Received: 3 June 2020 | Revised: 13 January 2021 | Accepted: 9 February 2021

DOI: 10.1002/qj.4005

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

# The fractional energy balance equation

**Shaun Lovejoy**[1] | **Roman Procyk**[1] | **Raphael Hébert**[1,2] | **Lenin Del Rio Amador**[1]

[1]Physics, McGill University, Montreal, Quebec, Canada

[2]Alfred-Wegener-Institut, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany

**Correspondence**
S. Lovejoy, Physics, McGill University, 3600 University Street, Montreal, QC H3A 2T8, Canada.
Email: lovejoy@physics.mcgill.ca

## Abstract

Classical Energy Balance Equations (EBEs) are differential equations of integer order ($h = 1$), here we generalize this to fractional orders: the Fractional EBE (FEBE, $0 < h \leq 1$). In the FEBE, when the Earth is perturbed by a forcing, the temperature relaxes to equilibrium via a slow power-law process: $h = 1$ is the exceptional (but standard) exponential case. Our FEBE derivation is phenomenological, it complements derivations based on the classical continuum mechanics heat equation (that imply $h = 1/2$ for the surface temperature) and of the more general Fractional Heat Equation which allows for $0 < h < 2$. Unlike some of the earlier "scale free" models based purely on scaling, the FEBE has an extra blackbody radiation term that allows for energy balance. It therefore has two scaling regimes (not one), it has the advantage of being stable to infinitesimal step-function perturbations and it has a finite Equilibrium Climate Sensitivity. We solve the FEBE using Green's functions, whose high- and low-frequency limits are power laws with a relaxation scale transition (several years). When stochastically forced, the high-frequency parts of the internal variability are fractional Gaussian noises that can be used for monthly and seasonal forecasts; when deterministically forced, the low-frequency response describes the consequences of anthropogenic forcing, it has been used for climate projections. The FEBE introduces complex climate sensitivities that are convenient for handling periodic (especially annual) forcing. The FEBE obeys Newton's law of cooling, but the heat flux crossing a surface nonetheless depends on the fractional time derivative of the temperature. The FEBE's ratio of transient to equilibrium climate sensitivity is compatible with GCM estimates. A simple ramp forcing model of the industrial-epoch warming combining deterministic (external) with stochastic (internal) forcing is statistically validated against centennial-scale temperature series.

**KEYWORDS**

climate response functions, energy balance, externally forced variability, fractional equations, internal variability, scaling

  

## 1 | INTRODUCTION

The Earth is a heterogeneous system with a dynamically evolving atmosphere, ocean and land surface, it has a huge number of degrees of freedom. While general-circulation models (GCMs) are routinely used to model its evolution, these are highly complex and each GCM has its own climate, and these are different from the real-world climate. Simplified models have therefore been developed that exploit what at first sight appears to be a fairly simple yet fundamental constraint: the conservation of energy.

The corresponding Energy Balance Models (EBMs) are based on the (near) energy balance of the Earth with the Sun and outer space. For example, anthropogenic forcings are currently around 2.4 $W \cdot m^{-2}$ above the long-term mean solar forcing of $\approx 238\, W \cdot m^{-2}$ (both figures are for global averages), so that the perturbations away from balance are of the order of 1%. This makes it plausible that the responses to external forcings are roughly linear. This was confirmed by examining the evolution of the spatial patterns of 32 Coupled Model Intercomparison Project phase 5 (CMIP5) GCM responses when forced by the various Representative Carbon Pathway (RCP) scenarios (Hébert and Lovejoy, 2018). Even when these included forcing as high as 8.5 $W \cdot m^{-2}$, 3.5% of the reference mean, the response (the regional distribution) of each model – although different from the other models – was (nearly) a linear projection of its own past. The extent of the linear regime is unclear, and it is certain that due to albedo–temperature feedbacks and other nonlinearities, it will eventually break down. Consequently, strongly nonlinear (chaotic) models have been proposed, see Dijkstra (2013) for a review. These are based on Hasselmann (1976)'s rather general and now classical mathematical decomposition of the dynamics into slow/fast components where the fast (weather) scales stochastically drive the slow (climate) component (see Section 4 below).

