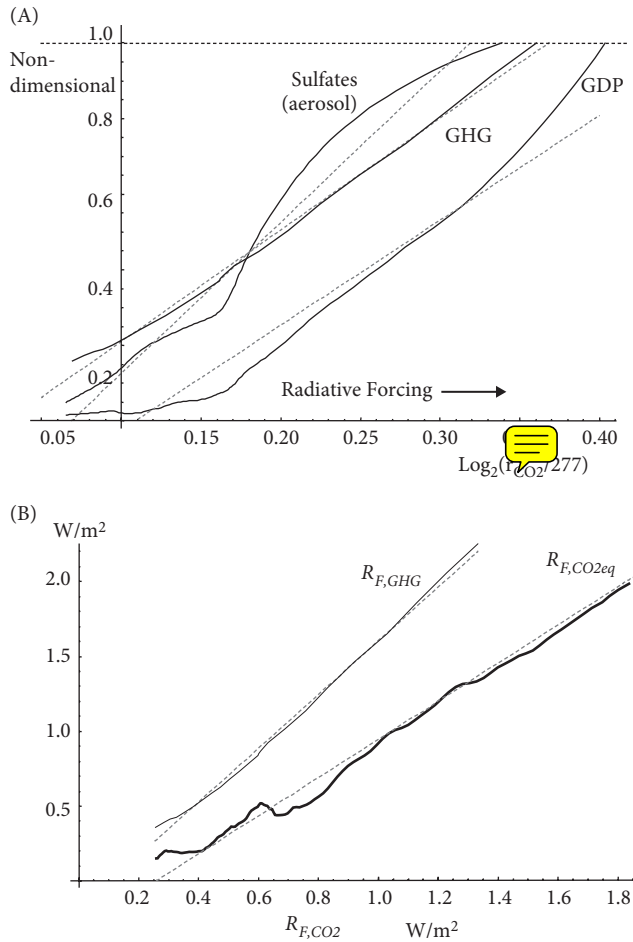
FIGURE 6.5 *(A) A comparison of the global gross domestic product (GDP), the total greenhouse gas (GHG) radiative forcings [the sum of carbon dioxide ($CO_2$) and other GHG forcings, mostly methane], and the global production of sulfates (a surrogate for sulfate aerosols). These were all normalized by dividing by their maximum values (the starting dates were all 1880, but the end dates were somewhat variable depending on the availability of data at the time the analysis was performed). Note that they were all plotted as functions of the number of $CO_2$ doublings. Although the lines are not perfectly straight, they do show the strong relationships between $CO_2$ doublings, the world economy, the total sulfate production, and cumulative GHG emission.[27] (B) The total GHG radiative forcing ($R_{F,GHG}$, top) and total anthropogenic radiative forcings ($R_{F,CO2eq}$; i.e., including estimates of the aerosol cooling effect) estimated as functions of the $CO_2$ radiative forcings ($R_{F,CO2}$), showing reasonable linearity.[28] The dashed lines are regressions, showing that all the GHGs effectively augment the $CO_2$ radiative forcings by 79%, whereas the lower line for the total $CO_2$ equivalent forcing, notably, includes estimates of aerosol cooling—effectively augmenting the $CO_2$ forcing by 12%.*

Figure 6.4A can therefore be interpreted in a quite different way, one that largely transforms a vice into a virtue and the simplistic into the *simple*: The $CO_2$ forcing is a surrogate for all the anthropogenic effects. In particular, it even takes into account indirectly the difficult-to-determine effects of aerosols. As long as the radiative effects of the aerosols—whatever they may be—remain a fixed fraction of the total forcing, then they are taken into account implicitly in Figure 6.4A. As we will see in Chapter 7, knowledge of the $CO_2$ concentration and past temperature data is enough to make accurate predictions and projections of future temperatures (see Plates 7.5 and 7.6).

## 6.4  Why the warming can't be natural

Figure 6.4B shows that, since 1880, we've had about 1°C of warming. This is close to the IPCC AR5 estimate of 0.85 ± 0.20°C.[29] What is the probability that a warming of this amount will occur in roughly 125 years?[30] Thinking in terms of a medical analogy, we can formulate this question in precise terms to apply statistical hypothesis testing techniques. In these terms, the null hypothesis is that this 125-year, 1°C temperature rise occurred purely naturally—that it is a GNF. If the probability (the "*p* value") of the GNF hypothesis is sufficiently low, then it can be rejected as being too unlikely.

Because we suspect that the recent period is not at all representative of the natural (preindustrial) variability, the only way to answer this question is to deduce the probability distribution of centennial-length temperature changes using preindustrial data: the multiproxies discussed in Box 5.3. We could note that for the preindustrial multiproxies, 125 years is apparently still within the macroweather regime (Fig. 5.10A), so that fluctuations are still tending to cancel, and a key characteristic is their fluctuation exponent *H*.

Figure 6.6 shows the result when multiproxies were used to estimate the northern hemisphere temperatures from 1500 to 1875. The series was broken into three typical preindustrial 125-year periods, each displaced for clarity. The mean of each was removed so that the periods vary about the indicated horizontal lines, with the typical changes over 125-year periods (0.2°C) indicated for reference. Above the three, at the top, is the industrial epoch global temperature since 1880, with the original 125-year length slightly extended to bring it more up to date (the same series was used as in Fig. 6.4A). We can see that this industrial epoch series is totally different from the preindustrial segments, showing an overall change of about 1°C over the same period of time. Also shown below it are the residuals obtained by removing the $CO_2$ contribution (reproduced from Fig 6.4C). As expected, the residuals appear (visually) to be very similar to the preindustrial multiproxies. What is the probability that the top series is just a random variation? A GNF?
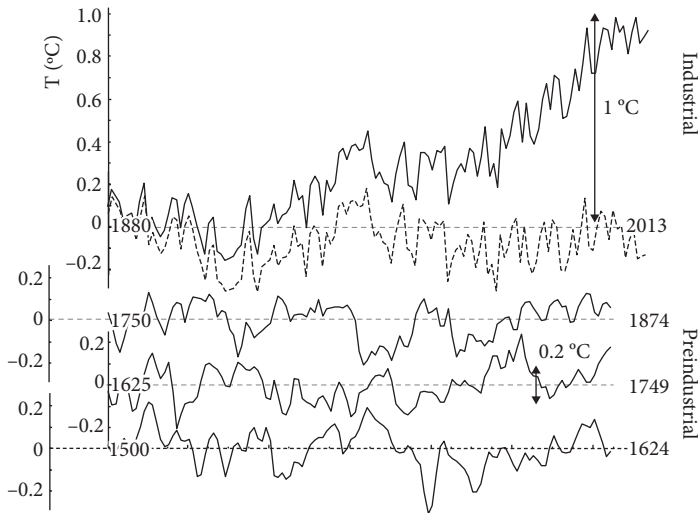
FIGURE 6.6 *The bottom three series (each displaced by 0.3°C for clarity) are the average of three post-2003 multiproxies (see Box 5.3, Fig. 5.15). They show preindustrial northern hemisphere temperatures; each segment has its mean removed so that they vary about the flat reference lines. The double-headed arrows show typical temperature changes (0.2°C) over 125 years. The top (dashed) curve is the globally averaged industrial epoch instrument temperature series; the top double-headed arrow shows the estimated anthropogenic warming from 1880 to 2013. The line below (fourth from the bottom) shows the industrial epoch macroweather (residuals) as in Figure 6.5C. They are quite similar to the preindustrial series.*

From Figure 6.6, we see that the conventional estimate of 1°C change during this period is quite plausible. If this change was simply a GNF with a bell curve probability distribution, then estimating its probability of occurrence would be easy. We would simply measure the standard deviation of preindustrial 125-year temperature differences—about 0.2°C (see the double-headed arrows in the figure[31]) and would find that the increase of 1°C is a five-standard deviations event (1/0.2 = 5). The probability of this is about 1 in 3,000,000. If this was particle physics, we would have just discovered the Higgs![gg] To put this probability in human terms, it is nearly the probability of a person being struck and killed by lightning in a given year. For most of us, this is an eventuality so remote that we don't lose sleep over it.

The assumption that temperature fluctuations follow the bell curve is conventional, but is it warranted? According to our previous discussion, a generic

---

[gg] The convention in particle physics is that an experiment purporting to discover a new particle is not considered to be statistically significant unless it occurs at this ridiculously high level of confidence. Usually the level is set at two standard deviations (i.e., 95% confidence). The IPCC sets it at 90% confidence (corresponding to 1.6 standard deviations).

consequence of scaling is the existence of power law probability distributions, associated with black swan extreme events (Box 3.1). What about the distribution of 125-year temperature fluctuations? Using three post-2003 multiproxies,[hh] we found that the extreme 3% of the temperature differences diverged from the bell curve (Fig. 3.1C). Indeed, as shown in Figure 6.7, we were able to bound the extreme multiproxy temperature changes between two power laws: one more extreme and one less extreme.
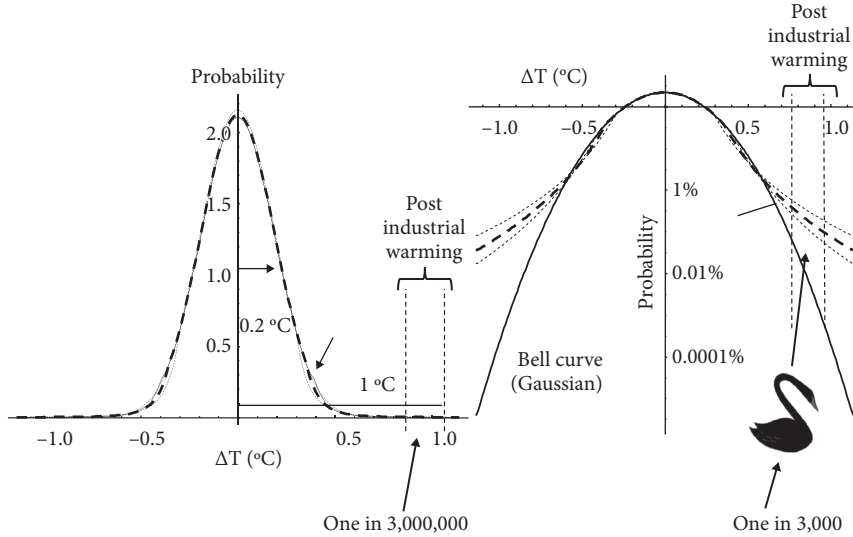


FIGURE 6.7  *A comparison of the bell curve distribution (solid black line, $q_D$ = infinity) of 125-year temperature changes with those with extreme power law tails (black swans; bottom, $q_D$ = 6; top, $q_D$ = 4[ii]). The actual multiproxy data are closer to the dashed black line ($q_D$ = 5, between the two; see Fig. 3.1C). The dashed vertical lines bracket the temperature change over the industrial epoch. Because the distributions are only significantly different for extremes with very small probabilities (less than about 3%), the usual presentation (left) doesn't distinguish them clearly. This is remedied on the right by the use of a logarithmic probability scale. The graphs show the probability densities; for statistical testing, it is the cumulative probability needed, and these are even more distinct (roughly 1 in 3,000,000 vs. 1 in 3,000).*

The basic power law states that the probabilities[32] of exceeding a given temperature threshold $s$ decrease as $s^{-q_D}$, where the exponent $q_D$ (see Box 3.1) quantifies the probabilities of extreme events.[33] The temperature probabilities were close to $q_D$ = 5, but in all cases they were found to be bounded between $q_D$ = 4 and $q_D$ = 6 (respectively, more and less extreme; see Fig. 3.1C). Although the probabilities of

---

[hh] It turns out that, as indicated in Box 5.3, Figure 5.14, all the multiproxies agree pretty well about the fluctuations up to 150 years, or the choice of multiproxies is thus not very important.

[ii] 1 + $q_D$ is the absolute slope of the tail on the extreme right-hand tail in Figure 6.7 (see Box 3.1).

the most common (nonextreme) 97% of the fluctuations are nearly identical to those of the bell curve, Figure 6.7 shows that the existence of black swan extremes (with $q_D = 5$) increases the probability of a 1°C GNF by a factor of 1,000, yet it is still less than 1 in 1,000. It is thus much more probable than being killed by lightening (1 in 3 million), but it is nevertheless not much more likely than being killed (in a given year) by poisoning.[34] This means the hypothesis that 1°C of warming could occur as a giant centennial-scale natural fluctuation can be rejected with more than 99.9% confidence[35]: It can be dismissed easily.

The fact that the difference between a standard bell curve and a black swan distribution might only affect the extreme 3% but still yield a 1,000-fold amplification in the probability of 1°C GNFs shows that adequately accounting for extremes can be a rather subtle affair. The situation is illustrated in Figure 6.8, in which two series of 5,000 points are compared, with each point representing the temperature change over 125-year periods. The bottom of the two uses conventional bell curve variables (corresponding to $q_D = $ infinity), whereas the top curve corresponds to $q_D = 5$. To bring out the point about the subtle difference associated with the extremes, the random temperature changes are chosen independent from each other.[36] The two standard deviation limits are indicated; this is roughly the limit where the two different probability distributions start to diverge. Because there are 5,000 points, we can easily estimate the values of the most extreme temperature changes that we expect to find. They should have probabilities of about[37] one in 5,000. For the top series with $q_D = 5$, this is about 1°C; for the bottom, it is about 0.71°C. Although these values seem close, the conventional bell curve probabilities fall off so quickly that it would (typically!) take a bell curve series nearly 1,000 times longer to show a 1°C GNF.[38]



Black swan probability: 1 in 5,000
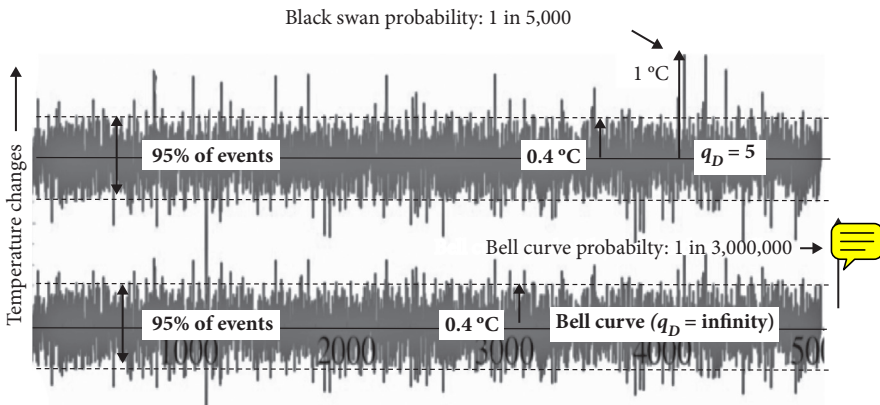
FIGURE 6.8  *(Bottom) A comparison of 5,000 randomly, independently chosen 125-year temperature changes. The bottom uses Gaussian (bell curve) variables, with 95% of the values within 0.4°C of 0 ($q_D = $ infinity). (Top) A more realistic situation in which the 5,000 independent values have the same fraction within 0.4°C, but with much more pronounced extremes ($q_D = 5$), including a 1°C giant natural fluctuation (with a 1 in 5,000 probability).*

## 6.5 What was shown

The original 2014 CD article went on to be the most highly "mentioned" *Climate Dynamics* article of the year.[jj] There was a deluge of email, some from colleagues, but mostly from the lay public, and this included legitimate requests for clarification as well as attacks by the organized deniers.

Because I couldn't answer all the queries, I put a Frequently Asked Questions tab on my website. I'll now summarize the more interesting points and misconceptions. The first part of the article was devoted to estimating the industrial epoch change (roughly the 1°C number mentioned earlier), but this was not necessary for the main part, which was on statistical hypothesis testing. The testing methodology was valid for any hypothesized century-scale temperature change, no matter how it was estimated. For this part, by far the most common misunderstandings were of the role of scale both in space and in time. As an example of the timescale error, I received numerous reminders that the temperature had been much higher in the past, and this purportedly showed that the industrial epoch warming could easily be a GNF; the medieval warming event was a commonly cited example. However, even if, 1,000 years ago, Europe's population enjoyed the same high temperatures as we do today, the implied high probability of 1°C millennial-scale temperature changes is irrelevant to the corresponding probability of the same warming in only 125 years: The rapidity of the warming—its timescale—is fundamental. This example also illustrates the other scale error: in space. It seems more and more likely that the medieval warming event was confined to Europe—that it was only regional in scope—that it was not a *global* warming event.

To make the point about spatial scale sharper, consider a criticism that the "Friends" used in their own press release in April 2014.[kk]

They noticed that between 1663 and 1762, the temperature in central England[39] increased by nearly 1°C (i.e., about the same as the global temperature during the past century) and concluded that such temperature changes were commonplace—that they are not at all rare. Unfortunately for them, central England (Greater London) is only about 0.01% of Earth's surface area, so it is not surprising that, during the same period, the northern hemisphere temperature only increased by 0.2°C, equal to its typical centennial change!

---

[jj] ~~Out of 750.~~ This is a category that weights an article's exposure in traditional news media as well as in blog postings, tweets, and other social media.

[kk] *Lovejoy Global Warming Paper 100% Wrong to Omit Previous Natural Warm Periods Say Friends of Science, Proving Natural Factors Affect Climate More than Humans*, http://www.prweb.com/releases/2014/04/prweb11767118.htm.

## 6.6  The pause

Basing itself on the hockey stick, the IPCC's AR3 (2001) declared that 1998 was the "warmest year of the millennium." At this point, the scientific case for the anthropogenic provenance of warming was already quite firm. Although estimates of the sensitivity of the climate to doubling $CO_2$ were still in the same wide range as in 1979 (1.5–4.5°C), this uncertainty was considered to be a purely technical issue to be solved with improved GCM resolutions expected with future supercomputers. The only significant exception was a problem that had first been raised in 1996[ll]: the apparent failure of satellite infrared atmospheric (not surface) temperatures to show the expected warming. But by 2006, the last of four flaws in this work had finally been laid bare. With the correct treatment of errors and biases (several were subtle), the satellite data ended up fully supporting the warming evident in the surface measurements. It was at this time that Michael Shermer, the influential editor of the *Sceptical Inquirer*, changed his position, saying that as a result of the advances in climate science, continuing to oppose anthropogenic warming on the grounds of skepticism would be unscientific.

By 2007, with the publication of the IPCC's AR4, the scientific case for anthropogenic warming had never been stronger, and the report was able to conclude that "it is very likely that human influence has been the dominant cause of the observed warming since the mid-20th century" (p. 4). Ironically, it was precisely at this time that—in apparent contradiction—the global temperature appeared to be leveling off (Fig. 6.4A),[mm] so that a clear understanding of space–time macroweather–climate variability was urgently needed.[nn]

By the end of the 2000s, it started to be clear that decadal GCMs temperature projections had been a little too high—by about 0.1 to 0.2°C (Fig. 6.9, although note the large model-to-model spread). The skeptics seized on the poor forecast and this

---

[ll] The chief protagonists in this affair were John Christy and Roy Spencer, who collaborated in the analysis. Both were upfront about their religion, with the latter signing the Cornwall Alliance's "Evangelical Declaration on Global Warming." This statement of scripture-based climate skepticism notably affi:

"Earth and its ecosystems—created by God's intelligent design and infinite power and sustained by His faithful providence—are robust, resilient, self-regulating, and self-correcting.

. . . We deny that Earth and its ecosystems are the fragile and unstable products of chance, and particularly that Earth's climate system is vulnerable to dangerous alteration because of minuscule changes in atmospheric chemistry. Recent warming was neither abnormally large nor abnormally rapid. There is no convincing scientific evidence that human contribution to greenhouse gases is causing dangerous global warming." See https://cornwallalliance.org/2009/05/evangelical-declaration-on-global-warming/.

[mm] In around 2007, the mainstream press started to pick this up. See, for example, Has global warming stopped? Whitehouse, D. *New Statesman,* December 19, 2007. http://68.171.46.155/WVFossils/Reference_Docs/has_GW_stopped.pdf.

[nn] Various articles soon appeared demonstrating that similar "pauses" were expected to occur simply as a result of natural cooling fluctuations temporarily counteracting long-term warming. But, these were based only on perusing GCM outputs, not on a systematic theory or understanding of the variability. See, for example, Hawkins, E. Our evolving climate: Communicating the effects of climate variability. *Weather* **66**, 175–179 (2011).

apparent "slow down,"[40] "hiatus,"[41] "pause"[42,oo] to proclaim joyfully that the models were useless and that the warming had ended. By 2013, articles started to appear that explained the GCM overforecast,[43] but by then—at least for the public—the damage had been done. The skeptics milked this issue for all it was worth and, as late as 2014—before the spectacular 2016 record-high global temperature that ended any pretense of a pause—the skeptics were still loudly proclaiming: "Global warming stopped naturally? 16+ years ago" (Fig. 6.1, lower right). In the absence of a clear understanding of the space–time statistics, the IPCC could only resort to the vague explanation: "In addition to robust multi-decadal warming, global mean surface temperature exhibits substantial decadal and interannual variability. Due to natural variability, trends based on short records are very sensitive to the beginning and end dates and do not in general reflect long-term climate trends."[44] Ironically, Box 7.7 shows that, with only knowledge of the $CO_2$ concentration and past temperatures, the temperatures could be forecast (starting in 1909!) through the 2000s with high accuracy (to within 0.2°C).
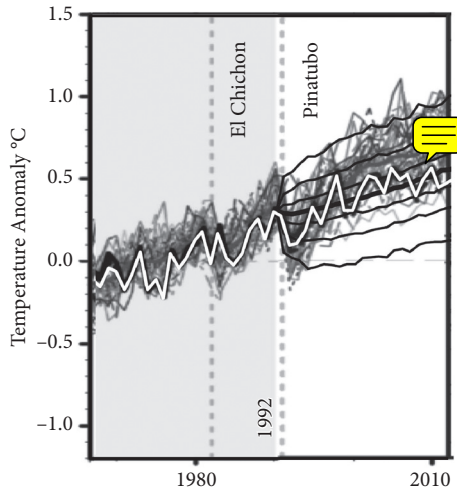


FIGURE 6.9  *A comparison of individual Global Climate Model (GCM) runs (from the Climate Model Intercomparison Project 3 models: the cloud of background lines). These runs were "harmonized" to start from the same value in 1990.[45] The GCM average (the "multimodel mean") is shown in thick red. The yellow line is the National Aeronautics and Space Administration GISS series used in Figure 6.4A. This is compared to the scaling-based hindcast from 1992 using the Scaling Macroweather Model (SLIMM, Chapter 7).[46] The black line is the hindcast/forecast mean (50% probability of exceedance), the blue, purple, and red lines are for 20%, 40%, and 49% above and below the mean (see the details in Chapter 7). The 1998 El Niño year marking the beginning of the pause corresponds to the high yellow peak. The GCMs tend to be 0.1 to 0.2°C too high (thick red) and are near the 90% mark for the SLIMM hindcast (thick black), which is quite accurate. The dashed vertical lines indicate the times of two major volcanic eruptions.*

---

oo There is even a debate about the most appropriate term to use.

With the aid of the previous sections—and especially Figure 6.4A, C—the basic reason for the pause is pretty obvious. Whether we follow the actual temperature (Fig. 6.4A) or the residuals representing the macroweather (natural) variability (Fig. 6.4C), we can see that the temperature was set for a return to the long-term anthropogenically induced climate trend (Fig. 6.4A). Alternatively—from the point of view of macroweather—we expect a canceling downward swing that would partially cancel the ongoing upward anthropogenic (climate) trend. Either way, it didn't take sophisticated analysis to see that we could expect a relatively long period of fairly stable temperatures while the downward macroweather fluctuations were canceling the continued anthropogenic increase. More quantitatively, Figure 6.4A, C shows that, between 1992 and 1998 (the "prepause"), the global mean temperature rocketed up by about 0.42°C, well above the long-term anthropogenic trend. Indeed, 1998 marked a particularly strong El Niño event, which are generally associated with large macroweather temperature upswings. Removing the superposed anthropogenic warming trend (Fig. 6.4C), we see that the macroweather (natural) warming was about 0.33°C superposed on an anthropogenic warming climate trend of 0.09°C. In comparison, during the 1998 to 2013 pause period, the overall temperature decreased by 0.01°C, a consequence of a macroweather cooling fluctuation of –0.28°C, which more than offset an anthropogenic climate warming of 0.27°C.[47] See Box 6.2 for more details.

---

ʙᴏx 6.2  Return periods and the pause

Several mechanistic explanations have been proposed for the overly warm post-1998 decadal GCM forecasts[48] (Fig. 6.9); but, regardless of the reason for this GCM problem, we can revisit the past two decades with our understanding of macroweather statistics. We can now improve on qualitative interpretations of Figure 6.4A, C. If the pause was really just a decadal-scale macroweather cooling fluctuation that largely offset the post-1998 anthropogenic warming, then to be convincing, we should be able to quantify it. What was the probability of this rough cancelation? If the probability was too low, this explanation must be rejected.

This was the motivation for the CD follow-up article: "Return Periods of Global Climate Fluctuations and the Pause"[49],[pp] (denoted GRL1[qq]). Only a few months after the CD article was published, this new publication shed a different light on the prevailing sociology and politics of science. The new article used the same basic approach as the CD article and provided the first estimates of how long one would expect to wait to observe natural global temperature fluctuations of various amplitudes over various time periods. It concluded that a cooling of the amplitude necessary to explain the pause occurs naturally every twenty to fifty years so that,

---

although the cooling was indeed strong, it was not highly unusual, especially because we saw that it immediately followed an even larger natural "prepause" warming event.

Figure 6.10 shows a simplified graph of the return periods for natural fluctuations of different amplitudes adapted from GRL1. The return period of an event is the typical time one must wait before the event repeats itself (here, between consecutive, large global annual temperature changes, with amplitudes given on the horizontal axis). For extreme events, it is usually assumed that successive extremes are uncorrelated, so the event return time can be estimated as the inverse probability of its occurrence. Figure 6.10 was therefore readily determined from the same preindustrial multiproxy probabilities that were used to test the GNF hypothesis in the CD article.[50]
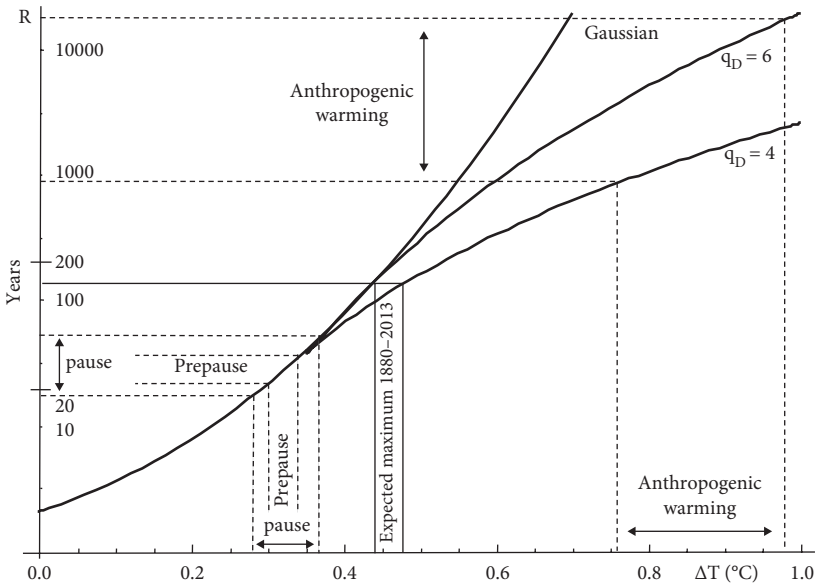


FIGURE 6.10 *The return periods for fluctuations of the amplitude are indicated on the horizontal axis. The curves are the empirical curves deduced from the preindustrial multiproxies. The plot shows the results for the bell curve (Gaussian, $q_D$ = infinity) and the bounding power law extremes curves (black swan, $q_D$ = 4, 6). The dashed vertical lines correspond to various events, from right to left: global warming since 1880 (range, 0.76–0.98°C), the largest event expected in the 134 years since 1880 (0.47°C), the postwar cooling (0.42–0.47°C, the prepause of 1992 to 1998 (0.30–0.33°C), and the pause of 1998 to 2013 (0.28–0.37°C). The horizontal lines indicate the corresponding return periods.*

According to Figure 6.10, anthropogenic warming (1880–2004, estimated as 0.76–0.98°C, shown by the dashed lines to the right) has a return period of 1,000 to 20,000 years (using the bounding distributions with exponents $q_D$ = 4, 6). Although this is a sufficiently long period that natural variability can confidently be rejected as an explanation for the warming, it is nevertheless much shorter than the ≈3,000,000-year return period obtained using the classic (Gaussian) assumption (the top line).

What is the largest fluctuation that we should expect over the period 1880 to 2013? Such an event would have a return period of 134 years; hence, according to Figure 6.10, an amplitude of ≈0.47°C. Comparing this estimate with Figure 6.5A, C, we see that—as expected—it is comparable to the postwar (1944–1976) cooling event of 0.42 to 0.47°C. Turning to the pause (0.28–0.37°C), from Figure 6.10 we see that the return period for such an event is twenty to fifty years. Although in themselves such cooling events are not unusual, they become altogether probable when they immediately follow comparable warming events. Because in this macroweather regime successive fluctuations tend to cancel, this is already a statistical explanation for the pause.[51]

GRL1 showed that the probability of a 0.28°C macroweather cooling over sixteen years is about 2% to 5% (corresponding to fifty- to twenty-year return periods), which was enough to explain the pause as a relatively common occurrence. But this probability was "unconditional." It did not take explicitly into account the preceding "prepause" temperature data. It applied to a macroweather cooling of 0.28°C from a date chosen at random, regardless of the temperature in the preceding years. Yet inspection of Figure 6.4A shows that it followed an even larger prepause macroweather warming. To quantify the probability of the pause properly, we should take into account this prior information. We should use *conditional* probabilities [i.e., conditioned on the actual (historical) record]. This is equivalent to making a stochastic (statistical) forecast; it was the subject of another article, GRL2,[52] that appeared a year later (and is discussed in Chapter 7). In GRL2, the pause was forecast ("hindcast") retroactively over the decade following 1998 to within better than ±0.05°C. This vindicated both the macroweather forecast procedure[53] and also the interpretation of the "slowdown" as "just a pause."

## 6.7  The $100,000 GNF

By showing quantitatively that the pause was natural, the article "Return Periods of Global Climate Fluctuations and the Pause" (as mentioned earlier, referred to as GRL1[54]), demolished the remaining skeptic argument that warming had stopped in 1998. The article achieved media prominence and was notably featured on the front page of Montreal's daily English language newspaper with the head-line: "Global Warming slowdown just a 'pause.'"[rr]

At this point (July 2014), I was trying to promote some of the simpler graphs (e.g., Fig. 6.4A) as educational tools to help convince open-minded skeptics of

[rr] Seidman, K. Believe it: Global warming is real, global warming slowdown is just a "pause." *The Montreal Gazette*, p. 1, July 24, 2014, http://www.physics.mcgill.ca/~gang/popular.articles/Gazette.7.14/ Gazette.pause.all.final.jpg. This was followed by an Op-Ed in the *Gazette* in August by "The Friends," exposing their speculation about solar causation, accompanied with an ad hominem attack: *The Montreal Gazette*, August 28, 2014, http://montrealgazette.com/opinion/editorials/opinion-global-warming-pause-is-more-than-temporary-scientific-evidence-shows). It was signed: "Len Maier is an engineer and president of the Friends of Science Society in Calgary."

the reality of anthropogenic warming. In many countries—notably, Canada, the United States, Britain, and Australia—the skeptics still included substantial proportions of the general public, and I considered that battling the organized skeptics was important. I contacted colleagues particularly familiar with this kind of hand-to-hand combat, especially Michael Mann of hockey stick fame (Box 5.3). Mann himself was not shy of a fight and, ever since the publication of the hockey stick, he had been the main skeptic whipping boy.[ss]

  Mann was both helpful and encouraging, and put me in contact with a US-based nongovernmental organization (NGO) that was committed to bringing media attention to science that was deemed important in the fight against anthropogenic warming. Here I was in for a disappointment. Following Mann's advice, I forwarded the article, press release, and supporting material, hoping to get help in spreading the word. At first, I received a polite response saying they would consider it. After a few weeks, when I hadn't heard any answer, I prodded them with another email. This time, I received a laconic reply that, in their view, it was best to ignore the pause: Any mention of it—even to show that it was no more than a natural fluctuation—would give it undue publicity and make it appear that there was a debate where (according to them) there wasn't any. By this time, the environmentalists had become supremely confident and believed the skeptics were better ignored than confronted.

  Over the next year, I began using macroweather laws for monthly, seasonal, and annual forecasts (Chapter 7) and, in July 2015, in a new article, "Using Scaling for Macroweather Forecasting Including the Pause"[55] (hereafter GRL2), I showed how they could be used to make a decadal forecast of the pause with unprecedented accuracy. With this addition, a certain closure had been reached. The high-level laws had made the macroweather–climate distinction transparent, and had quantified the anthropogenic climate regime drivers. The understanding of the macroweather scaling laws—including their extremes—had allowed the industrial epoch GNF warming hypothesis to be rejected statistically while simultaneously *failing* to reject the natural cooling hypothesis for the (much smaller) post-1998 macroweather fluctuation responsible for the pause. Finally, GRL2 showed that macroweather laws combined with historical temperature data could be used to *forecast* the pause accurately (even from 1909; see Box 7.7). Together, these articles had eliminated nonanthropogenic explanations systematically, leaving the anthropogenic theory as the only viable explanation. The climate debate had reached scientific closure.

  On the media side, GRL2 confirmed there was a frustratingly low level of interest in the scientific aspects of the debate. Meanwhile, the deniers continued to spread their disinformation. Yet the closure synthesis was surely a milestone in the fight against the skeptics; it reduced them to mere "denialists."[tt] To bring

---

[ss] See his own compelling account: Mann, M. E. *The Hockey Stick and the Climate Wars: Dispatches from the Front Lines*. (Columbia University Press, 2012).

[tt] The term "skeptic" was too inaccurate—and too generous—to apply to a group that was by now simply following an obstructionist agenda that had nothing to do with genuine skepticism. The term "denier"—although used in other branches of science—evoked "holocaust denier," which had obvious

this closure to the attention of the wider community, in October 2015 I published "Climate Closure" in the AGU's (wide-distribution) membership journal *EOS*.[56] Although there was little media pickup, the piece generated more than 1,000 comments in various blogs. To kill the generally positive exchanges on the *EOS* site, Monkton himself was forced to intervene via an extensive trolling operation.[uu]

The confidence of the environmental NGOs grew even stronger when, in December 2015, 196 countries signed an agreement at the Conference of the Parties (COP)[vv] 21 meeting in Paris. Although the final document had numerous weaknesses—including its strictly nonbinding nature[ww]—after the previous failures, the mere fact of *any* agreement was an achievement. The feeling was that a historic turning point had finally been reached and, in Paris, there was palpable euphoria. The upbeat ambiance was relayed by the media, who almost totally ignored the deniers' well-funded presence.

For the scientific community, the agreement was bittersweet. For years, governments had refused to move, effectively gaining time by proclaiming the need for more climate research. In an attempt to convince governments to ignore the skeptics and, finally, to act, the scientific community had reacted by claiming loudly that the science was "settled"—that a scientific consensus existed on anthropogenic warming. Now that the world's governments had finally agreed collectively to phase out emissions, the scientists had suddenly become superfluous.[xx] Shortly after COP 21 (February 2016), this was more or the less the position of Australian government when they axed the Commonwealth Scientific and Industrial Research Organisation's (CSIRO's) climate division.[yy] In a gleeful blog, Australian climate denier Jo Nova captured the moment:

> CSIRO has announced it will axe 300 to 350 climate jobs, which will "wipe out" the climate division. The head of the CSIRO wants to focus on climate adaption and mitigation instead. Suddenly a lot of Profs who told us the debate was over are squealing that it needs more research. Climate science was "beyond debate" and in need of action, but now we "need to know more about the basic operation of the climate."[57]

negative connotations. The term "denialist" was an unsuccessful attempt at a neutral compromise. See Gillis, J. Verbal warming: Labels in the climate debate, *New York Times*, p. D1, February 17, 2015.

[uu] This involved several multipage interventions focused on my use of the term "denialist" to qualify his group. Although elsewhere Monkton refers routinely to scientists as the "climate communist faction," in his trolling he moaned (on and on and on) about how unfair and how unjust it was to be termed a "denialist" when in actual fact he was really just a sincere skeptic.

[vv] Part of the IPCC process, under the aegis of the United Nations.

[ww] Even if its nonbinding promises are respected—and notwithstanding the proclaimed aim at holding the industrial epoch temperature increase to 1.5°C (only 0.5°C more than what we already have)—the actual temperature of Earth would greatly surpass the previously proclaimed (COP 16, 2010) threshold of 2°C. It would likely exceed 3.5°C by the year 2100.

[xx] Adding to this was the fact both the 2010 COP 16 decision on the 2°C target and the (virtually unattainable) 1.5°C COP 21 target were determined by economic and political grounds—not by scientific considerations.

[yy] CSIRO is the federal government agency for scientific research in Australia.

The Australian reaction was admittedly extreme, but everywhere[zz] the ruling austerity policies had already led to frozen or receding research budgets, and even in Obama's United States, climate scientists were more exposed than most.

This was the situation when, on May 7, 2016, I received a personal invitation by British climate skeptic and professional statistician Douglas Keenan to make a submission to his "climate contest" with a $100,000 prize. Keenan kindly waived the $10 entry fee. Although I had been unaware of it, the contest had been launched in November 2015 and was to remain open until November 30, 2016.[aaa] Keenan's up-front motivation for sponsoring the contest was to bring public attention to his criticism of the IPCC and to his claim that the industrial epoch warming might have been no more than a GNF. The contest was widely promoted by the climate skeptic community and was used to publicize the more sophisticated skeptics' statistical arguments against anthropogenic global warming.

I was interested in the contest not only as an intellectual brainteaser, but for other reasons—the foremost of which was that it was the first time climate skeptics had made an explicit GNF model. For decades, skeptics had claimed that global warming was no more than a GNF, yet they were impossible to nail down. They never gave a precise model that might make their claim more than empty-phrase mongering. The contest offered a golden opportunity to invalidate their GNF claim. Another reason for my interest was that I was convinced that Keenan's model involved processes with long-range memories—scaling and the high-level laws—and this was an opportunity to highlight their impact on the statistical analysis of climate series and to improve IPCC uncertainty estimates.

Shortly after the invitation, I discovered a long-forgotten missive from Keenan dated April 12, 2014, less than twelve hours after Monkton's "emanation of the forces of darkness" assessment of the CD article. At the time, Keenan had just submitted a critique of the IPCC's trend analysis to the United Kingdom's House of Lords and—like Monckton—was also interested in combatting the dark forces. In his email to me, he stated:

> My critique and your paper "Stochastic and Scaling Climate Sensitivities" [*Climate Dynamics*, 2014] are obviously in conflict. I have looked at your paper, though not studied it closely. I do not understand some of the calculations. Of particular interest are calculations of the confidence intervals for correlations. Each climatic time series exhibits many autocorrelations: how are the autocorrelations handled, in the calculations?

Because (as explained later), the CD article did not use trend or autocorrelation analysis to estimate the uncertainties of any temperature time series trends[58]—and

---

[zz] This excludes several Asian countries that were able to ignore the International Monetary Fund's austerity strictures, including South Korea, Singapore, India, and, especially, China, with science budgets that continued their exponential growth.

[aaa] Except in the unlikely—in retrospect, impossible—event that the contest was won at an earlier date!

because at that moment I was being bombarded by hate mail—I did not pay close attention, giving the laconic reply: "I see no grounds for conflict with my paper."[bbb]

The contest involved 1,000 realizations[ccc] of a "stochastic model fit to the annually and globally averaged temperature series from 1880–2014."[ddd] The contest aim was to bring into question the IPCC statistical uncertainty assumptions and hence to undermine the empirical basis of the anthropogenic warming theory.[eee] In an accompanying explanatory blog, his basic logic seemed to be this: (a) trends were needed to establish anthropogenic changes, (b) a trend without an uncertainty is meaningless, (c) a model of the residuals is needed to estimate the uncertainty, and (d) the IPCC did not justify its choice of an uncertainty model. Keenan therefore concluded ipso facto that all the IPCC uncertainty estimates were meaningless, hence its trends were meaningless, and finally that its conclusions about anthropogenic effects were groundless.

Along with colleagues, we soon[fff] published "Giant Natural Fluctuation Models and Anthropogenic Warming"[59] (GRL3), the key conclusions of which were as follows:

1. Contrary to point (a), trends are not needed to test the GNF hypothesis statistically; this can be done directly with probabilities. Although points (b) and (c) may be valid, the contest therefore proves nothing about anthropogenic warming.
2. The IPCC assumed that, after removing trends, the residuals had only short-range correlations.[ggg] In contrast, as we had predicted, Keenan's model had long-range (power law, scaling) residuals. Taking these strong correlations into account allowed for tighter (less dispersed, *less uncertain*) trend estimates. Ironically, Keenan's model illustrated that, by adopting

[bbb] There is irony here. In the original explanation of the contest, Keenan stated: "I am sponsoring a contest: the prize is $100,000. In essence the prize *will be* awarded to anyone who can demonstrate, via statistical analysis, that the increase in global temperatures is probably not due to random natural variation" (italics added). However, several weeks after the start of the contest, he surreptitiously changed the terms to: "I am sponsoring a contest: the prize is $100,000. Anyone who can demonstrate, via statistical analysis, that the increase in global temperatures is probably not due to random variation *should be* able to win the contest" (italics added). According to the original ("will be")—but now modified ("should be") terms—the CD article would have won the contest.

[ccc] Keenan actually had two models: the basic (on average) trendless model $[T_{init}(t)]$, and another one $[T(t)]$ that consisted of randomly adding to $T_{init}(t)$ a trend of ±1°C/century. The aim of the contest was to identify which of the models had trends added: Nine hundred correct out of 1,000 were needed to win. It turned out that the basic model itself was a random mixture of four different submodels.

[ddd] In a later threatening email, he tried to distance himself from the idea that his model was intended to be realistic. Yet, as he clearly indicated, the model was "fit"—with considerable care and accuracy—to the data. It was indeed *statistically* realistic. The problem was that it was *scientifically* unrealistic, but Keenan is a statistician not a scientist.

[eee] Although Keenan was never completely explicit about this, the IPCC assumed that, after trend removal, the residuals had only short-range (exponential) correlations whereas in reality—and in Keenan's contest—the residuals had, on the contrary, long-range (power law, scaling) correlations.

[fff] We attempted a rapid publication in attempt to derail the contest by sharing our analyses publicly before the November 30, 2016, deadline.

[ggg] In AR5, there was a discussion of short-range versus long-range correlations and, indeed without much justification, the IPCC authors came down on the side of conventional (short-range) wisdom.

the short-range dependency assumption, the IPCC had *overestimated* the uncertainties in its trends![hhh]

3. Over the period 1880 to 2014, the variability of Keenan's model was quite realistic. Indeed, if the only information that existed about temperature variations were these 135 numbers (the globally, annually averaged temperatures), then his model could not be rejected. However, when his model was compared to preindustrial temperature data, its variability was far too large, predicting Ice Ages every 1,000 years or so rather than every 100,000 years (Fig. 6.11). His model could be easily rejected *scientifically*.
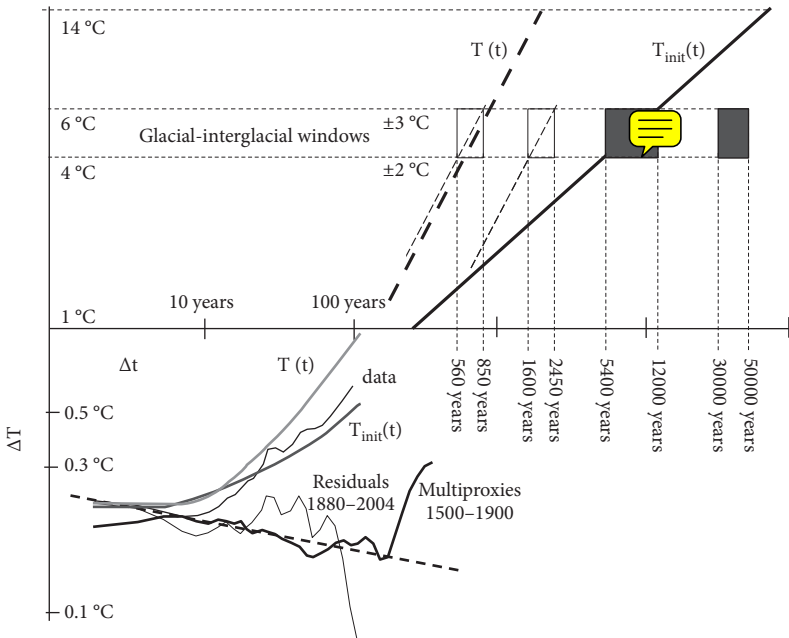


FIGURE 6.11 *The Haar fluctuation analysis of the Keenan's two models with and without the added random 1°C/century temperature trend [T(t), pink; T_init(t), brown]. The original GRL3 extrapolations were based on the assumption of a unique model (with no submodels), which is shown by a dashed black line. The red lines are the updates based on the final unveiled model. We can see that the T(t) model (i.e., with the randomly added trends) was well reproduced (dashed red and black lines on the upper left), but that the T_init(t) model was a little off (dashed black and solid red lines, right). In all cases, the glacial–interglacial window is shown (rectangles), indicating that the models have Ice Age variability far too often (the realistic window is shown in green, upper right). For comparison, the preindustrial multiproxies (from Fig. 5.14) are included as well as the fluctuation analysis of the residuals from Figure 6.4C, showing that even at much shorter timescales, the natural preindustrial macroweather variability is much lower than Keenan's model implies.*

---

[hhh] It's a bit puzzling that Keenan hadn't noticed this himself, because he must have known that the only way to win the contest was to use an alternative uncertainty model that specifically *reduced* the conventional IPCC uncertainties. It was only by reducing the random spread (uncertainty) in the trend estimates that the artificial addition (or subtraction) of 1°C/century trends could be detected.

Four months later, on November 30, 2016, Keenan announced that none of the thirty-three contest submissions were winners and he unveiled the computer code that generated the 1,000 series. His model turned out be actually quite bizarre: It consisted of a random shuffling of each of four rather different submodels, one of which was actually an IPCC numerical model output![60] Two of the other submodels were basically standard scaling models[61] whereas the third was a complicated homemade concoction that nevertheless was also scaling at long timescales. As predicted in GRL3, all the submodels (and hence the overall model) had strong (scaling, power law) statistical dependencies.[62]

After Keenan unveiled his model, we ran the published code but for 10,000 simulated years rather than the original 135, confirming with only minor differences the GRL3 conclusions (Fig. 6.11). For example, the temperature fluctuations of the model with added trends [$T(t)$] was predicted accurately (top, dashed line). Indeed, extending this model to Ice Age periods, it predicts a whopping ±200°C variation (a difference of 400°C between glacial and interglacial conditions, which is off-graph). In contrast, the basic model without the added trends [$T_{init}(t)$, ~~solid red~~ line] was a little less extreme than that predicted by GRL3. Figure 6.11 shows that it nevertheless predicts Ice Ages with half-period 5,000 to 10,000 years (i.e., about five times too frequently). The glacial–interglacial temperature variations (at 30,000–50,000 years for a half period) were typically about 14 ±7°C, which is about three times too high.

AQ: Please clarify reference to color.

After GRL3 was published, in the classic style of British aristocracy, Keenan threatened libel action against the *Geophysical Research Letters* editors and *Geophysical Research Letters'* parent organization (the AGU). Despite the fact that British libel law does not apply in the United States, the threat apparently succeeded in quashing a nonspecialist story planned for AGU's *EOS* magazine. Partly as a consequence, both the scientific community and the media were oblivious to the entire episode. This was a shame because, if analyzed properly, the patent failure of GNF models drives a final nail in the coffin of skeptics' attempts to find scientific alternatives to the theory of anthropogenic warming.

## Notes

1. Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014).

2. Lovejoy, S. Climate closure. *EOS* **96**, doi:10.1029/2015EO037499 (2015).

3. https://wattsupwiththat.com/2014/04/11/lovejoys-99-confidence-vs-measurement-uncertainty/.

4. This is the average over Earth's surface. Many regions—especially the Arctic—have warmed significantly more.

5. Callendar, G. The artificial production of carbon dioxide and its influence on temperature. *Q. J. Roy. Meteor. Soc.* **64** (275), 223–240 (1938).

6. These numbers are estimates of the ECS—the temperature rise that would occur if we waited long enough for equilibrium to be reached (Chapter 5). These early values are

close to the modern IPCC range of 1.5 to 4.5°C. Interestingly, Arrhenius himself later revised his own estimate downward by about 1°C: Arrhenius, S. Die vermutliche Ursache der Klimaschwanungen [The probable cause of climate fluctuations]. *Meddelanden fran K. Vetenskapsakademiens Nobelinstitut Band* **1** (2), 1–10 (1906).

7. Probably the very first climate skeptic argument—going back to the beginning of the twentieth century but still regularly repeated—claims that the $CO_2$ bands are "saturated." At saturated wavelengths, between the ground and outer space, the atmosphere is totally opaque: None of these surface photons can *directly* escape from Earth. This is, for example, generally the case for Venus, which has a huge radiative greenhouse effect (several hundred degrees). The skeptics then argue that because the atmosphere is opaque, further increases of $CO_2$ will change nothing. But both the premise and the conclusion of the argument are wrong. To start with, it turns out that, unlike Venus, most of the bands on Earth are *not* saturated. Yet even if they were, the consequences of an increase in $CO_2$ would be very much the same: The roof would still be raised, its temperature would still decrease, less heat would still be emitted, and Earth would still warm up.

8. Both atmospheric pressure and $CO_2$ concentrations fall off exponentially with altitude whereas temperature drops off linearly. This implies that raising the roof leads to a logarithmic dependence of the $CO_2$ forcing on the $CO_2$ concentration. But adding $CO_2$ not only raises the roof, it also promotes marginally sensitive $CO_2$ bands into bands that interact more and more strongly with $CO_2$. Both effects depend on the logarithm of the $CO_2$ concentration: Every $CO_2$ doubling increases the forcing by the same amount.

9. The IPCC was first established in 1988 by two United Nations organizations, the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP), and later endorsed by the United Nations General Assembly. It has several working groups; I frequently cite the regular ARs of the Physical Sciences working group (Working Group 1). The first five reports were published in 1990, 1996, 2001, 2007, and 2013; the sixth report is expected in 2022. The reports can be downloaded from www.ipcc.ch/.

10. Brulle, R. J. Institutionalizing delay: Foundation funding and the creation of U.S. climate change counter-movement organizations. *Climatic Change* **122**, 681–694 (2014).

11. For a discussion, see Oreskes, N. Playing dumb on climate change. *New York Times*, January 4, 2015, p. SR2.

12. At long-enough timescales, the albedo of Earth could change and this would complicate things. For example, the ice caps are melting, and this slowly darkens the planet as white snow and ice are replaced by dark ocean or land. This is a positive feedback that accelerates warming. The Daisyworld model discussed in Section 5.9 is an example of an analogous negative feedback mechanism.

13. About 2 W/m$^2$ greater than during the preindustrial epoch (i.e., an increase of about 1%; see Section 4.6).

14. Hebert, R. *A Scaling Model for the Forced Climate Variability in the Anthropocene*. MS thesis, McGill University (2017).

15. Hebert, "A Scaling Model."

16. The $CO_2$ concentrations were smoothed somewhat because only scales longer than the two-year atmosphere–ocean coupling scale are pertinent.

17. The relationship with the concentration is logarithmic (see Box 6.1).

18. The range indicated means that, with 90% certainty, the temperature increase with a $CO_2$ doubling is between 1.97°C and 2.69°C. It was deduced by considering the spread in

the value obtained by three different temperature series. For details, see Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014).