Two different linear EBM modelling approaches have evolved, the earliest going back to Budyko (1969) and Sellers (1969) who applied the continuum mechanics heat equation. This yielded one-dimensional (1D) (longitudinally averaged) models of the equilibrium surface temperature's latitudinal distribution, later extended to time (Dwyers and Petersen, 1973). An attractive feature of the Budyko–Sellers models is that the system can become nonlinear if the albedo (and hence forcing) is coupled with the temperature (e.g. via the position of the snow line). This has enabled it to be used for modelling past and (possible) future climates.

A second EBM strand was inspired by "box models" that related $CO_2$ emissions to atmospheric concentrations. In energy balance box models, one or more boxes are used to model the regional or – more usually – the globally averaged temperature. The simplest such "zero-dimensional" box models were first proposed by Hasselmann *et al.* (1993) and involve a single uniform slab of material – the "box" – representing the Earth, and that interacted radiatively with outer space according to Newton's law of cooling (NLC). Although the Budyko–Sellers and box approaches are distinct, when the time-dependent versions are reduced to zero-dimensions and linearized, they are mathematically equivalent to a (single) box model (for a review, see North *et al.* (1981), the update North *et al.* (2011), and the valuable monograph North and Kim (2017)). In addition, a stochastically forced version of the same relaxation equation (applicable not only to the temperature) was proposed by Hasselmann (1976) based on the weather–climate scale separation rather than on energy considerations.

A key difficulty for all EBMs is that some of the energy received from the Sun today is stored and emitted to outer space only at a later time. If we consider models of the Earth averaged over (at least) macroweather time-scales – that is, longer than the typical lifetime of atmospheric (weather) structures (Lovejoy, 2013) – then some of the long-wave/short-wave radiative flux imbalances are stored in the subsurface (mostly in the oceans) and some are transported horizontally. Conversely, energy that was stored long ago – for example in the deep ocean – may appear and contribute to today's long-wave emissions. The key to applying the energy balance principle is therefore to have a realistic model of the storage processes.

The box and Budyko–Sellers models handle the storage problem in quite different ways. The original Budyko–Sellers models had no storage at all, the flux imbalance was simply redirected meridionally away from the Equator. In the box models, the temperature ($T$) of each box is spatially uniform and the energy stored ($S$) is the product of temperature with the box heat capacity ($C$) so that all the storage occurs as thermodynamic internal energy. Since the model corresponds to the state of the atmosphere averaged over a month or longer and over spatial scales containing many (energetic) meteorological and oceanographic degrees of freedom, this seems unnecessarily restrictive. The model also unrealistically implies that energy fluxes into and out of the box instantaneously change the temperature of the entire box.

A recent series of papers showed how the realism of these Budyko–Sellers models could be greatly improved by introducing a vertical coordinate (Lovejoy 2021a, b, hereafter L1, L2). Even though the aim is to improve the modelling of the horizontal temperature distribution, the inclusion of a vertical coordinate has two advantages. First, it allows us to apply the correct conductive-radiative surface boundary condition to determine the fraction

of the radiative imbalance that is conducted into the subsurface and the fraction that is re-emitted at long wavelengths. Second, since the subsurface stores the heat, the third dimension simultaneously provides a precise storage mechanism. When this conductive-radiative boundary condition was applied to the classical 3D heat equation (the Budyko–Sellers model with an extra vertical coordinate), a surprising consequence was that the *surface* temperature obeyed the long-memory, Half-ordered Energy Balance Equation (HEBE: L1, L2). In the HEBE, the source of the surface temperature long memory is completely classical; it is simply a consequence of the slow diffusion of heat into and out of the subsurface.

Even if we use the standard integer-ordered heat equation, we are being forced to deal with fractionally ordered surface temperature equations. But why restrict ourselves to the classical heat equation? Indeed, Budyko–Sellers-type models are for macroweather Earth states, that is, they are averaged over a wide range of (weather) degrees of freedom, they are "effective material" type models, there is no compelling reason that they should be integer ordered. Indeed, it suffices to make a modest extension to the classical heat equation to obtain the Fractional Heat Equation (FHE), which is a fractional diffusion equation that has been studied in the statistical physics literature.