19. All views adapted from Lovejoy, S. Return periods of global climate fluctuations and the pause. *Geophys. Res. Lett.* **41**, 4704–4710 (2014).

20. Lean, J. L. & Rind, D. H. How natural and anthropogenic influences alter global and regional surface temperatures: 1889 to 2006. *Geophys. Res. Lett.* **35**, L18701, doi:10.1029/2008GL034864 (2008).

21. She used a *multiple*-regression technique.

22. The coupling time is the typical time that it takes for the ocean and atmosphere to act as a single system. For example, month-to-month fluctuations in global temperatures are virtually unrelated to the corresponding month-to-month fluctuations in global ocean temperatures. However, at periods of two years and longer, the fluctuations are strongly correlated so that the overall atmosphere–ocean system is increasingly coupled. See Figure 7 in Lovejoy, S., Del Rio Amador, L., & Hébert, R. Harnessing butterflies: Theory and practice of the Stochastic Seasonal to Interannual Prediction System (StocSIPS), In: *Nonlinear Advances in Geosciences* (ed. A. A. Tsonis), pp. 305–355. (Springer Nature, 2017).

23. The one- to two-year coupling timescale is the minimum time for which the response can be considered linear. In Chapter 7, we will see that a scaling relation is more realistic, and this is implies that, as a result of the long memory, even quite ancient forcings may still be felt.

24. This range of time lags was taken into account in Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014).

25. The usual way to describe forcings other than $CO_2$ itself is by their "$CO_2$ equivalent" ($CO_{2eq}$) forcings. For a given greenhouse gas or aerosol, the $CO_{2eq}$ is the equivalent amount of $CO_2$ needed to give the same radiative effects. $CO_{2eq}$ was not used here because the main difficulty remains: aerosols. For reasons explained in Chapter 7, $CO_{2eq}$ is used there instead. In any case, it was found (Fig. 6.5B), not surprisingly, that since 1880, $CO_2$ and $CO_{2eq}$ are indeed very highly correlated, so that, on average, $CO_{2eq}$ is about 12% more than $CO_2$.

26. Actually, the connection between the economy and $CO_2$ is between the annual economic activity and the annual rate of emissions, not the total atmospheric concentration. However, during the industrial epoch, economic activity has grown roughly exponentially, and high school math shows that the rate of change of an exponential is also an exponential. Hence, the logarithm of either the emissions or the cumulative emissions is proportional to the forcing.

27. The figure was reproduced with permission from Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014). It includes more details.

28. This was the best data available at the time of the analysis, which slightly predated IPCC AR5 (2013) with its improved historical reconstructions (see Chapter 7).] The figure is an updated version of one in Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014). The greenhouse gas relation was deduced from data in Myhre, G., Myhre, A., & Stordal, F. Historical evolution of radiative forcing of climate. *Atmos. Environ.* **35**, 2361–2373 (2001). The $CO_{2eq}$ data are from IPCC AR5.

29. From 1880 to 2012, with 90% confidence.

30. One hundred twenty-five years was used because of the availability of the data used in the original study. An update will not alter the conclusions very much.

31. One standard deviation corresponds to 68% of the fluctuations being less extreme. Also, the estimate 0.2 is for differences in temperatures over 125-year periods, such as those shown in Figure 6.6. Over shorter lags, a little lower.

32. For probabilities, the power law is only expected to apply to the extremes. Here, the extreme is 3%.

33. As $q_D$ increases, the shape of the distribution changes and the transition point to extreme power law behavior moves to lower and lower probability levels. We could thus think of the bell curve as effectively having $q_D$ = infinity, meaning a transition never occurs.

34. Statistics from the United States.

35. $p < 0.001$. In the CD article, the $p$ values for several slightly different hypotheses were estimated. They took into account the uncertainty in the amount of warming (it isn't exactly 1°C) and also the uncertainty in the nature of the extreme probability "tails." Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014).

36. These series are displayed to make the point about the nature of the probability extremes. They are not intended to be taken as models of time series of temperature changes, because successive real-world temperature changes over 125-year periods are correlated. These correlations were taken into account in the CD paper. Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014).

37. We say "about" because this is the probability, averaged over an infinite number of statistically identical series, each with 5,000 points. Of course, each of the individual series would have different maxima.

38. If we made 1,000 such series each with 5,000 points, the one with the most extreme maxima would show a 1°C GNF.

39. This result can be deduced from the famous Manley monthly climate series and is based on early instrument temperatures: Manley, G. Central England temperatures: Monthly means 1659–1973. *Q. J. Roy. Metereol. Soc.* **100**, 389–495 (1974).

40. Guemas, V., Doblas-Reyes, F. J., Andreu-Burillo, I., & Asif, M. Retrospective prediction of the global warming slowdown in the past decade. *Nat. Climate Change* **3**, 649–653 (2013).

41. Fyfe, J. C., Gillett, N. P., & Zwiers, F. W. Overestimated global warming over the past 20 years. *Nat. Climate Change* **3**, 767–769 (2013).

42. Slingo, J., et al. *The Recent Pause in Global Warming Parts 1–3.* (The Met Office, 2013).

43. Guemas, V., Doblas-Reyes, F. J., Andreu-Burillo, I., & Asif, M. Retrospective prediction of the global warming slowdown in the past decade. *Nat. Climate Change* **3**, 649–653 (2013). Schmidt, G. A., Shindell, D. T., & Tsigaridis, K. Reconciling warming trends. *Nat. Geosci.* **7**, 158–160 (2014). Mann, M. E., Steinman, B. A., & Miller, S. K. On forced temperature changes, internal variability, and the AMO. *Geophys. Res. Lett.* **41**, 3211–3219 (2014).

44. Quote from page 2 of AR5 summary for policymakers. This AR5 distinction between what was, in actual fact, macroweather and climate variability was not much better than earlier pronouncements on the subject. The following are extracts from the "Summaries for Policy Makers" from the earlier IPCC assessments:

AR1 (1990): "[T]he Earth's climate would still vary without being perturbed by any external influences. This natural variability could add to, or subtract from, any human-made warming; on a century timescale this would be less than changes expected from greenhouse gas increases."

AR2 (1995): "Any human-induced effect on climate will be superimposed on the background 'noise' of natural climate variability."

AR3 (2001): "Changes in climate occur as a result of both internal variability within the climate system and external factors."

AR4 (2007): "On [regional] scales, natural climate variability is relatively larger, making it harder to distinguish changes expected due to external forcings." Cited in Hawkins, E., Edwards, T., & McNeall, D. Pause for thought. *Nat. Climate Change* **4**, 154–156 (2014).

45. Adapted from the International Panel on Climate Change Fifth Assessment Report (2013), Figure 1.4.

46. Adapted from Lovejoy, S. Using scaling for macroweather forecasting including the pause. *Geophys. Res. Lett.* **42**, 7148–7155, (2015).

47. These are the numbers estimated using the NASA GISS temperature series (Fig. 6.5A) and using the zero-lag climate sensitivities. For more details and slightly different twenty-year-lag results, see Lovejoy, S. Return periods of global climate fluctuations and the pause. *Geophys. Res. Lett.* **41**, 4704–4710 (2014).

48. Guemas, V., Doblas-Reyes, F. J., Andreu-Burillo, I., & Asif, M. Retrospective prediction of the global warming slowdown in the past decade. *Nat. Climate Change* **3**, 649–653 (2013). Schmidt, G. A., Shindell, D. T., & Tsigaridis, K. Reconciling warming trends. *Nat. Geosci.* **7**, 158–160 (2014). Mann, M. E., Steinman, B. A., & Miller, S. K. On forced temperature changes, internal variability, and the AMO. *Geophys. Res. Lett.* **41**, 3211–3219 (2014).

49. Lovejoy, S. Return periods of global climate fluctuations and the pause. *Geophys. Res. Lett.* **41**, 4704–4710 (2014).

50. After sixteen years or so, the probability distributions of changes are nearly independent of the lag, so only the long-time (sixty-four-year) probabilities were used. Figure 6.10 is thus slightly simplified in comparison with the original.

51. A little after GRL1 appeared, Karl et al. [Karl, T. R., Arguez, A., Huang, B., Lawrimore, J. H., McMahon, J. R., Menne, M. J., Peterson, T. C., Vose, R. S., & Zhang, H.- M. Possible artifacts of data biases in the recent global surface warming hiatus. *Science Express* 1–4, doi: 10.1126/science.aaa5632 (2015).] produced a temperature series with new ocean and other bias corrections. In this warmer series, the amplitude of the corresponding natural cooling is 0.09°C less than that shown in Figure 6.5A (i.e., about 0.2°C instead of 0.3°C). As the return period for this smaller natural cooling is only about ten years (Fig. 6.10), decadal trends cannot (and did not) detect any statistically significant pause. Because Karl et al. *defined* a pause in terms of decadal scale trends, they made headlines by pronouncing the pause to be nonexistent!

52. Lovejoy, S. Using scaling for macroweather forecasting including the pause. *Geophys. Res. Lett.* **42**, 7148–7155 (2015).

53. This explanation of the pause in terms of (high-level) macroweather statistics is likely to be consistent with the (lower level, mechanistic) explanation by Steinman et al., who explained the pause by specific narrow-scale-range processes involved in the Atlantic multidecadal oscillation (AMO): Steinman, B. A., Mann, M. E., & Miller, S. K. Atlantic and Pacific multidecadal oscillations and northern hemisphere temperatures. *Science* **347**,

988–991 (2015). It is likely that these authors have simply identified the largest and strongest in a scaling hierarchy of fluctuating processes. This conclusion was supported by Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F., & Vallis, G.K. Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Clim.* **23**, 2418–2427 (2010). Our macroweather hindcasts account statistically for the whole scale range and show they are all important (although the largest and the slowest in the hierarchy are the most important).

54. Lovejoy, S. Return periods of global climate fluctuations and the pause. *Geophys. Res. Lett.* **41**, 4704–4710 (2014).

55. Lovejoy, S. Using scaling for macroweather forecasting including the pause. *Geophys. Res. Lett.* **42**, 7148–7155 (2015).

56. Lovejoy, S. Climate closure. *EOS* **96**, doi:10.1029/2015EO037499 (2015).

57. http://joannenova.com.au/2016/02/csiro-wipes-out-climate-division-350-scientists-to-go-since-its-beyond-debate-who-needs-em/.

58. The closest to a time series trend in the CD article is seen in Figure 6.4A (a variant of which had appeared in CD). However, it was not an analysis in time, but rather an analysis in the number of $CO_2$ doublings. Even there, the uncertainty of the trend (slope) was estimated by comparing three different temperature data sets (not by the statistical assumptions about each series to which Keenan was referring). In any case, no trends were used to reject the GNF hypothesis. This was done using the probabilities of temperature differences.

59. Lovejoy, S., del Rio Amador, L., Hebert, R., & de Lima, I. Giant natural fluctuation models and anthropogenic warming. *Geophys Res Lett.* **43**, doi:10.1002/2016GL070428 (2016).

60. This was only one of the 1,000. Each of the other three had 333 realizations.

61. One was a fractional Brownian motion (fBm) model (closely related to the fGn model, discussed in Chapters 5 and 7) and at long-enough timescales. Another was a standard random walk.

62. Things were actually even more complicated than this because Keenan spiced things up in a manner that is very difficult to analyze theoretically. In actual fact, he made 365 realizations of each submodel and, for each (using a nontrivial "excision" procedure), the thirty-two realizations with the largest variability were thrown away. The resulting $3 \times 333$ "clipped" series were then added to the fourth submodel (the unique GCM output) to take the total up to 1,000. The trends of the clipped submodels had a nonstandard—and nontrivial to analyze— probability distribution. In the GRL3 supplement, the trends were (appropriately) estimated by assuming that the residuals had long-range dependencies. But, when the series was classified into trended or nontrended, their statistical distribution was assumed to be Gaussian. This led to the prediction that $893 \pm 9$ could be classified correctly. It seems that the clipping invalidated the Gaussian assumption and reduced this to only 860 correct.

{ 7 }

# Macroweather predictions and climate projections

## 7.1 Predictability limits, forecast skill, and system memory

### 7.1.1 DETERMINISTIC PREDICTABILITY

"Does the Flapping of a Butterfly's Wings in Brazil Set off a Tornado in Texas?" This was the provocative title of an address[a] given by Edward Lorenz, the origin for the (nearly) household expression "butterfly effect."[b] It was December 1972 and it had been nearly ten years since he had discovered it,[1] yet its significance was only then being recognized. Lorenz explained: "In more technical language, is the behavior of the atmosphere unstable to small perturbations?" His answer: "Although we cannot claim to have proven that the atmosphere is unstable, the evidence that it is so is overwhelming."[c]

Imagine two planets identical in every way except that on one there is a butterfly that flaps its wings. The butterfly effect means that their future evolutions

---

[a] Lorenz, E. U. "Predictability: Does the Flap of a Butterfly's Wings in Brazil Set off a Tornado in Texas?" presented at the 139th meeting of the American Association for the Advancement of Science, Washington, DC, 1972.

[b] Ironically, despite this title, Lorenz can take credit only for the science, but not for the famous expression itself. Lorenz had apparently failed to provide a title for his invited talk by the conference deadline. Organizer Philip Merilees then chose it for him.

[c] Lorenz's argument is remarkably similar to the one Poincaré had given sixty years earlier:

Why have meteorologists such difficulty in predicting the weather with any certainty? Why is it that showers and even storms *seem* to come by chance . . . ? We see that great disturbances are generally produced in regions where the atmosphere is in unstable equilibrium. The meteorologists see very well that the equilibrium is unstable, but exactly where, they are not is a position to say; a tenth of a degree more or less at any given point and the cyclone will burst here and not there . . . . If they had been aware of this tenth of a degree, they could have known it beforehand, but the observations were neither sufficiently comprehensive nor sufficiently precise, and that is the reason why it all seems due to the intervention of chance. Here . . . we find . . . a very trifling cause that is inappreciable to the observer, [and leads to] considerable effects . . . [italics added]. (p. 66)

Poincaré, H. *Science and Method*. (Thomas Nelson and Sons, 1912). Poincaré was a self-avowed determinist, but—appreciating statistical mechanics—recognized that chance might nevertheless be needed to deal with complex multiple causes (see Box 1.1).

are "sensitively dependent" on the initial conditions, so that a mere flap of a wing could perturb the atmosphere sufficiently so that, eventually, the weather patterns on the two planets would evolve quite differently. On the planet with the Brazilian butterfly, the number of tornadoes would likely be the same. But on a given day, one might occur in Texas rather than Oklahoma. This sensitive dependence on small perturbations thus limits our ability to predict the weather. For Earth, Lorenz estimated this predictability limit to be about two weeks. From Chapters 4 and 5 and the discussion that follows, we now understand it as the slightly shorter weather–macroweather transition scale.[2]

In Chapter 1, we learned that the ratio of the nonlinear to linear terms in the (deterministic) equations governing the atmosphere is typically about a thousand billion. The nonlinear terms are the mathematical expressions of physical mechanisms that can blow up microscopic perturbations into large effects. Therefore, we *expect* instability. In Chapter 4, we examined instability from the point of view of the higher level statistical laws—the fact that, at weather scales, the fluctuation exponents $H$ for all atmospheric fields are positive (in space, up to the size of the planet; in time, up to the weather–macroweather transition scale at five to ten days). We saw that positive $H$ values imply unstable "wandering" behavior. In the weather regime, none of the atmospheric variables show any signs of convergence to stable, fixed values.

Lorenz was speaking at the dawn of the deterministic chaos revolution to an audience familiar with linear systems with a behavior that is exemplified by the clockwork motions of the planets in their orbits.[d] Although Lorenz's atmospheric model was highly nonlinear, it was nevertheless deterministic. If the starting conditions were perfectly known, the model could, in principle, be predicted perfectly for any time, no matter how far into the future. Lorenz invoked the question of stability to argue that, for the atmosphere, perfect Newtonian predictability[e] would hold in *principle*, but that in practice it would be academic. Even microscopic errors in the data (initial model conditions) would grow catastrophically and, after a time, they would spoil the forecast. Lorenz argued that there are *objective limits* to deterministic prediction (Box 1.1).

To understand these limits, let's consider a very simple chaotic system called the "logistic map"—a model that Lorenz himself studied[3] and one that is closely related[4] to the Mandelbrot set (Chapter 2). First, let's pick a starting number between zero and one. We'll call it $x_0$. Then, multiply it by its difference from one and then by a constant $r$ to obtain the next number, $x_1$, in the sequence:

---

[d] Even here, the reality is nonlinear. The equations of orbital motion are only exactly linear for two bodies, such as a single planet orbiting the sun. As soon as a third is added, the system is nonlinear. The real planetary system turns out to be chaotic, although this is only noticeable for periods of tens of millions of years. See Laskar, J. Large-scale chaos in the solar system. *A&A* **287**, L9–L12 (1994).

[e] Sometimes called "Laplacian" determinism, after Pierre-Simon Laplace (1749–1827), who famously postulated that (using Newton's laws), given the position and velocities of all the particles in the universe, a "divine calculator" could calculate the future and the past.

$$x_{n+1} = rx_n(1 - x_n).$$

As long as $r$ is in the range 0 to 4 and the starting value $x_0$ is between 0 and 1, the sequence remains within the interval [0,1].

In the example in Plate 7.1A, I took[f] $r = 3.72$. Now repeat the process, starting with $x_1$ to generate $x_2$, and repeat. Each repetition advances the clock. The number of repetitions $n$ is the time taken for the system to evolve from $x_0$ to $x_n$. Plate 7.1A shows the result of eighty iterations when three different starting values of $x_0$ are chosen. The values differ only by one part in one hundred million, so that the initial values shown in the figure are $x_0 = 0.4$, $x_0 = 0.40000001$, and $x_0 = 0.40000002$. At first, as expected, the value of $x$ evolves (almost) exactly the same way for each starting value. Yet, after thirty or so iterations, the values start to be visibly different from each other. After thirty-five iterations, they seem to be pretty much unrelated to each other.[g] This can be see more clearly in Plate 7.1B, which shows the result of 1,000 such series, each starting with nearly the same value of $x_0$ and each differing from each other by only one part in *10 billion*. Rather than showing the 1,000 series, I calculated their average (the central curve) and their typical spread.[5] Analysis shows that, at least initially, for every successive iteration (as $n$ increases by one), the spread of the initial values increases by a factor of 1.7. If the initial value was known only to one part in 10 billion, this would correspond to the growth of the forecast error and it would double every time $n$ increased by 1.3, until eventually it becomes large and it "saturates."

The logistic map can be interpreted as a simple model of the atmosphere that includes a weather regime, a weather–macroweather transition, and a macroweather regime. In this case, the values of $x$ might represent the temperature. Plate 7.1B shows that in this model, with appropriate units, the temperature varies over the range of about 2 to 9°C. Imagine starting a weather forecast at time $t = 0$ (the leftmost dashed vertical line in Plate 7.1B). At this point, the spread in $x_n$ is about 0.01, corresponding to initial temperature measurements that are in error by 0.1°C. If $n$ represents the time in forty-hour units, then the forecast error doubles every $1.3 \times 40 = 52$ hours. After ten days (the second vertical line in Plate 7.1B), the spread of the temperature forecasts has increased by a factor of 25, to 2.5°C. The model has saturated (the errors no longer double) and beyond this "lead time,[h]"

---

[f] The full behavior of the logistic map is very complex and was not well understood until, with its help, Mitchell Feigenbaum and others discovered "universality" during the 1970s. One of the complications is that there are only certain ranges of the parameter $r$ that lead to chaotic behavior. For example, below a critical value, $r = 3.56995...$ , the behavior is instead periodic.

[g] In the logistic map, errors can grow exponentially fast so that, in principle—if it was a model for Earth's atmosphere—then it would be so unstable that a butterfly flapping its wings on Alpha Centauri could, via a nearly (but not quite infinitesimal) gravitational tug, provoke growing terrestrial perturbations!

[h] Sometimes called the "forecast horizon." This is usually defined as the time between the end of the data used in the forecast and the beginning of the period of the forecast. A "zero-lead time" one-month temperature forecast for the month of February would use data up until the end of January to make a forecast for the average temperature in February.

the forecast becomes statistical in nature. It can be based purely on past knowledge of the mean climatological temperature (the horizontal dashed line in Plate 7.1B) to be about 6.5°C. The red lines in Plate 7.1A give us the uncertainty, the expected error in the forecast. They tell us that the temperature would be between 4°C and 9°C, with 68% confidence. Beyond the weather–macroweather transition, there is still a little residual skill (see the zigzagging part in Plate 7.1B just beyond the right-hand dashed line that continues to a bit beyond $n = 40$).

Although the system is perfectly deterministic in the sense that the rule tells us how to calculate $x_{n+1}$, where $x_n$ is simple and unambiguous, beyond ten days, the result is purely statistical. Ten days is the deterministic predictability limit. If, at scales of one day, the observations were in error by a tenth of a degree, it would be the longest time that we could ever *hope* to predict the system. Notice that things are not improved much with better observations. If the daily observations had an error eight times smaller (i.e., if they were one eightieth of a degree), then the fore-cast limit would be improved by three error doubling times: $3 \times 52$ hours = 6.5 days (i.e., from 10–16.5 days). The fifty-two-hour error doubling time is a useful number because a real-life forecast system could not make skillful predictions to lead times more than a few times this value.[i] It is a useful benchmark.

In this chapter, we are interested in macroweather forecasts, so the error doubling time isn't the most useful way to characterize forecast skill. Let me intro-duce a different metric that will allow us to compare both deterministic weather and stochastic macroweather forecasts. Although for a given lead time $t$ there are many ways to quantify the skill, a simple, commonly used metric is the mean square skill score (MSSS) or "skill," defined as

$$Skill(t) = 1 - [Average\ E(t)^2 / (Average\ E(\infty)^2],$$

where the forecast is made at time $t = 0$, $E(t)$ is the error at time $t$ after the fore-cast, and $E(\infty)$ is the error infinitely far in the future when the error is at its max-imum. At the time of the forecast, there is no error $[E(0) = 0,\ Skill(0) = 1)]$ and the formula shows that the skill always goes to zero at long times.[6] Note that the typical error in the predictions is the same as the spread in the predictions. It is the dis-tance between the middle (mean) and one-standard deviation (top and bottom) curves in Plate 7.1B.

With this definition, the skill of the deterministic forecast is shown in Plate 7.2A (bottom curve). We see that, as expected (using the time units described earlier; $n = 1$ corresponding to forty hours), it goes from 1 (perfect skill) to 0 (no skill) at around ten days.

---

[i] We mentioned two predictability limits: the error doubling time (fifty-two hours) and the "error saturation" time (ten days) beyond which there is little skill. They are not very different and either one could be used, depending on the application.

## 7.1.2 STOCHASTIC PREDICTABILITY

Because the high-level turbulent laws are stochastic, let's now consider the predictability of a simple stochastic macroweather model. In this case, the model starts with the value at two consecutive time steps $x_n$ and $x_{n-1}$ and uses a simple but stochastic rule to determine the next value, $x_{n+1}$. The rule is simply that $x_{n+1}$ is the weighted sum of the previous $x_n$ and $x_{n-1}$, with an extra random variation $\gamma_n$, called an "innovation."[7] Technically, this type of model is called "autoregressive," or AR[j] for short.

[...].1 describes how innovations arise from energy stored as heat in the oceans, cryosphere, and soil. Let's say that the process was started a long time ago, and that we know the weights are 0.1 and 0.72, so the rule to determine the next value in the series is

$$x_{n+1} = 0.1x_n + 0.72x_{n-1} + \gamma_n.$$

From past measurements, we would like to predict the future. Clearly, if we measure the process at two consecutive times, then—knowing the rule—we could apply it to get a *possible* future by "flipping a coin" to generate possible future innovations. In contrast to the deterministic model, we see that even if the measurements are absolutely error free—so that, for example, $x_0$ and $x_1$ are known exactly—the future given by our rule is only one of an infinite number of possible futures. This is because, although we know the rule that is used to choose the innovations (we know their probabilities), the future innovations themselves are unknown.

Plate 7.2B shows three simulations of a process that had the values $x_0 = 1$ and $x_1 = 2$ when measured at the two times $n = 0$ and $n = 1$ (upper left). ~~Because~~ each simulation is forced to be identical for the first two (measured) values, each series represents a possible future conditioned on these measurements. We can see there are significant undulations, which are the result of the strong correlations between the variables induced by the weights. In many ways, the result is similar to the deterministic model we just examined. However, here, what ruins this "conditional prediction"[k] is not imperfect initial information, but rather the (unknown) future innovations $\gamma_n$. It is an objective rather than a subjective limitation.

A simple way to predict the future of the model is to make a large number of conditional simulations and use the average as the forecast. The uncertainty can be determined by the spread around this average. Figure 7.1 shows the result using 1,000 different series just as in the deterministic case (Plate 7.1B). Notice that the mean (middle) decays quite quickly to its long-term value (0). With these weights, on average, the distance from the long-term average is halved every $n = 3.46$ steps.[l] If we took the time step $n$ to be 2.9 days, then, on average, the

---

[j] Not to be confused with assessment reports, which are all numbered.

[k] The prediction conditioned on knowledge of past values.

[l] AR models with a finite number of terms generally decay rapidly (exponentially) quickly. This rate of exponential decay is a fundamental timescale for the process. Because of this characteristic time, AR processes are scalebound. Slower (power law) decays are discussed later.

model would return halfway to its long-term value every ten days, corresponding to the weather–macroweather transition scale (see Fig. 7.1). In this interpretation, the model represents a stochastic forecast of a macroweather series of consecutive, ten-day averaged temperature anomalies.
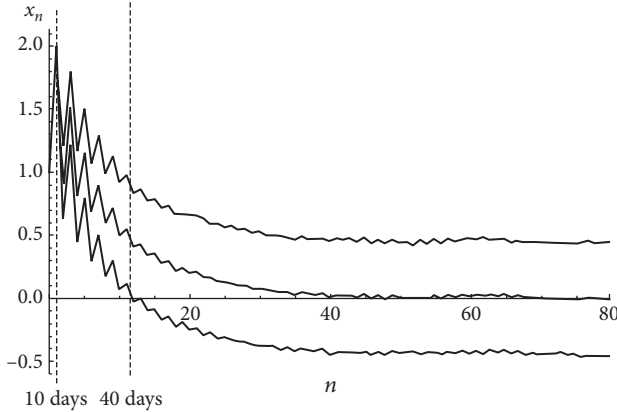


FIGURE 7.1 *The mean and spread of 1,000 stochastic simulations of the same macroweather anomaly model as in Plate 7.2A, with the middle curve showing the mean and the top and bottom curves, the spread (68% of the series between the two). The mean is an exponential return to the long-term value (0); it corresponds to a decrease of a factor of 2 every n = 3.46 time steps. The timescale shown in the model by the dashed vertical lines corresponds to n in units of 2.9 days, with the observations representing the average of the first ten days (the beginning of the macroweather regime).*

To determine the skill of the model, notice that the spread (Fig. 7.1, top and bottom) of the conditional series starts off small, but grows at the longer times, decreasing the skill.[m] Notice that, although the deterministic logistic mapping skill is negligible after ten to twenty days, there is still stochastic skill up to forty days or more.

### 7.1.3 SYSTEM MEMORY

In both the deterministic and stochastic cases, the information contained in the initial conditions (the observations) is used to make a forecast, but the information decays and is eventually lost as the forecast moves further into the future (for the physics, see Box 7.1). The memory of the system is finite. To clarify this, let's investigate the memory of a slightly different stochastic process, the MA model, where the value $x_n$ of the process after $n$ iterations is a weighted sum of past *innovations* instead of the weighted sum of past *values*. For example,

$$x_n = \gamma_n + 0.5\gamma_{n-1} + 0.1\gamma_{n-2}.$$

---

[m] Note that if the magnitude of the innovations was changed, or if the starting (conditioning) values determined by the observations were different, we would still obtain the same skill curve. The rate that it decays with the forecast horizon (lead time) depends on the model weights only (0.1 and 0.72).

As before, physically, the innovations might be past temperature perturbations caused by storms, volcanic eruptions, ocean currents, and so on. According to this model, the value $x_n$ of the process at time $n$ depends on the current innovation ($\gamma_n$) as well as on the previous two innovations ($\gamma_{n-1}$ and $\gamma_{n-2}$). In this example, we see that the next value (at time $n+1$) is

$$x_{n+1} = \gamma_{n+1} + 0.5\gamma_n + 0.1\gamma_{n-1},$$

so that one time step later, the perturbation caused by the innovation $\gamma_n$ has lost half of its importance. Two steps later (at time $n+2$), it has only 10% of its original importance,

$$x_{n+2} = \gamma_{n+2} + 0.5\gamma_{n+1} + 0.1\gamma_{n+1},$$

and three steps later,

$$x_{n+3} = \gamma_{n+3} + 0.5\gamma_{n+2} + 0.1\gamma_{n+1}.$$

plays no role whatsoever. In this model, the role of the innovations is clear: The innovations define the system memory. If we knew all the past innovations, we would know the past values of $x$. Conversely, knowing the past values of $x$ allows us to determine the past innovations. We could then use these past innovations to make a forecast by the same method as in the previous AR case, by making a large number of simulations and conditioning them on the past innovations.

The original AR and MA processes were popularized by an influential book by Box and Jenkins,[8] which introduced the topic to the engineering, hydrology, and other geoscience communities. Unfortunately, Box-Jenkins models are scalebound, having only short-term[9] memory. This is an inevitable consequence of considering models with only a finite number of innovations. In typical applications, only a fairly small number of innovations are used. By adjusting the parameters, the critical scalebound scale over which the skill is reduced by a factor of 2 can be adjusted. But, whichever critical scale is used, when one goes much beyond it, the prediction skill is lost rapidly.[10]

Let's now return to the fGn process, which is scaling and which is essentially an MA process but with an infinite number of innovations. If the weights are chosen to fall off in a power law manner, so will the skill [i.e., much more slowly than in the (scalebound) case of a finite number of innovations]. Plate 7.2A (top) shows this slow fall-off, which allows much longer forecasts (Box 7.2).

BOX 7.1 Earth's (fractional) energy balance, storage, relaxation, memory, and innovations

The way and rate that Earth stores energy is the source of its memory. In this box we examine how storage can be used to understand the long-range memory that is the basis for both macroweather prediction and climate projections (Sections 7.2 and 7.3): the relation between the "internal" and "externally forced" variability.

Earth receives energy from the sun and emits thermal radiation to outer space. If the two were exactly in balance for a long-enough time, it would be in thermodynamic equilibrium, and this would determine its temperature. However, Earth's axis of rotation precesses, its orbit wobbles, and the sun has spots and cycles, so that, even at the top of the atmosphere, incoming radiation is variable. To this we can add volcanic eruptions that shroud Earth by propulsing reflecting particles into the stratosphere. In addition, Earth's constantly evolving clouds, oceans, ice sheets, and vegetation change the rate that energy is absorbed and re-emitted. The forcing—the difference between the rate at which energy is absorbed and emitted—is thus constantly changing. Each forcing corresponds to a different equilibrium temperature that would be attained if the forcing stopped varying and remained at a fixed value. A positive forcing means more energy arrives than leaves. It leads to a rising temperature that, in turn, leads to a rising emission that tends to close the absorption–emission gap [i.e., to reduce the forcing (and conversely for a negative forcing)]. If left long enough, equilibrium will thus eventually be established. The temperature thus responds to positive or negative forcings by "relaxing" to a new equilibrium.[11]

Relaxation is not instantaneous. At any given moment, the rate that energy arrives differs from the rate at which it leaves: Energy is either stored or, conversely, previously stored heat energy is emitted to outer space. Most of the storage (about 93%) is in the oceans. This is not only because the oceans cover 70% of Earth's surface, but also because they are much darker than the average of the land surface. Like the atmosphere, the ocean is a turbulent fluid with eddies and gyres that span wide ranges of scales (and depths); there is a vast range of fast, medium, and slow storage processes. Over land, storage also occurs over a wide range of scales—notably, in land ice, soil moisture, and other land processes.

When it is forced away from equilibrium, how does the climate system adjust (relax) its temperature to its new equilibrium value? In the conventional relaxation model—the "energy balance equation" (EBE)—Earth's storage is treated simplistically as consisting of a thick slab of solid material, such as a bar of metal, that slowly warms or cools depending on the net energy balance. The memory of past forcings is in the slab that stores heat energy. The EBE implies that the process is exponential—that it is fundamentally characterized by a "relaxation time" that characterizes the rate at which equilibrium is reestablished. A jump in the forcing to a new value defines a new equilibrium temperature and then, after one relaxation time, the temperature relaxes halfway to its new value. After two relaxation times, the difference is reduced to one fourth of the original difference, so that three fourths of the adjustment occurs during the first two relaxation times, seven eighths in three relaxation times, and so on.[n]

Now consider the more general situation in which, instead of suddenly increasing the forcing to a new, constant value, it varies continuously as a result of a combination

---

[n] For example, if—as a result of the forcing—the new equilibrium temperature is 4°C greater than the old one and the relaxation time is one year. Then, after a year, the temperature will rise so it is only 2°C below the new equilibrium value. One year later, it will only differ by 1°C and, after the third year, by 0.5°C.

of an anthropogenic, volcanic, or solar forcing, with each new value of the forcing defining a new equilibrium temperature. At any instant, as the temperature starts to relax toward its new equilibrium value, the value itself must be updated to reflect even newer forcing.

By now, you might suspect that, once again—as a result of scaling symmetries—the relaxation process is not exponential, but rather a qualitatively different long-memory power law process, with only a small fraction of the return to equilibrium occurring within the analogous (power law) relaxation time.[12] However, in contrast to exponential relaxation, with each additional relaxation time taking us twice as close to the equilibrium temperature, power law relaxation requires us to wait *multiplicatively* longer. For example—depending on the exponent—it might take ten (power law) relaxation times to halve the difference between the actual temperature and its equilibrium value. In this case, it would take $10 \times 10 = 100$ relaxation times to halve it again to a quarter, and $10 \times 10 \times 10 = 1{,}000$ times to relax it to within an eighth. Power law relaxation is thus a much slower process.

Although exponential relaxation implies a short memory (the effects of forcings further in the past than a couple of relaxation times are almost negligible), power law relaxation is associated with long memories. Rather than approximating Earth as a giant uniform slab of material that gradually warms or cools, one assumes that the oceans have turbulent currents organized into a hierarchy of gyres and eddies, each of which transfers its heat at a rate that depends on its size and depth. Heat transfer over land can also be considered to occur via a hierarchy of land ice, soil moisture, and other processes with rates that depend on size.

Surprisingly, although the new power law relaxation model is simple and physically more realistic, it has only recently been proposed.[o] I suspect that the reason for its neglect is that physicists are trained only to consider mathematical models that involve differential equations, such as the EBE, that are of integer order. Although ubiquitous, it turns out they are the only equations that lead to exponential behavior. As soon as one uses relaxation equations of fractional order—the fractional energy balance equation[13] (FEBE)—one obtains power laws.[p]

What is the evidence for the FEBE? We have indicated that the FEBE relaxation time has a rather different interpretation than in the standard (exponential) EBE. Although in the EBE it represents a transition from a strong memory (scales less than the relaxation time) to virtually no memory (longer timescales), for the FEBE, it represents a transition between two power laws, hence between two regimes, each with strong but somewhat different memories. In Section 7.2, I show how macroweather forecasts (from weeks to years) can be made by exploiting the

---

[o] At the time of writing, this was still work in progress in my McGill group.

[p] It may be surprising to learn that, going back to the nineteenth century, fractional (even complex ordered) integrals and derivatives existed. In the past decade or so, they have been used increasingly in physics, and some theory of fractional differential equations exists. See, notably, Podlubny, who treats the deterministic version of the fractional relaxation equation. [Podlubny, I. *Fractional Differential Equations*. Vol. 198. (Academic Press, 1999).] As far as I know, the stochastic version—the "noise-driven fractional relaxation equation"—needed to take random internal variability into account had not been considered previously.

long memory. For the globally averaged temperature, this is characterized by the exponent $H \approx -0.1$, whereas in Section 7.3 I argue that the decadal, centennial global temperature response to anthropogenic and other external forcings (projections) can best be made with a different exponent, $H_R \approx -0.4$. The plausibility of the FEBE is not only that it is compatible with these strong power law memories, but that it explains both because, theoretically, it implies that the two exponents are linked via the relation $H = -(1/2 + H_R)$.

The FEBE thus unites the variability that is a consequence of externally forcing (anthropogenic, solar, volcanic, or other external changes to the rate of energy coming into or going out of the system) with the random "internal variability," the source ("innovations") of which is otherwise mysterious. The innovations are thus simply internal sources and sinks of heat associated with heat transport toward or away from the surface (mostly, but not only, deep ocean currents). Just as the external forcing drives the system by providing an excess or deficit of heat energy, internal sources (or sinks) of heat—such as currents delivering heat energy to the surface—similarly drive the system.[q]

---

BOX 7.2  Finite and infinite memories

To understand process memories more fully, Figure 7.2 gives an example showing how to construct MA models that are better and better approximations to fGn. Starting with the bottom curve, I show a series of independent random variables: the innovations. This is effectively an MA model with a single innovation. All the curves above this one are built by adding more and more of these innovations, with weights decreasing such that those further in the past become progressively less important. As we move from the bottom to the top series, the number of innovations used to determine a given value in the series increases from 1 to 256. In Figure 7.2, this is indicated by the solid horizontal bar.

With the increase in the number of innovations, the system memory increases, but the memory doesn't only depend on the number of innovations that are used; it also depends on their weights. In this example, they were chosen to fall off slowly (in a power law way[14]) so that, in the limit of an infinite memory, the model respects a scaling symmetry (power law weights). Moving from bottom to top, we see that the series are correlated over longer and longer stretches—indeed, up to, but not exceeding, the number of innovations. Sections farther apart than this are, on the contrary, statistically independent of each other. With a finite number of innovations, the series are all scalebound; however, they can approach the scaling fGn result in the limit of a large number of innovations.

The weights for the older and older innovations fall off with the lag so that the fraction of the total memory for each approximation compared to the full fGn model is 18%, 34%, 49%, 61%, and 71% for 1, 4, 16, 64, and 256 innovations, respectively (bottom to top in Fig. 7.2). Although we see little difference between the series with 64 innovations and the series with 256 innovations (the top two series), we are still

---

[q] The fundamental difference between the two is only that, on average, the internal forcing is zero whereas there is no such constraint on the external forcing.

not so close to the infinite innovation fGn model. The memory of the top series is still missing 29%. This missing memory is associated with very old but numerous innovations. The reason we don't see any obvious consequence is that they almost exclusively affect the very long time undulations (the very low frequencies). The missing 29% essentially moves the entire series up and down. Over the finite range of interest—here, over a series 1,024 points long—this could be (nearly) summarized in a single number: the overall vertical displacement of the series. Although these really ancient innovations may still be "felt," their collective effects are easy to account for even with a relatively small set of past measurements. This is the idea behind the stochastic forecast explained later.
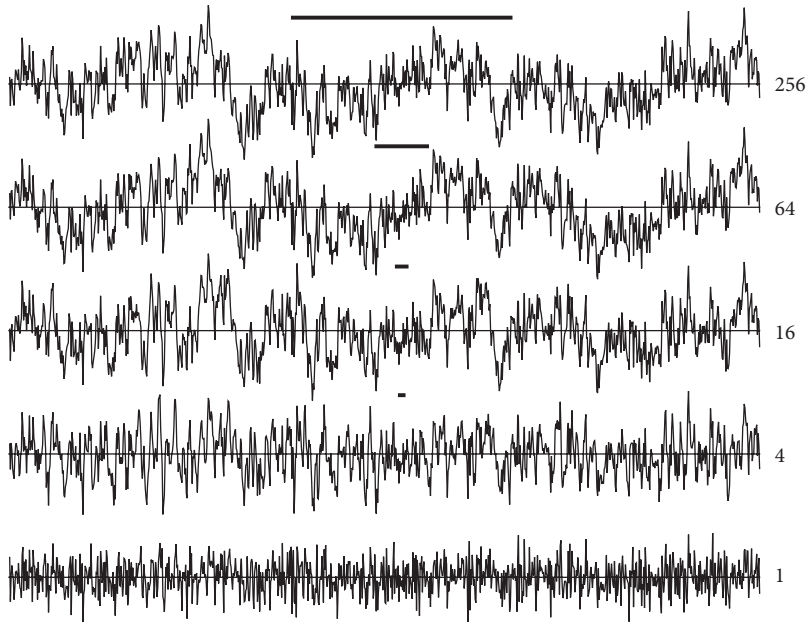


FIGURE 7.2 *This figure shows the memory effect with power law weighting (H = −0.1) using an identical set of 1,024 innovations for each series. As we move from bottom to top, the number of innovations used in the model for $x_n$ increases. This number is indicated by the number on the right-hand side and by the length of the central black segments that are given for reference. As we move from bottom to top, we get an increasingly accurate approximation to the fractional Gaussian noise model that corresponds to an infinite number of innovations.*

## 7.2 Harnessing butterflies: Stochastic macroweather forecasting

### 7.2.1 THE ROLE OF SPACE

In the previous section, I simplified the prediction problem to compare and contrast deterministic and stochastic predictability limits, and to show they can both be quantified using the same measure of skill. In both cases, when forecasts are made

further and further into the future, there is a decay in the initial information—a loss of memory. It seemed plausible that, although beyond the ten-day weather–macroweather transition, the deterministic skill is low, but there may, nevertheless, be significant stochastic skill. Inasmuch as the randomness is a result of the small-scale sensitivity—the butterfly effect—the stochastic forecast effectively harnesses the butterflies to predict the future skillfully.

Before discussing the real-world implementation of this idea, there is a key aspect to address: the spatial part of the problem. Forecasting involves predicting the evolution of spatial structures, not simply forecasting a single time series at a fixed location. In weather forecasting, the size of structures (eddies) determines how far ahead they can be forecast. Lorenz[15] used this to develop a model of inverse cascades of prediction errors.[16] In this scenario, the flapping of a butterfly's wings leads to a perturbation in a structure/eddy a little larger than the butterfly. The effect of this perturbation roughly doubles over the (longer) lifetime of this larger eddy. As discussed in Chapter 4, this lifetime is determined by the energy rate density and by the size of the eddy.[17] At first, this error corrupts the forecast of the small scales just beyond the short lifetimes of the butterfly-scale structures. This in turn leads to the growth of errors in forecasting the corresponding slightly bigger eddies. Over the time it takes for these slightly larger structures to grow and decay (their lifetimes), the forecast will be ruined at this larger scale, so this error is propagated to still larger and larger scales until, eventually, planetary scales are reached. As the forecast progresses to longer and longer lead times, the small-scale details become progressively corrupted. Finally, only the forecast of the planetary-scale structures has any skill, and even this decays rapidly after the corresponding ten-day lifetime.[18]

As explained in Chapter 5, for weather and macroweather forecasting, GCMs deal with the spatial aspect by specifying complete information about the state of the atmosphere at an initial instant. These models are then integrated forward in time. Mathematically, we are dealing with an "initial value problem." However, using the emergent laws for macroweather forecasting, we need *past* data; initial data are not enough. If complete information of the atmospheric state was required not only at an initial instant, but also at all past times, the need for this extra (past) data would be a great *dis*advantage. However, thanks to a space–time symmetry called "space–time statistical factorization," there is a huge simplification of the spatial part of the problem.[19] STSF implies that, although strong spatial relationships (correlations) may exist between the atmosphere at even distant locations—"teleconnections"—whenever long-enough past data exist at a given location, this information is already included in the series. Using series from other spatial locations explicitly as "copredictors" does not necessarily improve the forecast. For example, although a seasonal forecast of Montreal temperatures might be influenced by an El Niño event off the coast of Ecuador, a long-enough series of Montreal temperatures already includes the information relevant for Montreal macroweather forecasts. The Montreal series may be forecast without this additional information.[r]

---

[r] The STSF applies only to the parameter being forecast—in this case, the air temperature. It turns out that a little extra skill is indeed obtained by using climate indices such as the SST off the coast of Ecuador (an El Niño index). This is current work in progress!

### 7.2.2 STOCHASTIC MACROWEATHER FORECASTS: A KILLER APP?

The possibility of using a stochastic approach for modeling and forecasting the atmosphere beyond the ten-day deterministic predictability limit goes back to an influential article, published in 1976, by Klaus Hasselmann (b. 1931).[20] Since then, his approach was developed significantly and applied to macroweather forecasts in the framework of "linear inverse modeling" (LIM), sometimes also called the "stochastic linear framework."[21] These models are essentially extensions of the AR process described earlier, but to continuous times and to a potentially large number of coupled variables, such as atmospheric temperatures at different locations, ocean temperatures at various grid points, El Niño, or other climate indices.

Typical implementations[22] involve twenty components (implying hundreds of empirical parameters). From our point of view, the key point is that LIM assumes implicitly that the system is scalebound (decaying exponentially)—in other words, there is no long-term memory. Consequently, typical applications have little skill beyond a year or so.

Because its theoretical basis is weak and it involves a large number of empirical parameters, LIM is an example of what is commonly termed an "empirically based" approach. Other such approaches have been proposed,[23] and they have had some success by using carefully chosen climate indices related linearly to macroweather variables of interest and by using empirically determined time delays. In contrast, our approach is based physically on the high-level laws that emerge as a result of the space–timescale symmetries respected by the dynamical equations and reproduced statistically in GCMs.

According to the scaling storage and relaxation model (Box 7.1), macroweather can be well approximated by fGn. fGn is also the simplest scaling macroweather model, and its application to macroweather forecasting, including the numerical algorithm to make the optimal forecast, is SLIMM.[24] To see how this works, let's take the example of predicting the annually, globally averaged temperature in Figure 6.4A. As we saw in Chapter 6, most of the change in this temperature since 1880 is the result of low-frequency anthropogenically forced climate—not macroweather variability. Our first step, therefore, is to remove this and to forecast the residuals (the internal variability; Fig. 6.4C). Failure to remove the anthropogenic trend[25] leads to poor results.[26]

In Box 7.2, we saw that fGn can be represented as a sum of an infinite number of past innovations that have weights that decrease in a power law way determined by the fluctuation exponent *H*. How can we use this knowledge to make forecasts? Let's say we have monthly resolution temperature data and we want to make a forecast three months ahead. First, from fluctuation analysis (as in Chapter 2), we can estimate *H* (= −0.1 here), which determines the innovation weights. Next, we need the past innovations—in principle, from the infinite past. Using this information, the temperature three months from now is given by the weighted sum of these past innovations plus a contribution from the (unknown, future) innovations over the next three months.

These future innovations could be chosen randomly; they are *possible* innovations that determine *possible* temperatures three months into the future. They would yield one of an infinite number of possible futures. The actual future would not be exactly the same, but it might be close. Just as in Figure 7.1, we could then repeat the procedure 1,000 times, generating 1,000 possible temperature series, each of which would be identical over the entire past, but would differ over the future. However, just as in the examples in Plate 7.2B, they would start off near each other and generally become more and more different further into the future. Our final forecast would be the average; the skill of the forecast would be determined by the spread of their futures conditioned on the past data.[27]

When this approach is applied to global, annually averaged temperatures for the critical "pause" period following 1998, we can obtain the forecast shown in Figure 7.3. We see that the temperature over the next fifteen years (bottom right) is reproduced quite accurately.[28] Indeed, although the GCM predictions were about 0.2°C too high (Fig. 6.9), Figure 7.3 shows the detailed pause hindcast was accurate to within about ±0.05°C.[29] This shows that the overforecasting of the pause temperature was purely a GCM problem. The atmosphere itself behaved exactly as expected!
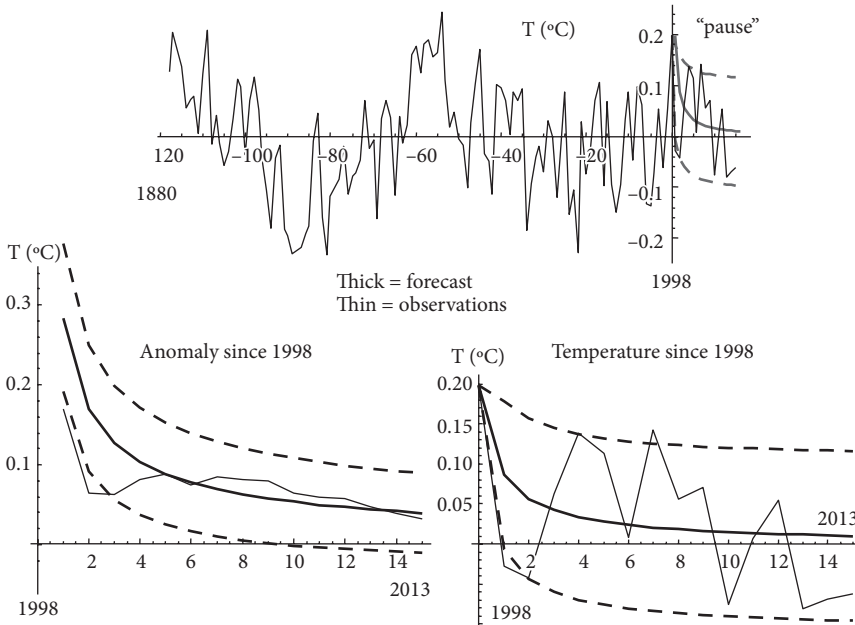


FIGURE 7.3 *(Top) The residuals temperature of Figure.6.4C after the low-frequency anthropogenic rise has been removed (thin line) with the hindcast from 1998 (thick line; the dashed lines are ±34% confidence intervals of the forecast, which is the likely range of their error). (Bottom left) The anomaly defined as the ~~natural temperature (i.e.,~~ residual, ~~Fig. 6.4C)~~ over the hindcast lead time (thin line). The thick line is the hindcast, the dashed lines are the ±34% confidence limits. (Bottom right) The temperature since 1998 (thin) with the hindcast (thick). It is a blowup of the hindcast part of the top right.[30]*

This successful forecast based on the fGn model works, but it can be improved in several ways. Box 7.3 examines the SLIMM approach, which has been operational since April 2016.[s]

The discussion in the box shows that it works extremely well on GCM control runs. Control runs provide a good testing ground because they have pure macroweather statistics. There are no low-frequency, slow climate forcings to worry about.

Up until now, scaling approaches to the atmospheric sciences have primarily provided a new way for understanding the weather and climate, for making models with realistic statistical properties. What has been lacking is a practical application of scaling that improves on conventional alternatives, an application that confers palpable advantages. Macroweather forecasts have applications to agriculture, renewable energy, environmental risk management to name only a few. Any prediction technique representing even a modest improvement over traditional methods would provide a relative advantage and would be profitable. Will the economic advantages of SLIMM-based stochastic macroweather forecasts turn out to be scaling's "killer app?"

---

BOX 7.3　SLIMM: The role of the distant past and forecasting GCM control runs

---

The forecast approach used in Figure 7.3 assumes that we had an infinitely long series of temperature data $T_n$ for times $n \leq 0$ and that each is a weighted sum of the preceding innovations $\gamma_n$. To make the forecast, it "inverts" the $T_n$ series to determine the innovation series $\gamma_n$. From there, a possible future is generated by resumming the innovations with different weights and with randomly chosen future innovations. This process is then repeated many times, and the forecast is the average over many of these possible futures; each has the same (measured) series of past innovations, but a different, randomly chosen set of future ones. Although this forecast approach works, in the practical case in which the amount of past data is limited, there is a better way. We can avail ourselves of the mathematical solution of the finite fGn forecast problem.[31] This solution has a remarkable feature that arises because of the huge fGn memory. Let's say we have temperature data for twenty years at monthly resolution—240 monthly values in all—and we want to use them to make a forecast, knowing the process was an fGn with exponent $H$. We want to combine the past data in a way that makes the average prediction error as small as possible.[32] It turns out this can be done by an appropriate weighted sum of the past values.

　　It seems obvious that if we want to forecast next month's temperature, that the most recent temperature—the average over the month that just ended—should be accorded

---

the most importance, given the highest weight. Indeed, this is a general feature of stochastic forecasts. Figure 7.4 shows the weights for the fGn with the familiar value $H = -0.1$ (appropriate for globally averaged temperatures), confirming this feature (the spike at the right of the figure). Not surprisingly, the figure shows that older temperatures in this series generally receive much less weight. However, the figure also shows that, starting with the oldest 10% or so of the data (in this example the first two years of the data at the left of Fig. 7.4)—remarkably—the weights rise sharply so that the oldest month (240 months in the past) gets a comparable amount of weight to the most recent month! In the words of the mathematicians who discovered this effect, this is because the most ancient data "is the closest witness to the distant past."[33]
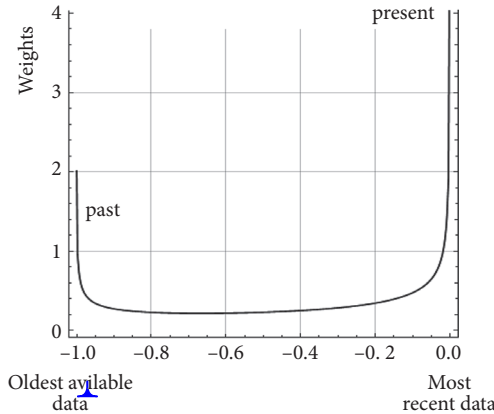


FIGURE 7.4 *This figure shows how the weights used to forecast the next time step decay, as we move from the present (at the far right) to the distant past (far left), rise sharply. The horizontal timescale is the fraction of the interval over which past data are available, with –1 indicating the oldest. The plot is for H = –0.1.[34]*

The data from the distant past are largely used to determine the overall offset of the series from its long-term value (0). Therefore, before applying SLIMM to make real-world macroweather temperature forecasts, we must first remove extraneous nonmacroweather features—in particular, low-frequency anthropogenic climate variations. Before dealing with this real-world complication, let's see how SLIMM does on the pure macroweather problem of forecasting GCM control runs that, by definition, don't have any climate forcing. If SLIMM is indeed a good high-level stochastic model of low-level GCM behavior, then it should be able to predict the control runs to their theoretical stochastic predictability limits.

My colleagues and I used thirty-six globally and monthly averaged Coupled Model Intercomparison Project Phase 5 (CMIP5) control run outputs. For each, we estimated the relevant exponent $H$ and found a mean $H = -0.11 \pm 0.09$[35]; this value near zero implies a huge memory. Using this method at a one-month resolution, we produced hindcasts with lead times from one to twelve months,[36] verifying them against the control run. Figure 7.5 compares both the expected (theoretical) skill and the actual hindcast skill. Although there is a fair bit of variation from GCM to GCM, the theoretical and actual skills are, on average,

very similar and they are very high, with more than 50% skill on average at twelve-month lead times.[37] These results make it quite plausible that the theoretical stochastic predictability limit (Plate. 7.2A) really is an upper bound on the skill of macroweather forecasts.
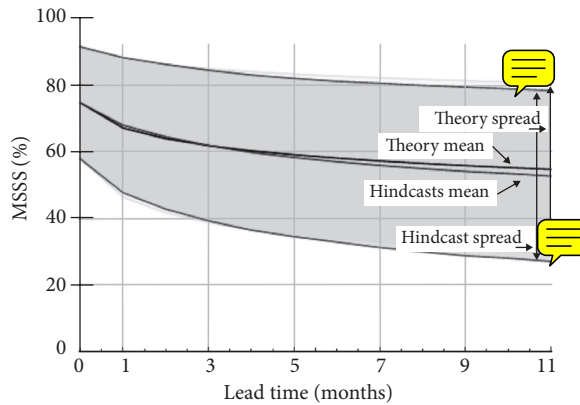


FIGURE 7.5 *The mean square skill score (MSSS; skill) for hindcasting thirty-six Coupled Model Intercomparison Project 5 Global Climate Model (GCM) control runs, each at least 2,400 months long. Each GCM had a slightly different H value and, hence, different theoretical predictabilities. The graph shows that both the means and the spreads (shaded areas) of theory and practice (Scaling Macroweather Model hindcasts) agree very well. A lead-time zero forecast is for the month immediately following the latest available data.[38]*

#### 7.2.3 REGIONAL FORECASTING

The ability to forecast GCM control runs to their stochastic predictability limits and the successful hindcast of the pause in the globally averaged temperatures show that fGn is a reasonable macroweather model and that SLIMM's numerical algorithms are as good as could be hoped. What about applying this to regional macroweather forecasting, for example, on 5° × 5° grids at monthly scales? I mentioned that, as a result of the statistical space–time factorization symmetry, if long-enough series are available, they could be forecast directly—that using information at other locations as "copredictors" would not increase the overall skill because the past data at the location already contain the information from the spatial correlations.

The various steps in the forecast are illustrated in Figure 7.6, using the pixel over Montreal as an example. The first step is to remove the low frequencies that are not a result of (scaling) macroweather variability to be able to apply the SLIMM forecasting algorithm. This is not totally trivial for two reasons. First, the low frequencies have both a mean component of anthropogenic origin, but also—because of its long memory—one resulting from long-term macroweather variability. Second, at 5° × 5°, there are strong annual cycles, and they are not perfectly fixed; they evolve slowly from one year to the next.[39]
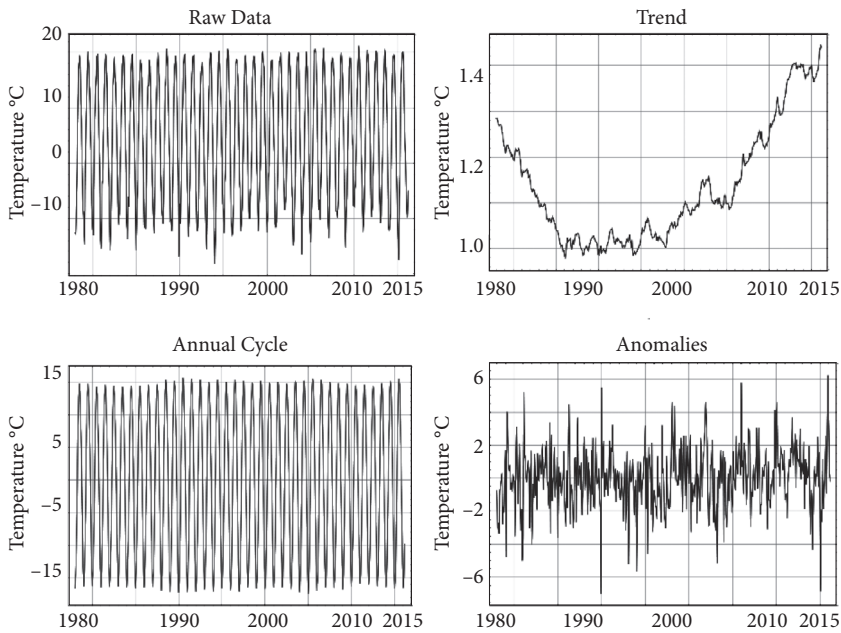
FIGURE 7.6 *An example of forecasting the temperature at Montreal.[40] The top left shows the raw monthly data, the bottom left shows the mean annual cycle as deduced using a thirty-year running estimate, the upper right shows the low-frequency trend,[41] and the bottom right shows the resulting anomalies forecast by SLIMM.[42]*

The solution that we adopted was to use the fact that the anthropogenic effects are only strong at decadal scales, so that by filtering they can be largely separated from the macroweather variability.[43] The annual cycle can be dealt with by using the past thirty years of data to make running estimates of the latest cycle and then use it for next year's forecast. Finally, the anomalies that result from removing the anthropogenic warming and the annual cycle can be forecast using SLIMM (lower right, Fig. 7.6). These steps in the "preprocessing" stage needed to obtain a set of residuals (anomalies) with behavior as close as possible to a pure fGn process are shown in Figure 7.6 for the Montreal grid point.

To test the method, for each grid point, the fluctuation exponent $H$ (Fig. 5.5) of the preprocessed anomalies (bottom right, Fig. 7.6) were then estimated and monthly resolution hindcasts with lead times up to one year were made, with one such hindcast for every month for thirty years. It turns out that the average hindcast skill is quite close to the theoretically maximum possible skill.[44] Using this approach, since April 2016, we have regularly made regional macroweather temperature forecasts.[t]

---

[t] Forecasts are made for monthly, seasonal, and annual lead times, initially at $5° \times 5°$ and, since 2017, at $2° \times 2°$ resolutions (about 200 km). See http://www.physics.mcgill.ca/StocSIPS/.

### 7.2.4 ~~STOCHASTIC SEASONAL TO INTERANNUAL PREDICTION SYSTEM AND CANADIAN SEASONAL TO INTERANNUAL PREDICTION SYSTEM~~

We saw that before applying the SLIMM algorithm to forecast the scaling fGn-like macroweather component, we must first estimate, remove, and forecast separately the nonmacroweather low frequencies and annual cycles. The overall model, including this preprocessing, is called the "Stochastic Seasonal to Interannual Prediction System" (StocSIPS).[u] Although it is stochastic, it is comparable in scope to deterministic "long-range" or "extended-range" macroweather GCM forecast systems. In the following, I compare StocSIPS to a typical GCM,[v] the Canadian Seasonal to Interannual Prediction System (CanSIPS[45] ). For more information about how CanSIPS makes such forecasts, see Box 7.4; for a summary CanSIPS-to-StocSIPS comparison, and the new CanSIPS–StocSIPS hybrid model: CanStoc (see Table 7.1).

---

BOX 7.4  GCMs and macroweather forecasts: CanSIPS and CanStoc

---

Since 2010, CanSIPS has been the official extended long-range forecast product of Environment Canada, Canada's national environmental and weather service. CanSIPS products are based on a "multimodel ensemble (MME)" consisting of realizations of two different GCMs: ten from the Canadian Climate Model 3 (CanCM3) and ten from CanCM4—twenty members in all. The publicly available maps are only over Canada and are for monthly through annual temperature and precipitation. However, in the comparisons included here, I accessed the global products and verified the global hindcasts since 1980.

Let's look in detail at how CanSIPS forecasts are made. The first step is the initialization; CanSIPS starts with reanalysis data from the ECMWF. Reanalyses are effectively data–model hybrids that we used frequently in earlier chapters. CanSIPS uses reanalyses obtained by assimilating meteorological observations into the ECMWF GCM.[46] The problem is that the ECMWF and CanSIPS models each have their own climates, so the CanSIPS GCMs cannot ingest the ECMWF reanalyses directly. Instead, the ECMWF initial values must be converted into ECMWF anomalies.[47] These anomalies are then added to the CanSIPS climate to obtain the CanSIPS initial values, the "actuals."[48]

The CanSIPS forecasts are then made using the twenty-member CanSIPS ensemble followed by complex (and computationally demanding) postprocessing that corrects primarily for the "model drift" and poor local estimates of climate sensitivity to anthropogenic warming. "Model drift" refers to the tendency of

---

[u] Since April 2016, StocSIPS has provided monthly, seasonal, and annual temperature forecasts. See http://www.physics.mcgill.ca/StocSIPS/.

[v] CanSIPS is similar in its functioning and comparable in its skill to the European Seasonal to Interannual Prediction System (or EuroSIPS), although each is based on a different GCM. According to the WMO, there are twelve international "producing centers."

model temperatures (even in control runs; see Chapter 5) to display low frequency variations that are attributed to artifacts arising from slow (mostly ocean) processes that are not fully "balanced" when the model is initialized. An additional problem is that the model does not have perfect representation of the sensitivity to anthropogenic effects, leading to systematic errors that contribute further to a low-frequency "drift." Both are partially removed using hindcasts over the previous five-year period in an attempt to estimate (and remove) spurious linear trends.[49] This is part of a complex GCM postprocessing procedure that attempts to remove known problems.

Despite these manipulations, the final forecast result—an "actual"—is seriously in error, typically by a factor several times larger than the amplitude of the anomalies (i.e., the ~~climatological~~ variability).[w] The fundamental reason for this is that the GCM climate is ~~wrong~~, so the macroweather anomalies fluctuate around the wrong values. Because the model seasonality is also poor, this wrong climate value depends on the season. As a result of these flaws, seasonal variations and climate biases are removed from the forecast model to yield anomalies. The publicly available macroweather forecasts are of the anomalies, not the actuals (see Table 7.1).

Finally, these anomalies are usually given a final layer of processing— "downscaling"—which is often needed because the GCM outputs are on grids hundreds of kilometers across, and often one is interested in forecasts for towns or cities that are at scales much smaller than this. The much larger grid-scale forecast must be scaled down to the city or local scale. This downscaling is performed either by using empirical relationships established with the help of long series of local meteorological station data or—a much more costly solution—by "nesting" a regional-scale GCM inside the global GCM (a kind of mini GCM within a GCM). The lateral boundaries of the regional model are constrained by the outputs of the full (global-scale but lower resolution) GCM. In comparison, as Table 7.1 points out, a final advantage of StocSIPS is that it avoids downscaling; if available, it can forecast the local station data directly.

An advantage of CanSIPS is that it incorporates an ocean model, and the ocean weather– ocean macroweather transition (and hence the deterministic predictability limits) of the ocean component is longer than the roughly ten day (atmospheric) weather–macroweather transition time.[50] This explains why CanSIPS has somewhat higher skill than StocSIPS over some areas of the ocean (Fig. 7.7), at least for lead times of several months. This leads to the possibility of combining StocSIPS and CanSIPS into a single hybrid model that has better skill than either taken separately.

---

[w] When the errors are larger than the climatological spread of the anomalies, the MSSS becomes negative. MSSS values of –200% or less are frequent for GCM forecasts of actuals.

TABLE 7.1   Summary of deterministic (GCM) versus stochastic macroweather forecasts.

| CanSIPS (GCM) | StocSIPS |
|---|---|
| 1. CanSIPS data input: Reanalyses from the ECMWF. Reanalyses are complex model–data hybrids using satellite, aircraft, station, and other data. It puts them onto a grid with a big model. | 1. StocSIPS data input: Start with the monthly station series. |
| 2. CanSIPS input anomalies: Remove ECMWF climatology to find ECMWF anomalies, which are used as CanSIPS input anomalies. | 2. Preprocessing: Separate the anthropogenic warming trend and annual cycle from the internal (macroweather) variability by using historical data at one-month resolution (this could be reduced to two weeks). |
| 3. Data assimilation into CanSIPS: Add the ECMWF anomalies to the CanSIPS climatology to estimate the CanSIPS actuals. Between $t = -12$ hours and $t = 0$, the actuals are assimilated into two different GCMs: CanCM3 and CanCM4. | 3. Internal variability forecast: Use the system memory to forecast the macroweather series from months to years as necessary. Both the mean and uncertainty are forecast together. |
| 4. Model integration: Each GCM is coupled with an ocean model and integrated one year into the future for ten different realizations. These twenty members define the ensemble. | 4. Warming trend and annual cycle forecast: This forecasting can be done, for example, by simple persistence of the latter from the previous year. |
| 5. Recovery of forecast anomalies: The CanSIPS climatology is removed from the forecast to recover the forecast anomalies. | 5. Final forecast at the station: Add the two forecasts. |
| 6. Removal of some artifacts: Postprocessing partially removes artifacts such as model drift. It also rescales the anomalies so the amplitude of their fluctuations is correct. It cannot correct for poor seasonality. The mean of the twenty outputs is the final forecast; their spread defines the forecast uncertainty. | **CanStoc (hybrid)** |
| 7. Downscaling: The forecast is for a grid point of roughly 100,000 km$^2$. For most purposes, it must be downscaled. This is done either by ad hoc multiple-regression techniques using local station data or by using another nested regional model of similar complexity to the original GCM. This regional GCM takes its boundary data from the full global GCM. | CanStoc is a hybrid GCM–stochastic model. The CanSIPS model is used as a copredictor and the CanStoc forecast is the weighted average of the StocSIPS and CanSIPS forecast. |

CanCM, Canadian Climate Model; CanSIPS, Canadian Seasonal to Interannual Prediction System; CanStoc, CanSIPS-StocSIPS hybrid; ECMWF, European Centre for Medium-range Weather Forecasting; GCM, Global Climate Model; StocSIPS, Stochastic Seasonal to Interannual Prediction System.
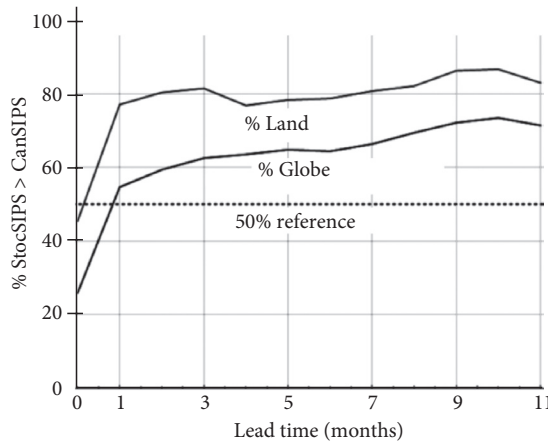
FIGURE 7.7 *The relative skill of the Stochastic Seasonal to Interannual Prediction System (StocSIPS) and the Canadian Seasonal to Interannual Prediction System (CanSIPS) anomaly hindcasts (1980–2010) over the globe and over land only, showing that the StocSIPS relative advantage increases systematically with lead time and is particularly strong over the oceans.[51]*

Box 7.4 shows that the main problem with the GCMs is that their climates, including their annual cycles, are rather different from those of the real world. Because of this, the publicly available GCM macroweather forecasts are of the anomalies, not the real temperatures (the "actuals"). For these anomalies, the comparison with StocSIPS is much closer (see Plate 7.3). However, Plate 7.3 shows that even for anomalies, over most of the globe, for two months and longer, StocSIPS has higher skill. The relative skill of StocSIPS is particularly high over land, which is probably a result of the fact that the CanSIPS ocean model is still within its deterministic predictability limit of one to two years, making its ocean forecasts reasonably accurate. This impression in bolstered by Plate 7.4, which compares CanSIPS at six months and StocSIPS at two years (the skills are comparable), and also by Figure 7.7, which shows that relative advantage of StocSIPS grows with lead time and is particularly strong over land.[52]

The overall saving in computational speed is estimated to be on the order of a million. At the same time, StocSIPS directly forecasts the conditional ensemble average (i.e., effectively, the results of an infinite ensemble) whereas CanSIPS has only twenty members in its ensemble. StocSIPS can also be used to make forecasts directly at individual measuring stations, avoiding otherwise time-consuming and error-prone "downscaling" procedures.

StocSIPS can be extended directly to other fields such as wind or precipitation, which are known to have macroweather statistics roughly satisfying the SLIMM requirements.[53] But, the main advantage of StocSIPS may be its ability to forecast other fields directly, such as insolation, wind power, or degree-days, which are

currently only forecast indirectly by GCMs. Other future extensions of StocSIPS could include drought indices and the prediction of extremes.[*]

## 7.3 ~~The world our children will inherit~~: Scaling and climate projections

### 7.3.1 PREDICTIONS AND PROJECTIONS

Over the weather and macroweather range, the problem of predicting the atmosphere is one of understanding and exploiting its internal space–time variability. But what about multidecadal climate predictions to 2050 or 2100? New, slow anthropogenic changes will be dominant, so external "climate forcings" are now crucial. As discussed in Box 6.2, some of these (such as solar variability and volcanism) are natural, but in the Anthropocene, the main forcing is a result of emissions of greenhouse gases and aerosols.[54] Making a statement about the state of the atmosphere in 2050 or 2100 therefore depends on economically and politically driven scenarios that specify the future anthropogenic impacts on the atmosphere (especially emissions). They are *projections,* not predictions.

The aim of the projections is to determine the consequences of our actions. We would like to separate anthropogenic effects from the unimportant and distracting effects of macroweather "internal" variability as well as from the responses to past natural forcings (solar and volcanic). To a first approximation, we can identify the higher frequencies with the weather and macroweather, and the lower frequencies with the projection (i.e., the climate). This was the motivation for using the empirical notion of "climate state," defined by averaging over weather and macroweather scales (e.g., over the climate-normal period; Plate 5.2). However, this definition of climate state doesn't completely remove the effects of macroweather (internal) variability,[y] nor does it account for the consequences of our actions on the climate that are delayed, notably as a result of the oceans (see Box 7.1 for this question of storage and memory).

To separate the anthropogenic responses and macroweather variability cleanly requires a stochastic model.[z] Once the model is given, one can make an ensemble of many (stochastic) realizations, each differing only by the macroweather variability.

---

[*] ~~Ironically, although there were no funds available for the scientific research for developing StocSIPS, when the basic research had been done, there were funds to help determine its commercial potential!~~

[y] Recall that macroweather fluctuations decrease with scale, so that at ten years globally, they are about 0.2°C (= ±0.1°C). In the Anthropocene, this is roughly equal to the decadal-scale forced response (Fig. 5.6); it defines the beginning of the climate regime. Yet even at forty years, macroweather variability has only decreased to 0.15°C, whereas the corresponding (historical) anthropogenic fluctuations were about 0.4°C, so the two are still comparable (Fig. 5.6)

[z] When this is done using GCMs, averages are made over many runs and even over different GCMs. For projection purposes, the GCMs are thus treated implicitly as yielding stochastic responses to deterministic forcings.

In principle, the realizations can be averaged over the ensemble to yield the projection that we seek. For GCMs, the different stochastic realizations are produced by running the model many times with identical forcings, but with slightly different initial conditions.[aa] Because each run is computationally expensive, to eliminate the weather and macroweather variability better, ensemble and temporal averaging are combined. The result is a statistical ensemble of possible climate states conditioned on the forcing. Finally, the usual procedure is to broaden the ensemble by extending it to include the outputs from different GCMs all over the world: a multimodel ensemble, MME. The MME mean is then used for the projection, and the spread of the climate states about the mean quantifies the uncertainty (see Box 7.5).

---

BOX 7.5  Climate projections with the MME

For GCMs, the main difficulty in projecting the future climate is the modeling of the complicated set of water vapor, cloud, and radiation feedbacks. Growing levels of greenhouse gases increase the temperature, which in turn increases the evaporation. Higher humidity leads to stronger greenhouse warming (the direct effect), but also to increased cloud cover, which can either warm or cool Earth (the indirect effect), depending on their type. The resulting wide dispersion of sensitivities is referred to as "structural uncertainty"—the fact that each GCM has a different climate. The strengths of the various feedbacks are intrinsic parts of each model's dynamics so that GCM climate sensitivity cannot be adjusted or "tuned"; it is an "emergent" model property.

GCMs divide Earth into grids hundreds of kilometers across, with kilometric-scale vertical layers. Although the basic GCM is a model of atmospheric circulation, for climate modeling GCMs are coupled to ocean circulation models and, increasingly, to dynamical cryosphere and carbon cycle models. It is usual to make projections using dozens of different models grouped together into an MME. For example, the CMIP5—the basis of the IPCC's AR5—considered more than forty different GCMs. The MME mean is used as a surrogate for the real climate, and its spread about the mean is taken to quantify the uncertainty in the projected mean.

To make projections, we need assumptions about future human economic activity and the implied anthropogenic forcings. The IPCC therefore created four "representative carbon pathways" (RCPs) that are, effectively, sophisticated sets of assumptions about how future emissions might evolve throughout the century (and up to the year 2300 in their extended versions),[55] while attempting to take into account future economic development.[bb] The same scenarios are used in economic modeling in integrated assessment models that predict the coupled global atmospheric and economic evolution. Each RCP is labeled by the specified radiative forcing in the year 2100 (measured in watts per square meter). For example, the most benign is

---

[aa] For the past part of the model run, this is based on estimates of the historical forcings; for the future part, it is prescribed by the scenarios.

[bb] For our purposes, an important aspect of this is the assumption that aerosol pollution will decline as a result of cleaner factories and automobiles. This means that it is assumed that past relationships between greenhouse gas emissions, the economy, and aerosol emissions will no longer hold. Hence, for regional projections, we need to account separately for greenhouse gas and aerosol forcings. For global projections, the aerosols are implicitly taken into account in the $CO_{2eq}$ forcing.

RCP2.6, which assumes the total forcing in 2100 will only be 2.6 W/m$^2$ more than preindustrial levels, whereas the most dangerous (RCP8.5) assumes it will be more than three times higher (see Fig. 7.8). In the following, to simplify, I mainly discuss the two extreme RCPs (2.6 and 8.5) and the (more likely) intermediate one[56]: RCP4.5.
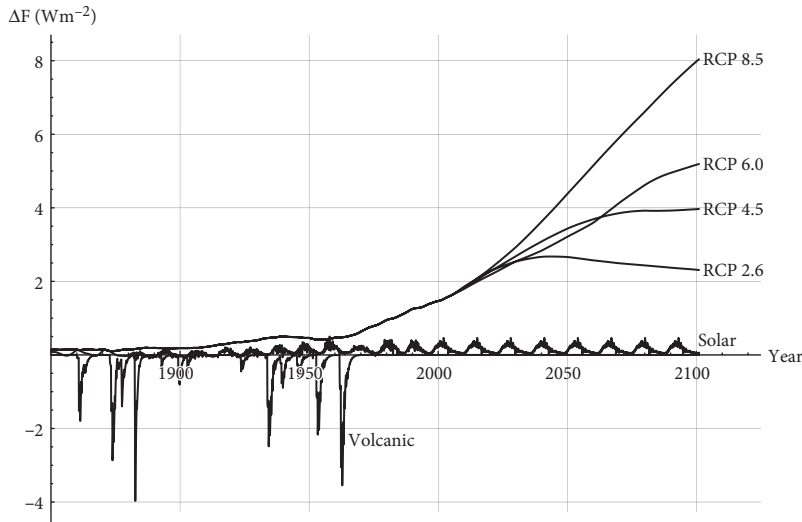


FIGURE 7.8 *The International Panel on Climate Change forcing scenarios: the representative carbon pathways (RCPs). Also shown are the main natural forcings: volcanoes and solar. Notice that the solar forcing is small compared to the anthropogenic forcing.*[57]

With the exception of RCP2.6, they are all monotonically increasing up until 2100. RCP2.6 assumes massive deployment of "negative emission" technologies to suck the $CO_2$ out of the atmosphere. The ability to do this at the required scale is highly speculative,[cc] so this scenario is probably not realistic.[dd] The solar activity is the only natural forcing, ~~which~~ is assumed to continue its current eleven-year solar cycle indefinitely into the future. Finally, to make projections, one also needs estimates of "reconstructed" past forcings. The IPCC's AR5 provides these from 1750, and they are also shown in Figure 7.8. They include reconstructions of past solar and volcanic forcing.

**DIMINISHING RETURNS**

The first modern estimates of climate sensitivity (i.e., the expected temperature change following a doubling of $CO_2$ concentration), derived in the 1970s, yielded the classic consensus range of 1.5 to 4.5°C. Since then, there has been a million-fold increase in computational speeds, and new evidence has supported the validity of this uncertainty range, but has failed to narrow it. Only for the IPCC's AR4 was there

---

[cc] The favored approach is BECCS. It would involve planting and regularly harvesting biomass over an area one to two times that of India, burning it to produce energy, capturing the $CO_2$, and then burying it for millennia.

[dd] However, all the IPCC economic scenarios rely optimistically on this speculative technology. See Anderson, K. & Peters, G. The trouble with negative emissions: Reliance on negative-emission concepts locks in humankind's carbon addiction. *Science* **354** (6309), 182–183 (2016).

sufficient agreement among experts to reduce the consensus range to 2.0 to 4.5°C, based mostly on evidence from the Coupled Model Intercomparison Project Phase 3 GCMs. However, AR5 backtracked, mostly because of more direct observations-based estimates of climate sensitivity, which showed that a lower 1.5°C value could not be rejected using historical observations. In fact, historical simulations from the MME of thirty-two GCMs of the more recent CMIP5 yielded a mean transient climate sensitivity (TCS)[58] about 15% higher than observations (Fig. 7.9).
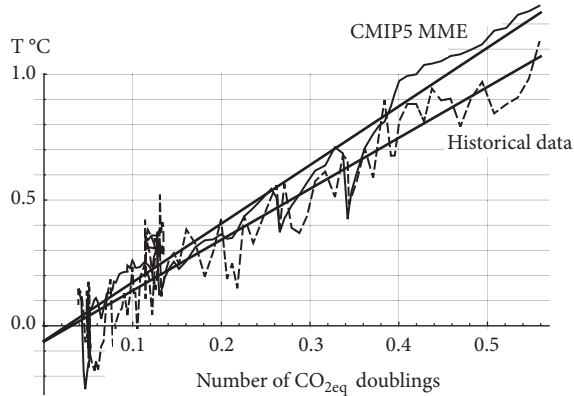


FIGURE 7.9 *The annual global temperature change since 1880 as a function of the number of doublings of carbon dioxide equivalents ($CO_{2eq}$).[ee] Four different temperature series were used (bottom) and thirty-two Coupled Model Intercomparison Project 5 (CMIP5) models were used (top). Because the $CO_{2eq}$ value is a little larger than the $CO_2$, the result (the TCS) is similar to Figure 6.4B, with a slope that was the EffCS, but with a somewhat lower value: 2.02 ± 0.08°C for the observations and 2.30 ± 0.5°C for the models ($CO_{2eq}$ doubling). The spreads are the 90% confidence intervals for the four series and the thirty-two models, respectively. The MME is a bit too high by about 15%.[59]*

Investments in the GCM approach to decrease the uncertainty now appear to be a classic case of diminishing returns, and the main proposal by the modelers is for more of the same: bigger, faster computers. It is claimed that if the resolution can be reduced from 10 to 50 km down to 1 km, that the models will be "cloud resolving" so that, finally, these limitations will be overcome.[60]

However, projecting the state of the atmosphere involves averaging GCM outputs over simulated decades. Will the inclusion of kilometric structures that live typically for less than an hour—a millionth of this time—really improve projection realism?[ff]

---

[ee] For example, the value 0.5 corresponds to half a doubling—meaning, there has been a factor of $2^{0.5} = 1.41$ increase in the $CO_{2eq}$ above the preindustrial value of 277 ppm (i.e., it corresponds to 391 ppm).

[ff] Beyond the ten-day deterministic predictability limit, the GCMs are stochastic, so that it is sufficient that the models generate realistic *statistics*. Realistic small-scale structures and morphologies are not needed and may not be possible anyway. Although GCMs may be excellent research tools, they are not—at least at the moment—the best way to predict either seasonal or annual temperatures, or to project temperatures to 2050 or to 2100.

But using the MME is problematic. It isn't obvious that the average of the climate sensitivity over diverse GCMs, each with its own climate, is a good estimate of Earth's actual climate. Surveying these sensitivity distributions, Knutti and Hegerl[61] commented: "These distributions reflect the uncertainty in our knowledge of sensitivity, not a distribution from which future climate change is sampled."

The problems arising from the limitations of GCMs are far from academic. At the moment, they are the only technique available for projecting the climate to 2050 or 2100 and, because of their differing climates, each GCM makes significantly different projections, leading to large uncertainties. Figure 7.10 shows that, over the coming decades, the uncertainties are sufficiently large that widely divergent scenarios of economic development lead to significantly overlapping projections. Both high and low emission pathways have non-negligible probabilities of staying within the agreed 2°C threshold for dangerous effects. If we use (standard) 90% confidence ranges, then in 2050, any of the scenarios used in IPCC AR5 (2013) could be claimed to be compatible with any threshold in the range 1.85 to 2.2°C.
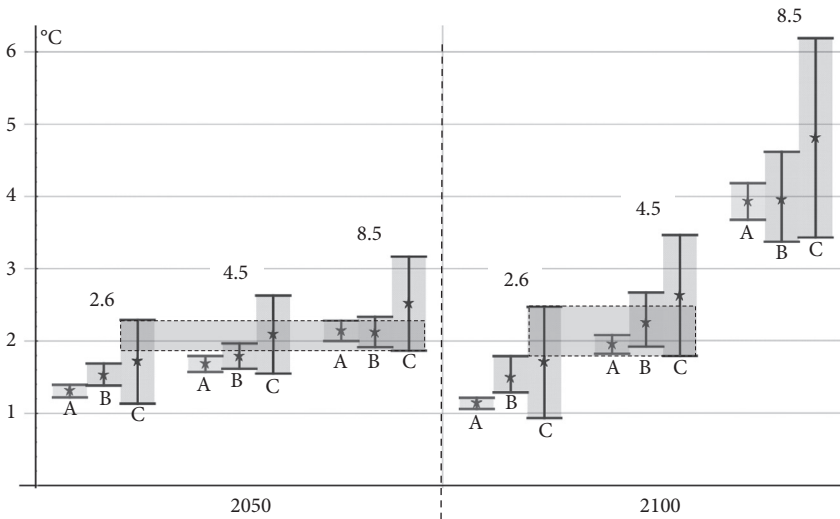


FIGURE 7.10 *A comparison of the Coupled Model Intercomparison Project 5 projections ("C"), the simple historical method (five-year memory, "A"), and the generalized historical method using scaling climate response function projections ("B") for 2050 (left) and 2100 (right) for three representative carbon pathway (RCP) scenarios, with the numbers indicating the watts per square meter of the scenario forcing in 2100. The vertical bars indicate 90% confidence intervals. For each projection date, the horizontal gray rectangles (bounded by the dashed lines) show the regions of overlap of the GCM ("C") projections. For 2050, all the scenarios are compatible with a warming anywhere in the range of 1.85 to 2.2°C. For 2100, with the exception of the 2100 RCP8.5 projection, all scenarios are compatible with a warming anywhere in the range of 1.8 to 2.5°C. In comparison, for a given date, both historical methods ("A" and "B") give (nearly) completely distinct projections, allowing policy choices (the RCPs) to be made according to the desired outcomes (temperatures). For the year 2100, the GCMs ("C") can exclude only the "business as usual" (RCP8.5) scenario. In contrast, the historical projections ("A" and "B") have projected ranges that are sufficiently narrow that they separate policies (RCPs) and consequences (temperature ranges) cleanly.[62]*

At the moment, policies ranging from very low to very high emissions—
RCP2.6 to RCP8.5—a range of more than a factor of 3 (see Fig. 7.8 and Box 7.5)
could all be said to be compatible with any of these thresholds. Policy choices
(emissions) do not map uniquely onto outcomes (temperatures; Fig. 7.10). But,
there is a further *qualitative* uncertainty: IPCC projections rely on a unique
(GCM) modeling approach, and this is not perfect. From the point of view of
scientific methodology, this reliance is a weakness that cannot be redressed
without the development of qualitatively different approaches. This much lower
uncertainty is possible and emphasized in Box 7.6, which shows that, using the
alternative discussed later, a one hundred-year "projection" made in 1909 could
have been extremely accurate. In the next sections, we look at how these histor-
ical (and scaling-based!) alternatives can reduce uncertainties while providing a
second path to projections.

---

BOX 7.6  The International Committee for Projecting the Consequences of Coal
Consumption: Hindprojections with Representative Coal and Petroleum Pathway 1.45

---

In the 1896 article in which Arrhenius estimated that doubling atmospheric $CO_2$
would increase global temperatures by 5 to 6°C, he commented that over the next
3,000 years, coal burning would lead to a 50% increase of $CO_2$ levels. Later that year
he noted: "We would then have some right to indulge in the pleasant belief that
our descendants, albeit after many generations, might live under a milder sky and
in less barren surroundings than is our lot at present."[63] Later, in 1908 (at a time of
much higher rates of global coal consumption), in a popular book, he revised this
millennial estimate downward to centuries.

For fun, let's imagine what might have happened if this idea had been seized
upon by his heat-starved Scandinavian colleagues. Let's say they possessed a
modern estimate of the EffCS (2.33 ± 0.36°C/$CO_2$ doubling, 90% confidence[64]), a
modern estimate of the fraction of the $CO_2$ emissions that end up in the atmosphere
(about[gg] 43%), and reliable global data since 1880. Looking back, an early 21st
century historian might have written the following account:

*To investigate global warming, in 1909, Arhennius and his colleagues constituted*
*the International Committee for Projecting the Consequences of Coal Consumption*
*(ICPCCC). The committee economists drew up a series of scenarios, the "representative*
*coal and petroleum pathways" (RCPPs), and numbered them by their radiative forcing*
*in the year 2000. Although Arhennius had originally only mentioned coal, petroleum*

---

[gg] Knowledge of the atmospheric fraction is needed to estimate the relationship between $CO_2$ emis-
sions and the increase in atmospheric concentrations. The rest of the $CO_2$ ends up in biomass (12%) and
in the oceans (45%), where it contributes to ocean acidification, See DeVries, T., Holzer, M., & Primeau,
F.,Recent increase in oceanic carbon uptake driven by weaker upper-ocean overturning. *Nature* **542**,
215–218 (2017).

*was included because it had already become clear that it would soon overshadow coal as a source of energy.*

*Three RCPPs were proposed, corresponding to low, medium, and high economic growth: RCPP0.85, RCPP1.45, and RCPP2.75. The low- and high-growth RCPPs were believed to be overly pessimistic or overly optimistic, respectively, so policymakers focused on the intermediate one (RCPP1.45), which assumed an increase in radiative forcing of 1.45 W/m² by the year 2000.*

*Using the $CO_2$ concentrations from the RCPP1.45 scenario, ICPCCC scientists made global temperature predictions using a new technique called StatCLIP (Statistical Climate Prediction). These predictions were based on the $CO_2$ concentration scenario and on the preceding thirty years of global data. These data were the oldest that were judged reliable. The otherwise arbitrary use of thirty years as a "regular climate" period was justified, because it was believed to be a convenient duration over which to average the temperature and gauge the state of the planet. The StatCLIP result was a genuine prediction conditioned only on the $CO_2$ emission scenario. It predicted both the macroweather variability as well as the temperature response to future atmospheric $CO_2$ concentrations.*

*At the time, it was expected that other consequences of coal burning—notably, aerosols—would also affect the predictions and, in addition, it was expected there would be errors resulting from the time delay associated with the ocean absorbing the heating and then re-emitting it years later. Pointing this out, some of the more sophisticated committee members derided the ICPCCC approach as simplistic. Extremists even claimed that atmospheric prediction was not even possible and ridiculed the entire enterprise.*[hh]

*Today, this daring ICPCCC focus on $CO_2$ can be checked by comparing their predictions with the temperatures over the following century. As it happened, the ICPCCC economists got lucky. The one hundred-year ICPCCC RCPP1.45 scenario was amazingly accurate. It ensured that the results of their one hundred-year prediction (shown in Fig. 7.11A) was accurate, with an uncertainty dominated by the global, annual resolution macroweather variability (i.e., about 0.2°C, which is much less than the 0.8–1°C rise over this period). Recognizing that the macroweather variability was a distraction, the ICPCCC scientists had also made a centennial-scale projection simply by using the EffCS and the scenario-based $CO_2$ concentration levels (Fig. 7.11B). By comparing it with the real data passed through a thirty-year smoother*[65] *(the thick line in Fig. 7.11B), it can be seen that the projection was surprisingly accurate. It was only 0.017°C too high and it had only a small spread (of 0.010°C). The uncertainty came from the spread in the estimates of the EffCS.*

---

[hh] Richardson's first attempt to use the equations of fluids was more than a decade away, and there were even serious doubts that it might be at all possible. See the discussion in Lynch, P. *The Emergence of Numerical Weather Prediction: Richardson's Dream*. (Cambridge University Press, 2006).
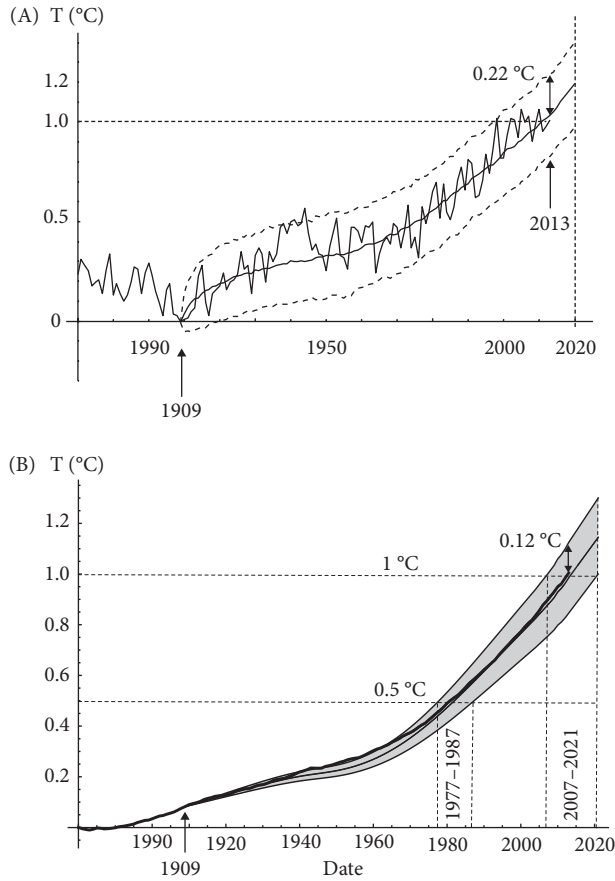
FIGURE 7.11  *(A) If, in 1909, one had a reliable scenario about the future carbon dioxide ($CO_2$) emissions, and one knew the globally, annual averaged temperature back to 1880 (This information mainly helps improve the forecast in the 1909 to 1919 decade. Beyond that, the hindcast depends almost only on knowledge of the $CO_2$ concentration. Then the global temperature could be predicted (the smooth middle line) to within 0.22°C with 95% certainty (see the dashed lines).[ii,66] (B) Hindprojecting the temperature from 1909 to 2021 using the effective climate sensitivity 2.33 ± 0.36 (upper and lower bounds) and a (thick) thirty-year running smoother (same $CO_2$ and temperature data as in Fig. 6.3). This is a "hindprojection" rather than a "hindcast," because a thirty-year smoother was used in an attempt to remove the high frequencies to isolate the responses to anthropogenic forcing.[67]*

In retrospect, the success of the ICPCCC prediction and projection was assured because the $CO_2$ turned out to be ⬚⬚y correlated with economic activity that was

---

[ii] This figure was tweeted in response to Scott Pruitt, who had just been sworn in as director of the US Environmental Protection Agency (March 2017). On the occasion, he was famously quoted as denying that there was a relationship between the levels of $CO_2$ and Earth's temperature.

highly fossil fuel dependent and, hence, with the other anthropogenic effects. It was an excellent surrogate for these.

After the ICPCCC projections were published, it started to become clear that the world was on a dangerous path. RCPP1.45 scenario indicated that the threshold of 0.5°C of warming would be exceeded between the years 1977 and 1987 (90% certainty), and that if measures were not taken, the threshold of 1°C might be reached as soon as 2007, but no later than 2021.[68] In comparison, the high-end[69] RCPP2.45 scenario predicted that 2°C would be exceeded sometime around the year 2010 (with a thirteen-year uncertainty either way).

Today's 1°C warming has already turned out to be dangerous enough; it is fortunate for us that the 1909 policymakers were sufficiently wise to reject the 2°C RCPP2.45 scenario. This was no mean feat. At the time, they had to resist well-financed lobbying efforts, particularly representatives of Standard Oil Company. This lobbying was based mainly on the unsubstantiated claim that almost all the fossil fuel $CO_2$ would end up being absorbed by the oceans and, consequently, would not increase atmospheric $CO_2$ concentrations. It was also claimed that, except over geological epochs, the climate was known to be constant. It is sobering to realize that from 2017 until he resigned in 2018, the ex-Chief Executive Officer[jj] of the descendant of Standard Oil (Exxon-Mobil) personally occupied one of the highest political positions in the world's most powerful economy, and he did everything in his power to block attempts at limiting emissions.

The most important point of this fable is that the main source of uncertainty in the ICPCCC projection came from the uncertainties in projecting the economy—not in uncertain climate science. Faced with the ICPCCC projections, early twentieth-century policymakers would at least have been given a clear choice. Based on whether they were ready to accept 0.5°C, 1°C, or 2°C of warming one hundred years in the future, they could choose the corresponding RCPP. This is unlike today's situation, in which the GCM-based IPCC projections have such large spreads that there is no longer a strong link between climate policies (emissions) and climate outcomes (temperatures; see Fig. 7.10); there is a looming "uncertainty crisis" (Box 7.7).

This example underscores an important advantage of the historical approach to projections, which are based on the relatively well-reconstructed historical $CO_2$ concentrations. This approach avoids assumptions about the (poorly reconstructed) aerosol levels, as well as debatable assumptions about the magnitudes of the direct and indirect aerosol radiative effects. However, it does make simplistic assumptions about the way the atmosphere–ocean memory works. It assumes that globally averaged temperatures react to the forcings over a period of only a few years.[70] In reality, some of the response is over timescales much further into the past. If we stopped emissions tomorrow, as a result of the "warming in the pipe" (e.g., in deep, slow ocean currents), the temperature would continue to rise for decades or longer (Box 7.1). These effects can be taken into account using a more sophisticated scaling-based historical method (see Box 7.8).

---

[jj] Although US Secretary of State Rex Tillerson no longer works officially for Exxon-Mobil, the company gave him a $180 million retirement package.

BOX 7.7  A looming uncertainty crisis

The range of 1.5 to 4.5 °C for a $CO_2$ doubling has been with us for so long that we are used to its imprecision, but the implications of this large uncertainty for projections and hence for policymaking are unacceptably large. For example, the COP 21 agreement (2015) included in its preamble the noble aim of keeping global temperatures to less than 1.5°C. When will we reach this level? Analysis of the CMIP5 GCMs used in the IPCC AR5 (Plate 7.5) shows that if we use a plausible emissions scenario, such as RCP4.5, that with 90% certainty the 1.5°C threshold will be exceeded between 2010 and 2050. This interval is so large that it can already be reduced simply by using the knowledge that today—several years after 2010—the increase is only about 1°C. This basic point has already been discussed by Andersen and Bows,[71] who pointed out that the amount of decarbonization required to remain within a 1.5°C or 2°C limit varies substantially, depending on which part of the GCM-derived probability distribution of climate sensitivity is used, underscoring the deleterious consequences of the considerable policy "wiggle room" allowed by this wide range of climate sensitivities. Without another approach, it seems unlikely that the large range of projected futures will be reduced.

How, then, can we justify the strongly negative emission policies needed to follow RCP2.6 if the consequences are difficult to disentangle from high-emission RCP4.5, which—even in 2100—has a significant overlap in the projected temperature ranges?

The current lack of alternatives has another deleterious policy consequence: the methodological consequences of relying on a single projection tool. In science, when only a single method or model predicts a given result, scientists are generally reserved, if not outright skeptical. Climate deniers posing as skeptics have skillfully exploited this situation to dismiss the projections and undermine public confidence not only in the projections, but also in the anthropogenic provenance of the warming itself. Policymaking flowing purely from GCM projections will be exposed politically and hence will be weak. However, using the qualitatively different historical projection method, the primary source of uncertainty is the reconstruction of past climate forcing rather than with the model dynamics. With some moderate (but interesting) differences, the new method gives reasonably similar projections. Therefore, our confidence in the common conclusions takes a qualitative leap. Skeptics now must reject two scientifically independent approaches.

Plate 7.5 shows the 1.5°C and 2°C thresholds that, according to international agreements, we should, respectively, *aim* to stay below and *actually* stay below.[kk] Comparing the implications of the historical projections' lower mean and smaller uncertainty for RCP4.5, we see that the 1.5°C threshold will be crossed sometime in 2034 to 2052 (GCMs) and 2044 to 2052 (historical) (with 90% certainty), and that

---

[kk] Neither goals were chosen on scientific grounds, but both on political grounds. The 1.5°C threshold was inserted at the last minute into the COP 21 (2015) preamble as a laudable goal, but did not figure in the convention itself (which was, in any case, nonbinding) (https://unfccc.int/files/meetings/paris_nov_2015/application/pdf/paris_agreement_english_.pdf). The 2°C goal was ordained officially at the COP 16 meeting (held in Copenhagen December 7–18, 2010) as an agreed threshold that should not be crossed.

the 2°C threshold will be crossed between 2040 and 2100 (GCMs) and between 2070 and 2100 (historical).

The historical method has a further advantage: its simplicity. GCMs are highly complex constructs, built by generations of scientists, each of whom contributed (at most) modules emulating specific climate processes. Few, if any, completely understand the resulting fully integrated system. For the vast majority of GCM users, they are "black boxes." Because the final projections are not even based on a single GCM, but on an MME, to evaluate the ensemble projections properly requires an understanding of the detailed similarities and differences between all the models in the ensemble. In comparison, the historical method is simple and can be understood easily. It is graphically represented in Figures 6.4B and 7.9. The anthropogenic effect is approximately linear with $CO_2$ forcing, whereas deviations represent natural variability. The basic method also works on the GCMs, proving that it is not exotic and respects GCM physics. Finally, the method can be improved to take into account its main limitations: its treatment of system memory (Box 7.8) and its handling of aerosols.

---

BOX 7.8  Using scaling to improve the memory of the historical approach

The simple historical method had shortcomings. First, it ignored the long-term system memory, especially as a result of the fact that most of the heat goes into the oceans, so there is a delay in atmospheric response. Up until the present, long-term memory effects have not been so important. This is because—as a result of rising emissions—the forcing was constantly increasing and the responses to recent forcing effectively swamped the lingering (old) responses to past forcing.[72] However, once we start mitigating the emissions, the memory effect will be more evident. This is why the simple historical method generally worked well on the MME scenario runs, except for the IPCC RCP2.6 scenario—the only one that used sufficiently massive negative emissions to reverse the forcing trend.

Another weakness of the simple historical method is that it assumed implicitly that however much aerosols contribute to the forcing, their contribution remains a fairly constant fraction of the total. If there is a likely "cleanup" over the upcoming decades (especially over Asia), then, for regional projections, we need to account separately for the aerosols. The RCPs assume that the aerosol forcing pattern evolves as western countries and then China clean up their emissions; it will take longer for India and Africa. This explains some of the discrepancy observed between the EffCS GCM regional projections and their actual integrations.

The basic method to overcome these limitations is to use a climate response function (CRF) that is scaling: an SCRF (Fig. 7.12, the curve).[73] Although the simplest historical CRF has one parameter—the effective sensitivity to $CO_{2eq}$ doubling—the SCRF has two extra parameters: a small-scale cutoff estimated at two years (the shortest timescale for which the linearity assumption holds) and the key scaling (power law) exponent,[74] which determines the long-term memory. To estimate it, we need to give special consideration to volcanic[75] and aerosol forcing.[76]
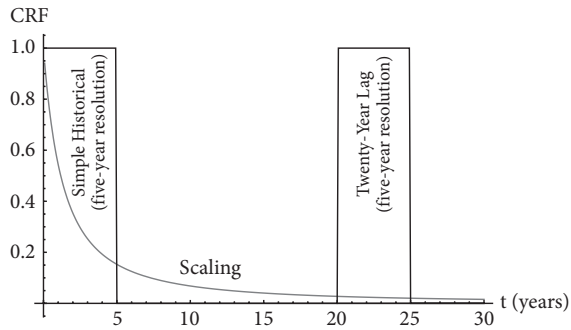
FIGURE 7.12 *Examples of the shapes of various linear climate response functions (CRFs). The rectangles are the simple historical responses at five-year resolutions (their widths). The one on the left is without a lag and the one on the right has a twenty-year lag, meaning that today's response depends only on the forcing twenty to twenty-five years ago. The lower curve is the scaling response with exponent $H_R = -0.4$ truncated at two years, the ocean atmosphere coupling scale. In this figure, only the relative shapes are important. The actual CRF is obtained by multiplying these by a climate sensitivity factor.*

For the (most plausible) RCP4.5 scenario (Plate 7.5), the results of these improvements are shown with extra embellishments. We see that with respect to the previous direct historical projection method, the mean projection is a little higher (but only about 15% less than the CMIP5 MME, compared to 25% for the historical method), and that the spread is a little wider than before (but it is still about half that of the CMIP5 MME). As before, the origin of the uncertainty in the SCRF method is in the historical reconstructions.[77] It is larger than before, mainly because aerosol emissions are so uncertain. Another important contribution to the uncertainty comes from the scaling (memory) exponent that determines the importance of the long-term memory. Its effect is more apparent when the forcing relaxes, as in the IPCC RCP2.6 scenario. In Plate 7.5, we can see that the SCRF method—unlike the CMIP5 GCMs—reproduces the 1998 to 2015 "pause" well (when a natural cooling offset anthropogenic warming temporarily (see Box 6.2[78]). For reference, Plate 7.5 also shows the 1.5°C threshold. We can see that the range of dates when it is projected to be exceeded is much narrower for the SCRF (2030–2042) compared to the large range of the MME (2010–2052; see Fig. 7.10).

### 7.3.2 PROJECTING TO 2100: THE HISTORICAL METHOD

We saw that projections require a model that can distinguish the internal variability from anthropogenic responses. GCMs do this by attempting to take everything into account, integrating basic equations at subhourly time steps and then averaging the result over a factor of 100,000 in time (decades) as well as over many realizations, each differing by slightly different initial conditions (the stochastic ensemble). The alternative to GCMs was already discussed in Chapter 6. It was to

exploit the linearity approximation, justified theoretically by the fact that anthropogenic forcings are about 1% of the average total solar power, justified empirically by the observed linearity of the temperature–forcing relationship (Figs. 6.4A and 7.9), and confirmed impressively by the hindprojection discussed in Box 7.6.

The linear framework is not as restrictive as it might seem at first sight. It simply states that the future response is a combination of past forcings, each weighted by an amount that depends on how far in the past the forcing occurred. The weights as a function of their lags are called the CRF (see Fig. 7.12 for examples).

The linear framework uses past data weighted by the CRF. It is a general "historical" method equivalent to specifying an infinite number of weights. Without additional simplifying assumptions, it is intractable. In Chapter 6 we examined the simplest historical assumption: to assume the responses were essentially "instantaneous," albeit using five-year-resolution forcing data. Figure 7.12 graphically shows these weights[79]; they fall to zero for times longer than five years, which is not too realistic. A simple alternative is to lag the response. The right-hand rectangle in Figure 7.12 shows an example with a twenty-year lag also at five-year resolution, so only the forcings at twenty to twenty-five years earlier are important.[80]

You should not be surprised to learn that the best way to make the CRF tractable is to constrain it by temporal scaling symmetry. The resulting SCRF is essentially a new emergent law that describes how the globally averaged temperature responds to external forcings. It can be justified here in a direct way by appealing to the scaling of the forced response or, more convincingly, via a scaling energy storage mechanism (Box 7.1). In the latter case, the FEBE determines both the scaling of the internal (macroweather) variability and the externally forced (climate) variability. However, there are various technical issues that make the SCRF a bit cumbersome, so let's first examine the consequences of the simple historical assumption that worked so well on the globally averaged industrial epoch temperatures described in Chapter 6 and in Box 7.6.

Let's compare our projections to the IPCC's CMIP5 GCM simulations. Instead of the $CO_2$ forcing, let's use the IPCC $CO_{2eq}$ forcing, which Figure 6.5 showed was about 1.12 times larger than the pure $CO_2$ forcing. From 1750 to the present, the forcings are reconstructions from a variety of sources that cover both anthropogenic reconstructions (notably of past $CO_2$, $CH_4$, and aerosols) and reconstructed solar and volcanic forcings (Box 7.5). Figure 7.9

shows the plot analogous to Figure 6.4A, confirming the near linearity of the temperature response, but with a slightly different effective sensitivity to doubling[81] of $CO_{2eq}$: 2.02 ± 0.08.[82] Figure 7.9 also shows the corresponding results for the CMIP5 MME, demonstrating that it overestimated the warming in the historical period by about 15% (we return to this later).

### 7.3.3 GLOBAL-SCALE PROJECTIONS

To project the temperatures into the future, the IPCC created four emission scenarios, as mentioned earlier, called RCPs (Box 7.5). The RCPs are numbered

by the total climate forcings (measured in watts per square meter) that the scenario assumes for the year 2100. For example, the most benign, RCP2.6, assumes that in 2100 there will be 2.6 $W/m^2$ of anthropogenic forcing, whereas the most dangerous—RCP8.5—assumes more than three times as much.

Following the method of Figure 7.9 (using $CO_{2eq}$) we obtain Plates 7.5 through 7.7, which compare this historically based method (labeled TCS for "transient climate sensitivity," based on $CO_{2eq}$ forcing[83]) with the CMIP5 GCM projections obtained in the usual way—by integrating numerically the coupled atmosphere–ocean models. The figure shows both the mean[84] and the spreads (90% confidence), representing the uncertainties. We can draw several conclusions. The first is that the mean of this simple historical method is within the 90% confidence level of the GCM projections. Statistically, the two methods agree with each other. The second is that the mean GCM projections are about 25% greater than the historical method mean. Third, the uncertainty of the historical method is reduced by a large factor (about 5).

Detailed analysis suggests that the MME is a too high because the GCM parameters were adjusted using data over past decades, which involved overly strong aerosol cooling, which masked the fact that the model climate sensitivities were too high. This would not matter if, in the future, the proportion of aerosol to total greenhouse gas forcings was the same. However the RCPs assume that aerosols (pollution) will soon be largely "cleaned up." Therefore, with little future negative forcing (cooling), the model oversensitivities become more apparent. As for the uncertainties, the reason why the historical projections have much smaller uncertainties (spreads) than the GCMs is that they are projected from the real world—not GCM climates. Their main uncertainty is from the temperature and forcing reconstructions, which lead to uncertainties in overall climate sensitivity. In comparison, the GCM projections use common reconstructions as inputs. Their projection uncertainties arise because each GCM has its own climate, and the uncertainty reflects this large spread (see Box 7.8 for more details).

Also indicated in the plot are the 1.5°C and 2°C thresholds that, according to international agreements, we should *aim* to stay below and *actually* stay below, respectively. We can now compare the GCM and historical-based projections to the implications of the historical projections' lower mean and smaller uncertainty. Avoiding the unrealistically extreme RCPs, we see that for RCP4.5, with 90% certainty, the 1.5°C threshold will be crossed in 2034 to 2052 (GCMs) and in 2044 to 2052 (historical). The 2°C threshold will be crossed in 2040 to 2100 (GCMs) and 2070 to 2100 (historical).

### 7.3.4  REGIONAL PROJECTIONS

We can now see how well the historical method does at a regional scales. Plate 7.8 (top left) shows the result when the regression method of Figure 7.9 is applied at each grid point to estimate the local TCS. Overall, we see that nearly all the grid points have warmed, especially at the high latitudes. The only exceptions are two

cooling areas off Greenland and near Antarctica (blue), which is likely a result of a stratification of the ocean, with freshwater flowing from the accelerated melting of the ice sheets.[85] Because we are currently at about one half of a doubling of $CO_{2eq}$, the plot is easy to interpret: By dividing the numbers in the figure by two, we obtain an estimate of the historical warming at each grid point. This projection just says that whatever we've had in the past will continue in the future.

We can compare this historical estimate of past warming to the CMIP5 estimates (Plate 7.8, top right). Plate 7.8 (bottom) shows there is a large spread between the approaches.[86] Comparing the GCM sensitivities to the historical ones (Plate 7.8, top), we see that the GCMs have overestimated the warming since 1880 over most of the globe, but they also underestimated it over some significant regions, such as Northwest Canada and Central Asia. This is confirmed directly in Plate 7.8 (bottom), which shows the map of the difference between them. For more than half of the grid points, the GCMs and the historical data are statistically in disagreement. The GCMs reproduced the past poorly, and at the regional level even more so. This gives us confidence in the slightly lower historical projections made by extrapolating the historical sensitivities into the future (Plate 7.8, upper left).

We can now use TCSs at each pixel to project the temperatures to the period 2080 to 2100 for each of the RCP scenarios (Plate 7.9). For the historical method, the spatial patterns are all the same because the forcings are all applied globally. The difference between the RCP scenarios is only in the total $CO_{2eq}$. In comparison, the CMIP5 projections have some spatial variation, notably because they involve a stronger reduction in aerosols over China and the Indian subcontinent. These differences have important policy consequences (discussed in Box 7.7).

# Notes

1. Lorenz, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).

2. Roughly the lifetime of planetary structures. See Section 7.2 for this inverse cascade of error growth.

3. Mathematically, the logistic model is a "mapping." It evolves in discrete time. This means the model determines the state only at specific intervals of time, not at in-between times. [Lorenz, E. N. The problem of deducing the climate from the governing equations. *Tellus* **16**, 1–11 (1964).] As he was studying logistic mapping, Lorenz was also developing nonlinear models of systems that were continuous in time (mathematically, "flows"). Notably, he discovered the famous "Lorenz model": Lorenz, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).

4. It can be mapped onto the real axis of the Mandelbrot set by a nonlinear transformation of variables.

5. So that 68% of all the series are between the top and bottom red curves in Plate 7.1B.

6. In real forecasting using GCMs, the long-term error is often not well estimated, so they often have negative skill (see Section 7.3).

7. In this case, $\gamma_n$ is simply a random number drawn independently of the previous ones. Conventionally, the innovations are chosen to be from the bell curve.

8. Box, G. E. P. & Jenkins, G.M. *Time Series Analysis: Forecasting and Control.* (Prentice Hall PTR, 1970).

9. It falls off exponentially, which is faster than any power law.

10. A generalization to processes with power law memories does exist (the fractional autoregressive integrated moving average, or FARIMA), but most practitioners consider it to be exotic; it is not much used.

11. This model is linear. At long-enough timescales—centuries or millennia—there are nonlinear couplings between forcings and albedo. Also I'll only discuss the simplest ~~ap-proximation in which~~ Earth ~~is treated as a uniform body so that spatial~~ temperature ~~(and other) variations are ignored (the model is "zero dimensional")~~.

12. The value of this relaxation time is hard to measure and is not well known, but is likely to be about two years, because this is the time required for the oceans and atmosphere to be strongly coupled.

13. Another mathematical difference between the EBE and FEBE is that, as a result of the long memory, the relevant FEBE is a "past value problem," not the usual "initial value" problem.

14. For an fGn with parameter $H$, the weights of the innovations a time $j$ steps in the past fall off as $j^{(H-1/2)}$. Because the statistics depend on the variance (the second moment), the average importance of the innovations (their statistical weight) decays more quickly as $j^{(2H-1)}$. However, the statistical weight of all the innovations that are older than $j$ steps in the past only fall off as $j^{2H}$. In this example, $H = -0.1$, so the influence of the ancient past decays extremely slowly. Note that when $H > 0$, it diverges. It turns out that the basic $H > 0$ fBm model can nevertheless be saved by effectively subtracting out the infinity.

15. Lorenz, E. N. The predictability of a flow which possesses many scales of motion. *Tellus* **21**, 289–307 (1969).

16. Interestingly, Lorenz's article ("The predictability of flow") uses a statistical turbulent approach that assumes planetary-scale scaling, but the article is not well known. For a review, including a generalization to account for the effect of intermittency, (causing the predictability to be lost in "bursts") see Schertzer, D. & Lovejoy, S. Uncertainty and predictability in geophysics: Chaos and multifractal insights. In: *State of the Planet: Frontiers and Challenges in Geophysics* (eds. R. S. J. Sparks & C. J. Hawkesworth), pp. 317–334 (2004).

17. Recall from Chapter 4: Lifetime = (Energy rate density)$^{-1/3}$(Size)$^{2/3}$.

18. The overall time for the planetary-scale prediction to be corrupted is, in fact, the sum of the times for the smaller structures to have their predictions corrupted. However, the critical prediction time for structures of a given spatial scale (equal to their "lifetime") grows as the two-thirds power of their size, so that for all practical purposes, it is only the lifetime of the largest structures that matters. If the large-scale turbulence had been two dimensional—as is conventionally assumed—then in principle, the predictability would be much longer. This is because in two dimensions, the lifetime of structures only grows logarithmically with size, so the error propagation to larger scales is much slower.

19. The exact status of STSF is not clear, although empirically it is supported by temperature and precipitation data, and there are theoretical arguments that it holds at least as an approximation. See Lovejoy, S. & Schertzer, D. *The Weather and Climate: Emergent Laws and Multifractal Cascades.* (Cambridge University Press, 2013); Lovejoy, S. & de Lima, M. I. P. The joint space–time statistics of macroweather precipitation, space–time statistical factorization and macroweather models. *Chaos* **25**, 075410, doi:10.1063/1.4927223 (2015).

20. Hasselmann, K. Stochastic climate models, part I: Theory. *Tellus* **28**, 473–485 (1976).

21. The latter is somewhat a misnomer because it excludes ~~restrictively~~ fractional ordered (but still linear) terms, yet these fractional terms are crucial for obtaining realistic scaling. See, for example, Sardeshmukh, P., Compo, G. P., & Penland, C. Changes in probability associated with El Nino. *J. Climate* **13**, 4268–4286 (2000); Penland, C. A stochastic model of IndoPacific sea surface temperature anomalies. *Physica D* **98**, 534–558 (1996); Penland, C. & Sardeshmuhk, P. D. The optimal growth of tropical sea surface temperature anomalies. *J. Climate* **8**, 1999–2024 (1995); and Newman, M. An empirical benchmark for decadal forecasts of global surface temperature anomalies. *J Climate* **26**, 5260–5269 (2013).

22. Newman, M. An empirical benchmark for decadal forecasts of global surface temperature anomalies. *J Climate* **26**, 5260–5269 (2013).

23. Suckling, E. B., Jan van Oldenborgh, G., Eden, J. M., & Hawkins, E. An empirical model for probabilistic decadal prediction: A global analysis. *Climate Dynam.* 48, 3115–3138 (2017).

24. Lovejoy, S., del Rio Amador, L., & Hébert, R. The ScaLIng Macroweather Model (SLIMM): Using scaling to forecast global-scale macroweather from months to decades. *Earth Syst. Dynam.* **6**, 1–22 (2015).

25. Baillie, R. T. & Chung, S.-K. Modeling and forecasting from trend-stationary long memory models with applications to climatology. *Intl. J. Forecasting* **18**, 215–226 (2002).

26. fGn naturally involves strong low-frequency undulations and we have to try to distinguish these macroweather variations from the anthropogenic climate variations and forecast the two separately (see the next section).

27. Lovejoy, S., del Rio Amador, L., & Hébert, R. The ScaLIng Macroweather Model (SLIMM): Using scaling to forecast global-scale macroweather from months to decades. *Earth Syst. Dynam.* **6**, 1–22 (2015).

28. Lovejoy, S. Using scaling for macroweather forecasting including the pause. *Geophys. Res. Lett.* **42**, 7148–7155 (2015).

29. For the anomaly averaged over three years.

30. Adapted from Lovejoy, S. Using scaling for macroweather forecasting including the pause. *Geophys. Res. Lett.* **42**, 7148–7155 (2015).

31. Gripenberg, G. & Norros, I. On the prediction of fractional Brownian motion *J. Appl. Prob.* **33**, 400–410 (1996). This was investigated numerically in Hirchoren, G. A. & D'attellis, C. E. Estimation of fractal signals, using wavelets and filter banks. *IEEE Trans. Sign. Proc.* **46**, 1624–1630 (1998).

32. Technically, we minimize the square of the prediction error. We obtain a "minimum square" estimator.

33. Gripenberg, G. & Norros, I. On the prediction of fractional Brownian motion *J. Appl. Prob.* **33**, 400–410 (1996).

34. Reproduced from Del Rio Amador, L. *The Stochastic Seasonal to Interannual Prediction System.* PhD thesis, McGill University (2018).

35. With regard to the range, 0.09 indicates the spread of the *H* values for the thirty-six different models, and it is quite small. The actual range of values was –0.02 to –0.2, which is the model-to-model variation in *H*.

36. ~~One for each of the available periods in the control,~~ typically for each control run~~.~~ ~~Th~~ere were several thousand hindcasts; the MSSS is an average over all of them.

37. Figure 7.5 compares the mean and spread of the hindcasts, and the theory on the control runs. Although it cannot be deduced from this figure, further analysis shows that each individual control run was hindcast nearly to its theoretical limit.

38. Reproduced from Del Rio Amador, L. *The Stochastic Seasonal to Interannual Prediction System.* PhD thesis, McGill University (2018).

39. With globally averaged temperatures, the annual cycle is very small and, in anomaly data, it is very small.

40. Using the National Center for Environmental Prediction reanalysis data (at 5° × 5° resolution).

41. The annual cycles and low frequency trends use data over the previous thirty years so that, for example, the 1980 cycle and trend are estimated from data over the period 1950 to 1980.

42. Reproduced from Del Rio Amador, L. *The Stochastic Seasonal to Interannual Prediction System.* PhD thesis, McGill University (2018).

43. The trend in Figure 7.6 was deduced directly from the previous thirty years of temperature data. It turns out that a slight improvement can be made by, instead, estimating anthropogenic warming at each point using the $CO_2$ forcing and the observed response at the point. This improvement has been introduced in more recent versions of the StocSIPS.

44. The theoretical skill depends only on the value of *H*, which varies somewhat from place to place (Fig. 5.5). See the review by Lovejoy, S., Del Rio Amador, L., & Hébert, R. Harnessing butterflies: Theory and practice of the Stochastic Seasonal to Interannual Prediction System (StocSIPS), In: *Nonlinear Advances in Geosciences* (ed. A. A. Tsonis), pp. 305–355. (Springer Nature, 2017).

45. Merryfield, W. J., Denis, B., Fontecilla, J.- S., Lee, W.- S,. Kharin, S., Hodgson, J., & Archambault, B. *The Canadian Seasonal to Interannual Prediction System (CanSIPS): An Overview of Its Design and Operational Implementation.* (Environment Canada, 2011).

46. The assimilation in this case is highly sophisticated. It is called "4D Var" for four-dimensional (space and time) variational method. It uses the equations of the atmosphere (as embodied in the model) to "ingest" all available data with meteorological information (i.e., station data, but also data from aircraft and satellites). Because each model has different and rather specific algorithms, reanalyses based on different models are somewhat different.

47. This is done with the standard method that uses the previous thirty years of data, removing the annual cycle, as described in Chapter 5.

48. It may be helpful to write this symbolically. First, define the model climates and anomalies from the actuals:

$$T_{CanSIPS}(\underline{r},t) = \overline{T}_{i(t),CanSIPS}(\underline{r}) + T'_{CanSIPS}(\underline{r},t)$$

$$T_{ECMWF}(\underline{r},t) = \overline{T}_{i(t),ECMWF}(\underline{r}) + T'_{ECMWF}(\underline{r},t),$$

where the overbar represents the climatological temperature $\overline{T}_i(\underline{r})$ at position (grid point) $\underline{r}$; for the month number $i = 1, 2, \ldots, 12$; and the primes indicate the anomalies, which are functions of both position and time [$i(t)$ denotes the month number of time $t$]; $T_{ECMWF}(\underline{r},t)$ and $T_{CanSIPS}(\underline{r},t)$ are the real temperatures, the "actuals." The conventional way to define $\overline{T}_i(\underline{r})$ is to use the averages over the previous thirty $i^{th}$ months (at each location/pixel $\underline{r}$).

The CanSIPS initializing temperature $T_{CanSIPS}(\underline{r},0)$ uses the ECMWF anomaly at time $t = 0$:

$$T_{CanSIPS}(\underline{r},0) = \overline{T}_{i(0),CanSIPS}(\underline{r}) + T'_{ECMWF}(\underline{r},0).$$

49. Unfortunately, five years is too short to estimate the trend properly (the true trends are buried in the macroweather noise until a decade or so in scale), so some of the internal five-year variability is thus spuriously removed in the postprocessing. Merryfield, W. J., Denis, B., Fontecilla, J.- S., Lee, W.- S,. Kharin, S., Hodgson, J., & Archambault, B. *The Canadian Seasonal to Interannual Prediction System (CanSIPS): An Overview of Its Design and Operational Implementation.* (Environment Canada, 2011).

50. Depending on the location, it varies from several months to several years.

51. Reproduced from Del Rio Amador, L. *The Stochastic Seasonal to Interannual Prediction System.* PhD thesis, McGill University (2018).

52. Although we have not mentioned it, StocSIPS actually provides forecasts of the probability distributions [both mean (discussed up until now) and the standard deviation about the mean]. These are used for various probabilistic forecasts.

53. Low-intermittency, temporal macroweather scaling with $-1/2 < H < 0$ and space–time statistical factorization. See Lovejoy, S. & de Lima, M. I. P. The joint space–time statistics of macroweather precipitation, space–time statistical factorization and macroweather models. *Chaos* **25**, 075410, doi:10.1063/1.4927223 (2015).

54. Changes in land use also contribute, but to a much lesser degree.

55. For aerosols, this includes the geographical distribution of the sources.

56. Not surprisingly, the fourth scenario (RCP6.0) gives projections in between RCP4.5 and RCP8.5. However, fewer CMIP5 runs were available for this scenario and we did not pursue it.

57. Adapted from Hébert, R., Lovejoy, S., & Tremblay, B. An observation-based scaling model for climate sensitivity estimates and global projections to 2100. *Climate Dynam.* (under revision).

58. Recall that the TCS is the proportionality constant between the historical temperature response and the total ($CO_{2eq}$) forcing. In comparison, the EffCS is the corresponding constant with respect to the pure $CO_2$ forcing. Over the historical period, the TCS is about a factor 1.12 times smaller than the EffCS (Fig. 6.3).

59. Reproduced from Hébert, R., Lovejoy, S., & Tremblay, B. An observation-based scaling model for climate sensitivity estimates and global projections to 2100. *Climate Dynam.* (under revision).

60. See, for example, Shukla, J., Palmer, T. N., Hagedorn, R., Hoskins, B., Kinter, J., Marotzke, J., & Miller, M., & Slingo, J. S. Toward a new generation of world climate research and computing facilities. *Bull. Amer. Meteorol. Soc.* **91**, 1407–1412 (2010). For a critical discussion, see Katzav, J. W. & Parker, S. The future of climate modelling. *Climate Change* **132**, 475–487 (2015).

61. Myhre, G. Consistency between satellite-derived and modeled estimates of the direct aerosol effect. *Science* **325** (5937), 187–190 (2009).

62. Adapted from Hébert, R., Lovejoy, S., & Tremblay, B. An observation-based scaling model for climate sensitivity estimates and global projections to 2100. *Climate Dynam.* (under revision).

63. Cited in Sample, Ian. "The Father of Climate Change," *The Guardian*, June 30, 2005, https://www.theguardian.com/environment/2005/jun/30/climatechange.climatechangeenvironment2.

64. The 90% uncertainty limits imply the range: 1.97 to 2.69°C/$CO_2$ doubling.

65. This is equivalent to annually updating a thirty-year averaged climate state. Because there is only one Earth, the ensemble has a single member, so there is no ensemble averaging.

66. Adapted from Lovejoy, S. Using scaling for macroweather forecasting including the pause. *Geophys. Res. Lett.*, **42**, 7148–7155 (2015).

67. The year 1909 is taken as the zero point; it was a year of record-low residuals, so this choice leads to a particularly high systematic error. On average, the temperatures were 0.017°C above the mean projection. The spread of the projection was only 0.010°C (90% confidence).

68. It was indeed exceeded in 2015 (Fig. 7.11B).

69. The slow-growth RCPP0.85 would only have been projected to be about 0.5°C at that time (with a four-year uncertainty either way).

70. The $CO_2$ level was smoothed over a five-year timescale.

71. Andersen, K. & Bows, A. Beyond "dangerous" climate change: Emission scenarios for a new world. *Phil. Trans. R. Soc. A* **369**, 20–44 (2011).

72. Indeed, it can be shown mathematically that a forcing that is convex (upward curving) necessarily leads to a convex temperature response, both of which are observed.

73. The CRF dicussed here is an approximation to the FEBE discussed in Box 7.1. The ~~smallest~~ scale ~~for which any linear model is appropriate~~ is about two years: the ocean–atmosphere coupling scale. ~~Therefore, this minimum scale at which the power law memory holds is an important parameter in the model.~~

74. The FEBE (Box 7.1) shows that the energy storage mechanism—and the memory that it implies— is fundamentally the same for macroweather (internal variability) and externally forced (climate) variability. However, each can be approximated by a different scaling law, with the two laws (their exponents) related as indicated in Box 7.1. Here I am discussing the low-frequency power law memory valid for timescales longer than two years or so.

75. The monthly resolution volcanic forcings were too "spiky" (intermittent) and needed to be adjusted. More details are given in Hebert's thesis: Hebert, R. *A Scaling Model for the Forced Climate Variability in the Anthropocene.* MS thesis, McGill University (2017).

76. For the aerosols, rather than attempting to make detailed models of the direct and indirect radiative effects, we used the IPCC aerosol forcing normalized by a factor determined by historical observations. Fairly similar results were obtained by using the total global sulfate emissions as a linear proxy for the aerosol forcing. Hébert, R., Lovejoy, S., & Tremblay, B. An observation-based scaling model for climate sensitivity estimates and global projections to 2100. *Climate Dynam.* (under revision).

77. Previously, we didn't need explicit knowledge of the aerosol forcings because we assumed the aerosol fraction of the total forcing was constant. The uncertainty is the parametric uncertainty from the Bayesian parameter estimation.

78. Box, G. E. P. & Jenkins, G. M. *Time Series Analysis: Forecasting and Control.* (Prentice Hall PTR, 1970). Sardeshmukh, P., Compo, G. P., & Penland, C. Changes in probability associated with El Nino. *J. Climate* **13**, 4268–4286 (2000). Lovejoy, S., del Rio Amador, L., & Hébert, R. The ScaLIng Macroweather Model (SLIMM): Using scaling to forecast global-scale macroweather from months to decades. *Earth Syst. Dynam.* **6**, 1–22 (2015).

79. In addition, the $CO_2$ forcing was taken as a linear surrogate for all the forcings—the responses characterized by the EffCS to $CO_2$ doubling.

80. A statistical treatment of all such lags between zero and twenty years is found in Lovejoy, S. Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. *Climate Dynam.* **42**, 2339–2351 (2014).

81. Multiplying this value by 1.12 doesn't exactly match the EffCS for $CO_2$ doubling (which is 2.33; Fig. 6.4A), because the temperature data used here were slightly different. The value 1.12 was derived from the average of four different global temperature series. For details, see Hébert, R., Lovejoy, S., & Tremblay, B. An observation-based scaling model for climate sensitivity estimates and global projections to 2100. *Climate Dynam.* (under revision).

82. The spread in the historical-based projections is a result of the use of four different temperature reconstructions, each with slightly different EffCS.

83. The EffCS is the result of using $CO_2$ as a surrogate for all the forcings. The EffCS is also "transient" in the sense that it ignores the long-term memory in the system. The term TCS introduced here is the sensitivity estimate using all the forcings ($CO_{2eq}$), but—unlike the ECS—it ignores the memory.

84. For the GCMs, the central line is the average over all the thirty-two CMIP5 models: the MME mean.

85. For the cooling of the North Atlantic, Hansen et al. have proposed that it depends on the amount of freshwater melting from Greenland. This addition meltwater leads to stratification of the upper layer of the ocean, reducing the formation of deep water and thus reducing the amount of warmer deep water that can rise to the surface. This means the cooling trend can potentially sustain itself as global warming increases and that the Greenland ice sheet continues melting—that is, until it has completely melted. This cooling should not be reassuring, because it means that more heat is trapped in the deeper ocean waters and, most importantly, this injection of freshwater may eventually lead to a shutdown of the thermohaline circulation, resulting in an upsetting impact on the world's climates. Hansen, J., Sato, M., Hearty, P., Ruedy, R., Kelley, M., Masson-Delmotte, V., Russell, G., Tselioudis, G., Cao, J., Rignot, E., Velicogna, I., Kandiano, E., von Schuckmann, K., Kharecha, P., LeGrande, A. N., Bauer, M., & Lo, K.- W. Ice melt, sea level rise and superstorms: Evidence from paleoclimate data, climate modeling, and modern observations that 2°C global warming is highly dangerous. *Atmos. Chem. Phys.* doi:10.5194/acpd-15-20059-2015 (2015).

86. About ±25%, although this ratio has some spatial variability. Also, it is expected that the ratio will not be the same because the spread between GCMs is the spread between different GCMs, whereas the spread between the different data sets is the result of different observations and statistical errors.

# Conclusions: Richardson's dreams

From big to small, from fast to slow, we traveled through scales—through magnifications of billions in space and billions of billions in time. We looked at how the traditional scalebound approach singles out specific phenomena: structures at specific spatial scales with specific lifetimes. The approach attempts to understand each in a (scale) reductionist and (usually) deterministic manner. Yet it fails miserably to describe more than tiny portions of the actual variability, giving—at best—some qualitative insights. Viewing the big picture with the help of modern data, we saw that, quantitatively, the scalebound approach underestimates the variability by a factor of a million billion (Fig. 2.3A).

The alternative is the scaling approach, which attempts to understand and model the atmosphere over wide ranges of scale. This approach is based on space–time scale symmetry principles. It describes statistically the synergy of nonlinear processes that act collectively over wide ranges of scale. To apply the idea in space, we needed to generalize the notion of scale itself (Chapter 3)—notably, to be able to account for the stratification caused by gravity. The appropriate notion of scale is one that emerges as a consequence of strong nonlinear dynamics, rather than being imposed a priori from without. Applying scaling in time, we found that the familiar weather–climate dichotomy was missing a key middle regime: from ten days to twenty years. It is a weather, macroweather, climate *tri*chotomy.

When it comes to real atmospheric modeling, scientists have long realized the limits of the scalebound approach. When they "really need to know," they defer to NWP or GCMs, the embodiment of Richardson's dream of "weather prediction by numerical process." This is fortunate, because the NWPs and GCMs respect space–time scaling symmetries; without them, they would be hopelessly unrealistic. At least when used for their original purpose—weather prediction up to the ten-day deterministic predictability limit—respecting scaling allows them to be reasonably accurate.

But Richardson had another dream: of using higher level statistical scaling laws to understand and model the atmosphere over huge ranges of scales. Rather than modeling deterministically and explicitly as many of the bumps, wiggles, cells, eddies, and structures as possible, the high-level statistical approaches account

for the collective, turbulent statistical effects of billions of billions of them. The low-/high-level dichotomy is analogous to the statistical mechanics–continuum mechanics dichotomy that underlies all atmospheric modeling. In principle, one could use (low-level) statistical mechanics to account for the position and velocity of all the molecules in the atmosphere, but this is impractical. Instead, one uses the (higher level) laws of thermodynamics and continuum mechanics. Both levels are valid and are used according to the application. However, the continuum laws that describe individual eddies are not at the top of the hierarchy. There are even higher level laws that describe the statistics of large numbers of vortices.

Four decades ago, the nonlinear revolution promised insights and solutions. In the words of a prominent chaos, scaling, and fractal enthusiast, a new generation of scientists was exhorted to "junk your old equations and look for guidance in cloud's repeating patterns."[1] It was hoped that this would solve the turbulence problem that had taunted scientists since the nineteenth century. Since then, nonlinear geoscience and atmospheric and climate science largely parted company. On the one hand, by using the low-level laws and brute-force numerics, GCMs have been wildly successful; on the other hand, the nonlinear revolution has only lived up partially to its promise of solving turbulence. This book is an attempt to reunite these two historic strands, these two complementary theoretical levels of atmospheric science.

As detailed in these pages, the scaling paradigm has made progress. Using modern data and new data analysis techniques, and with the help of stochastic models, the paradigm has verified and quantified space–time scaling over wide ranges, and the scaling laws were found to be fundamental enough to apply to Mars as well as to Earth. An almost incidental byproduct was the vindication— nearly ninety years after it was first proposed—of Richardson's wide-range scaling 4/3 law of turbulent diffusion. Other successes include the extension of scaling to tame atmospheric stratification and the strong tendency for variability to be "spiky" (turbulent intermittency), which turned out to be multifractal. And all of this was linked to extreme black swan events via scaling mechanisms, including the multifractal butterfly effect (Box 3.1).

In Chapters 6 and 7, we examined how these scaling advances can be applied to practical problems—notably, the problem of global warming—testing (and rejecting) the skeptics' GNF hypothesis. We also saw how scaling can be used to exploit the newly discovered huge memory to make macroweather (monthly, seasonally, and annually) forecasts nearly up to (new) stochastic predictability limits. Finally, in the brave new world of the Anthropocene, we discovered how scaling can improve climate projections up to the year 2100.

## Note

1. Cvitanovic, P. Introduction. In: *Universality in Chaos* (ed P. Cvitanovic), pp. 3–34. (Adam Hilger, 1984); quote, p. 4.

# { LIST OF ABBREVIATIONS }

| | |
|---|---|
| AGU | American Geophysical Union |
| AMDAR | Aircraft Meteorological Data Relay |
| AMO | Atlantic multidecadal oscillation |
| AMS | American Meteorological Society |
| AR | Assessment Report (always accompanied by a number) |
| AR | autoregressive |
| BECCS | Bio-Energy Carbon Capture and Storage |
| $CaCO_3$ | calcium carbonate |
| CanCM$x$ | Canadian Climate Model $x$ |
| CanSIPS | Canadian Seasonal to Interannual Prediction System |
| CanStoc | CanSIPS–StocSIPS hybrid model |
| CCAF | Canadian Climate Action Fund |
| CCAR | Climate Change and Atmospheric Research Programme |
| CD | *Climate Dynamics* |
| CFCAS | Canadian Fund for Climate and Atmospheric Science |
| $CH_4$ | methane |
| CMIP3$x$ | Coupled Model Intercomparison Project Phase $x$ |
| COP | Conference of the Parties |
| $CO_2$ | carbon dioxide |
| CRF | climate response function |
| CSIRO | Commonwealth Scientific and Industrial Research Organisation |
| CVAS | climate variability across scales |
| DFA | detrended fluctuation analysis |
| DMS | dimethyl sulfide |
| EAH | Early Anthropocene Hypothesis |
| EBE | energy balance equation |
| ECMWF | European Centre for Medium-Range Weather Forecasting |
| ECS | equilibrium climate sensitivity |
| EffCS | effective climate sensitivity |
| EFS | Ensemble Forecasting Systems |
| EGS | European Geophysical Society |
| $^{18}O$ | oxygen 18 |
| ENSO | El Niño southern oscillation |
| EOF | empirical orthogonal function |
| EPICA | European Project for Ice Coring in Antarctica |
| EuroSIPS | European Seasonal to Interannual Prediction System |
| FARIMA | fractional autoregressive integrated moving average |

| fBm | fractional Brownian motion |
| FEBE | fractional energy balance equation |
| FFT | fast Fourier transform |
| fGn | fractional Gaussian noise |
| FSP | Fractals Sums of Pulses (model) |
| GASP | Global Atmospheric Sampling Program |
| GCM | General Circulation Model (sometimes and equivalently, Global Climate Model) |
| GDP | gross domestic product |
| GHG | greenhouse gas |
| GISS | Goddard Institute for Space Studies |
| GNF | giant natural fluctuation |
| GPS | global positioning satellite |
| GRIP | Greenland Ice Core Project |
| GRL1 | "Return Periods of Global Climate Fluctuations and the Pause" |
| GRL2 | "Using Scaling for Macroweather Forecasting Including the Pause" |
| GRL3 | "Giant Natural Fluctuation Models and Anthropogenic Warming" |
| GSI | generalize scale invariance |
| HadCRUT | Hadley Centre Climate Research Unit Temperatures |
| HAD4 | Hadley model, version 4 |
| ICPCCC | International Committee for Projecting the Consequences of Coal Consumption |
| IMO | International Meteorological Organization |
| IPCC | International Panel on Climate Change |
| IR | infrared |
| LIM | linear inverse modeling |
| MA | moving average |
| MME | multimodel ensemble |
| MOZAIC | Measurement of Ozone by Airbus In-service Aircraft |
| MSSS | mean square skill score |
| MTSAT | Multifunctional Transport Satellite |
| NAO | North Atlantic oscillation |
| NASA | National Aeronautics and Space Administration |
| NG | Nonlinear Geophysics (focus group) |
| NGO | nongovernmental organization |
| NOAA | National Oceanic and Atmospheric Administration |
| NVAG | nonlinear variability in geophysics |
| NWP | Numerical Weather Model |
| PAGES | Past Climate Change ~~(working group)~~ |
| PDO | Pacific decadal oscillation |
| PETM | Paleocene–Eocene Thermal Maximum |
| QBO | quasi-biennial oscillation |
| RCP | representative carbon pathways |
| RCPP | representative coal and petroleum pathways |

| | |
|---|---|
| RMS | root mean square (the square root of the average of the square) |
| SCRF | scaling climate response function |
| SLIMM | Scaling Macroweather Model |
| SOC | self-organized criticality |
| SST | sea surface temperature |
| StatCLIP | *Statistical Climate Prediction* |
| StocSIPS | Stochastic Seasonal to Interannual Prediction System |
| STSF | space–time statistical factorization |
| TAMDAR | Tropical Airborne Meteorological Reporting |
| TCR | transient climate response |
| TCS | transient climate sensitivity |
| $^{10}$Be | beryllium 10 |
| TRMM | Tropical Rainfall Measurement Mission |
| TSI | total solar irradiance |
| 20CR | Twentieth-Century Reanalysis |
| UNEP | United Nations Environment Programme |
| WMO | World Meteorological Organization |

# { GLOSSARY }

**aerosol forcing** ~~when~~ airborne particles, most importantly sulfates, lead to a negative forcing by reflecting solar energy

**anthropogenic forcing** all man-made environmental changes that affect the radiative balance of Earth, mainly through changes in the atmospheric composition of the Earth, emission of aerosols, and changes in land use

**anthropogenic variability** the temperature variability resulting directly from anthropogenic forcing, which is responsible for a large part of Earth's surface warming during the twentieth century

**assessment reports by the IPCC** published materials composed of the full scientific and technical assessment of climate change by the working groups of the IPCC. Five reports were published: in 1990, 1995, 2001, 2007, and 2013.

**bell curve** the informal name of the "Gaussian" (or "normal") probability distribution, so-called because its graph is in the shape of a bell. The bell curve is ubiquitous in statistics, but it has virtually no extremes and it is ~~usually~~ inappropriate for geo-applications.

**black swans** events that are so extreme that, according to standard (usually bell curve) theory, they should virtually never occur (their probability would be astronomically low). Originally, the term designated completely "out of the box"-type events, but increasingly the term has been used to designate events produced by processes with power probability tails. Several mechanisms have been proposed for generating such extremes, including the multifractal butterfly effect and self-organized criticality. They are typically associated with spatial and/or temporal scaling, they are also associated with multifractal phase transitions and divergence of moments.

**butterfly effect** Lorenz's colorful description of the property of nonlinear systems to be so sensitively dependent on their starting values that even a butterfly could alter their future evolution

**chaos** the converse of order, originally from Greek mythology. Various paradigms of chaos have been proposed, including deterministic and stochastic. In the former, randomness is only apparent, hiding a nonrandom rule or law, and manifested in sensitive dependence on the initial conditions (e.g., resulting from subjective measurement error) whereas in the latter, it is produced by an objectively random process or law.

**climate** the range of timescales beyond macroweather in which fluctuations grow in amplitude with increasing time interval. During the preindustrial epoch, it is probably longer than one hundred years (although the duration is variable from epoch to epoch and place to place), whereas during the industrial epoch, it is the scale at which the anthropogenic changes dominate the natural variability, which is currently about sixteen to eighteen years—a little less than the official climate-normal period of thirty years.

**climate forcing** the increase or decrease in the power per area (measured in watts per square meter) above or below a convenient long-term level. Climate forcing quantifies

the rate of energy accumulation or loss resulting from a forcing agent. Climate forcings change the equilibrium temperature of Earth.

**climate normal** thirty-year averaged atmospheric data, defined by fiat by the World Meteorological Organization. In the current epoch, it is close to the macroweather–climate transition scale, ~~so these~~ data can be used to define climate states.

**climate response function** relates climate forcing to a temperature response. It specifies how important old forcings are for today's response. A scaling climate response function specifies that the weights of old forcings fall off in a power law, scaling manner. Unlike ~~many approaches~~, it has no characteristic response time.

**climate state** the mean ~~climate~~ state around which macroweather variability fluctuates; an average over the critical macroweather–climate transition scale (or longer). It is conventional to approximate a climate state by taking an average over a thirty-year climate-normal period. For a GCM run in a control run (without any changes in the external forcing (including fixed atmospheric composition), at monthly and longer times, the model exhibits pure macroweather and the climate state is the average over an infinitely long run. It is different for each GCM; for a given GCM, it is different for each set of forcings.

**climate variability** ~~general term for climate change~~

**$CO_2$ equivalent** a way to reduce any climate forcing (usually measured in watts per square meter) to the forcing that would result from an equivalent carbon dioxide ($CO_2$) concentration in parts per million, generally using the approximate relationship that states that doubling the $CO_2$ concentration leads to an increase in forcing of about 3.71 W/m$^2$.

**control run** the output of a GCM when all the external parameters are fixed: no solar variability, no volcanism, no changes in atmospheric composition, no orbital changes. For scales longer than a month or so, control runs exhibit pure macroweather variability. They slowly converge to a climate state— the model's climate (which is different for each set of cliamte forcings and for each model).

**Coupled Model Intercomparison Project** a framework for comparing and evaluating global coupled ocean–atmosphere General Circulation Models from different modeling teams around the world. It began in 1995 under the auspices of the World Climate Research Programme's Working Group on Coupled Modeling.

**effective climate sensitivity** proportionality coefficient between the historical temperature series and the carbon dioxide ($CO_2$) forcing series; usually expressed in degrees Celsius per $CO_2$ doubling. It is useful because it is ~~much~~ more reliably ~~reconstructed~~ than the transient climate sensitivity (based on $CO_{2eq}$) and because the other anthropogenic forcings are highly correlated with it as a result of economic activity. It ignores the long-term memory.

**equilibrium climate sensitivity** the expected temperature change at equilibrium after a step doubling in carbon dioxide concentration; requires a climate model to be determined

**feedback** a retro-action loop in which a change in the system fosters other processes that can either reinforce the original change (positive feedback) or weaken it (negative feedback)

**fluctuation** a general notion of change in a process over a given space or time period. The general mathematical theory of fluctuations is wavelets, although there exist fluctuations (such as those used in detrended fluctuation analysis) that are not wavelets. In this book,

the main fluctuations are differences, anomalies, and Haar fluctuations (equal to the difference of the anomaly).

**forced variability** the temperature variability that results from all the external climate forcing.

**fractal** a geometric set of points that is symmetrical under scale changes. It is scaling and generally characterized by noninteger fractal dimensions.

**fractal dimension** the exponent that characterizes how the size of the fractal set changes with scale; it can quantify sparseness or wiggliness.

**General Circulation Model** a numerical climate model based on the Navier-Stokes equation of fluid dynamics for either the atmosphere, the ocean, or both. Sometimes the term "Global Climate Model" is also used when other climate-relevant processes are included.

**generalized scale invariance** a general system for defining the notion of scale and size in anisotropic scaling systems; requires only a unit scale ("unit ball") and a rule (the "generator") to go from one scale to a neighboring scale.

**Global Climate Model** see General Circulation Model.

**greenhouse gas forcing** man-made emissions of greenhouse gases (e.g., carbon dioxide, methane, nitrous oxide, halogens, chlorofluorocarbons) that alter the atmospheric composition of Earth, which decreases the amount of thermal radiation lost to space (aka the "greenhouse effect"), and therefore leads to a positive climate forcing.

**hindcast** a way to validate predictions by ~~using~~ past data.

**Holocene** the warm interglacial period since the retreat of the ice sheets, about ~~12,000~~ years ago.

**Intergovernmental Panel on Climate Change** a scientific and intergovernmental body under the auspices of the United Nations, established in 1988; an international group of experts on climate change that regularly reviews and assesses scientific, technical, and socioeconomic information relevant to the understanding of climate change.

**internal variability** the random fluctuations of the weather and macroweather that are linked to internal modes of heat storage of the climate and not to external climate forcing.

**macroclimate** is the climate regime between about 100,000 and 500,000 years in which fluctuations decrease quickly with timescale; apparently dominated by the responses to astronomical forcings, although it is not clear whether these responses are scaling or simply a broad quasi-periodic regime.

**macroweather** the dynamical regime, starting at timescales of about ten days, in which fluctuations in atmospheric variables (e.g., wind, temperature, precipitation, humidity) decrease with timescale ($H < 0$). Because fluctuations tend to cancel, the macroweather regime has the property ~~that states~~ that averages over longer and longer times tend to converge. However, after some critical timescale—the macroweather–climate transition—averages again tend to increase ($H > 0$). In the Anthropocene, this transition time is currently about twenty years. During the preindustrial epoch, it was variable geographically and also variable from epoch to epoch. During most of the past 800,000 years, it has varied from several centuries to millenia.

**macroweather state** the state of the atmosphere after it has been averaged over the weather scales (typically taken as one month) and then subtracted from a climate state (typically

defined by an average over 30 years; the climate normal, with the seasonal cycle also removed). It is thus an anomaly with respect to the current climate state.

**megaclimate** the regime from about 500,000 years to at least 550 million years in which temperature fluctuations grow with scale ($H > 0$). According to benthic paleodata, it is at least roughly scaling. It is thought to be dominated by biogeological processes.

**multifractal** a mathematical field or measure ~~of density~~ that is symmetrical with respect to generalized scale changes (scaling). It is usually characterized by a scale-invariant exponent that is a function. This is equivalent to an infinite number of exponents (one for each level of activity, weak or strong). This infinite hierarchy can often be reduced to three as a result of the existence of basic, stable attractive behaviors: multifractal universality.

**multifractal butterfly effect** analogous to the usual butterfly effect in that it arises from tiny, small-scale perturbations. In this case, the perturbations are distributed spatially (as though from a flock of butterflies), and the consequence is the occasional extreme (black swan) event, power law probability tails.

**multimodel ensemble** an ensemble of General Circulation Model ~~results with a~~ mean ~~that~~ is often used to project climate change and ~~with a~~ dispersion ~~that~~ is used to quantify the uncertainty (sometimes called "structural uncertainty").

**multiproxy reconstruction** the use of large numbers (typically hundreds or thousands) of paleoindicators distributed geographically to reproduce the temperature evolution over centuries and millennia. Multiproxy reconstructions are of large-scale (~~usually~~ hemispheric or global) temperature series.

**natural variability** is the response of the atmosphere to natural forcing agents, such as solar and volcanic forcing, that are external to the climate system.

**negative emissions** technologies that remove carbon dioxide from the atmosphere; a form of geo-engineering. To be useful for mitigating global warming (to offset, in part, the usual positive emission from fossil fuel burning), they would have to be deployed on a massive scale. At the moment, these are highly speculative although almost all scenarios of economic development include them.

**paleotemperature** a proxy such as an oxygen isotope concentration that is related to the temperature in a known way; often used as a proxy for the actual temperature.

**pause** the period following the strong warming of the 1998 El Niño event, during which natural cooling temporarily offset anthropogenic warming; it was poorly forecast by GCMs. Sometimes called the "hiatus" or "slowdown" in the warming.

**phenomenological fallacy** the ~~illogical~~ inference of dynamical mechanisms from phenomenological form. The fallacy arises in two distinct but related ways. In the first, one attributes qualitatively different mechanisms purely on the basis of morphologies, structures, at different space and/or timescales. In the atmosphere and climate, this is often unwarranted, because a single wide-range scaling mechanism may suffice (e.g., clouds). In the second, even at given space and/or timescale, different mechanisms may be attributed to weak and strong events. This occurs when there is scaling in the probabilities (see "black swans").

**predictions** forecasts of the evolution of the future state of the atmosphere. Macroweather predictions are over periods of typically months to years to decades. They generally exclude the response to changing climate forcings because they depend on anthropogenic forcing scenarios.

**projections** the evolution of the forced climate change for decadal timescales and longer. ~~Because~~ they depend on climate forcing scenarios (i.e., human actions).

**reanalysis** a data–hybrid; the result of assimilating meteorological observations into a Numerical Weather Model. This can help correct errors in measurements and allows the assimilation of satellite radiances and other indirect sources of information.

**representative carbon pathways** future emission scenarios designed for Coupled Model Intercomparison Project Phase 5 to allow for intermodel comparison. There are four scenarios (RCP2.6, RCP4.5, RCP6.0, and RCP8.5), which were established based on different socioeconomic assumptions. The number in the name refers to the total radiative forcing (measured in watts per square meter) expected in 2100 under each scenario (in comparison with the preindustrial level).

**scalebound** a way of thinking about the hierarchy of different processes and phenomena in the natural world; equivalent to the "new worlds" view. In the scalebound paradigm, one expects to find qualitatively different processes and phenomena every factor of 10 or so in scale. One attempts to isolate phenomena by restricting the scale range. Biology furnishes typical scalebound examples. For example, adult human beings are roughly in the range between 50 cm and 2 m in height; their height and size is scalebound.

**scale invariant** the exponent in a scaling system (such as the fractal dimension $D$ or the fluctuation exponent $H$) that does not change with scale. Scale can be defined in a rather general, anisotropic way via generalized scale invariance. Mathematically, it only requires a notion of size (a measure), not notion of distance (a metric). In multifractal processes, the scale-invariant exponents are not numbers; they are functions and are equivalent to a different exponent for every level of activity (e.g., strong, weak).

**scaling** when a property such as volume, mass, or temperature fluctuation varies in a power law way with spatial scale (or has the analogous property with respect to timescale or with respect to space-time). Because power laws have no characteristic size, they describe dynamics and phenomena that act over wide ranges of scale. Their exponents do not change with scale; they are scale invariant.

**scaling regime** a range over which the variability of a system respects a scale invariance symmetry; it is scaling. Over a scaling regime, there is no fundamental space or timescale and the same basic dynamics hold, repeating scale after scale, for example, in a cascadelike manner.

**simple climate model** a numerical climate model with few degrees of liberty in which climate sensitivity is usually specified as a parameter that can be found by comparison with the historical record.

**solar forcing** small variations in the sun's total irradiance resulting from solar turbulence and internal solar cycles (especially sunspot cycles with a period of about eleven years). The forcing is measured as deviations from a long-term average "solar constant."

**spectrum** sometimes called the "power spectrum"; the variance (mean square) of a process per frequency (time) or per wavenumber (space). It characterizes variability as a function of inverse time (frequency) or inverse length (wavenumber).

**transient climate** the constant of proportionality between the temperature response and the total climate forcing over the historical period. It ignores the longer term memory and is expressed in terms of the temperature change after a doubling in carbon dioxide equivalent concentration. The difficulty in estimating it empirically comes from uncertain past reconstructions, especially aerosol emissions.

**transient climate response** the temperature change after a gradual doubling in carbon dioxide concentration over seventy years. It can only be determined from a climate model.

**turbulence** the regime of chaotic fluid flow that occurs when the forces stirring a gas or liquid are much larger than the internal friction forces (viscosity), which tend to dampen the motions. The atmosphere is typically turbulent at all scales larger than a small viscous scale of less than a millimeter.

**volcanic forcing** when volcanic eruptions eject large amounts of sulfate aerosols up to the stratosphere, where they reflect solar rays and therefore lead to negative forcing.

**warming in the pipeline** the committed warming stored in the Earth system, mostly in the oceans. This stored energy will warm the atmosphere in the future even if the forcing were to stop today.

**warming limits of 1.5°C (or 2°C)** thresholds often invoked that should not be surpassed to ensure the habitability of our planet. Since the COP 21 meeting in Paris in 2015, the 1.5°C limit is one that the international community aims to respect whereas, since the COP 16 meeting in Cancun in 2010, the 2°C limit is a threshold that is imperative to respect.

**weather** the dynamical regime in which fluctuations in atmospheric variables (e.g., wind, temperature, precipitation, humidity) increase with timescale ($H > 0$). From milliseconds to the macroweather transition scale around ten days. This is the typical lifetime of planetary structures.

**weather state** the average state of the atmosphere for timescales up to the scale of the weather–macroweather transition of about 10 days. Typical weather maps are weather states reperesenting averages over 6 hours or so. The longer the weather state is averaged, the lower is its spatial resolution, because in the weather regime, space and time are connected statistically by the typical large-scale wind (about 10 m/s).