The conductive-radiative boundary condition applied to the FHE implies that the surface temperature satisfies the Fractional EBE or FEBE, generalizing the HEBE from $h = 1/2$ to $0 < h < 1$. Like the HEBE, the FEBE is fundamentally a 2D macroweather temperature anomaly model, whose zero-dimensional version is a fractional relaxation equation, itself a fractional generalization of the classical zero-dimensional box models that are the $h = 1$ special cases. When the FEBE is driven by a Gaussian white noise, the result is fractional Relaxation noise (fRn) that generalizes the classical Ornstein–Uhlenbeck process, and its high-frequency limit is a fractional Gaussian noise process (fGn) that generalizes Gaussian white noise. The resulting FEBE can be used advantageously for both macroweather forecasts (the high-frequency FEBE approximation: Lovejoy *et al.*, 2015; del Rio Amador and Lovejoy, 2019) and multidecadal projections (the low-frequency approximation: Hebert, 2017; Lovejoy *et al.*, 2017; Hébert *et al.*, 2021, and Procyk *et al.* (2020) for the full FEBE).

This article focuses on models that combine the physical principles of energy balance and scaling. These are distinct from both the classical models that use only energy balance as well as from various proposals to model the temperature using only the scaling principle. For example, Van Hateren (2013) proposed a "fractal" model as a scaling hierarchy of (energy balancing) box models whereas Hebert (2017) and Hébert *et al.* (2021) proposed a model whose approach to equilibrium was scaling, essentially a low-frequency approximation to the FEBE (see Section 3.3 below). This was achieved by truncating the scaling at high rather than low frequencies.

While both of these scaling models respect energy balance, and are stable to small perturbations, there are several scaling models that are not. For example, Rypdal (2012) proposed a model in which the temperature at decadal and multicentennial scales responds in a "scale free" (pure power-law) manner. This model is effectively the FEBE but with only the fractional storage term. Without the black-body emission term, energy balance is impossible. While the author acknowledges that this model is unstable to infinitesimal step forcings (it has an infinite Equilibrium Climate Sensitivity, ECS), he argues that unspecified "slow feedbacks" will eventually truncate the scaling behaviour at very long time-scales. Similarly, Rypdal and Rypdal (2014) argue that divergences will be avoided because eventually the linear scaling model will be "saturated by nonlinear effects" and break down (see Hébert and Lovejoy (2015) for a critique). It is revealing that when Rypdal *et al.* (2015) made a fractional generalization of North *et al.* (2011)'s heat diffusion model, that it was obtained precisely by removing the critical energy balance term.

The use of infinite ECS models has been also justified by the fact that over a limited range of scales, they can be approximated by multi-box models that have finite ECS (Fredriksen and Rypdal, 2017), so that they might still be useful. Alternatively, since even infinite ECS models have finite responses after finite times, their Transient Climate Responses (TCRs) will be finite, prompting Rypdal and Rypdal (2014) to argue that the TCR may be a more significant quantity than the ECS. In any event, an unsatisfactory consequence of the absence of energy balance is that when pure scaling models are used for climate projections, in order to avoid infinite temperatures, the forcings must return to zero (Rypdal, 2016; Myrvoll-Nilsen *et al.*, 2020).

Since the FEBE has different high- and low-frequency scaling regimes and the pure scaling model has only one, the two models are theoretically and empirically distinct. For example, model differences are important at poorly understood multicentennial/multimillennial scales so that empirical clarification of the climate variability over these scales is an important goal of the Climate Variability Across Scales (CVAS) Past Global Changes (PAGES) working group (since 2015, Lovejoy (2017b), Hébert *et al.* 2020; for reviews, see Franzke *et al.* (2020) and Lovejoy (2019a)).

Rather than deriving the FEBE from a three-dimensional continuum heat model (following L1,

L2), in this article it is instead derived phenomenologically by arguing that the basic energy storage mechanisms respect a scaling principle and we treat the consequences of the FEBE for the Earth's temperature. In Section 2 we derive the zero-dimensional FEBE, and in Section 3 some of its mathematical properties, consequences and solutions. This includes the stochastic internal variability that was examined from a more mathematical point of view in Lovejoy (2019b), hereafter L3. In Section 4, we discuss model parameter estimation/validation (high frequency, low frequency, annual periodicity). We also propose a simple ramp and plateau model of the industrial epoch that includes both stochastic internal forcing and deterministic external forcing.

## 2 | ENERGY STORAGE AND THE CLASSICAL ENERGY BALANCE EQUATION (EBE)

### 2.1 | Multibox models and climate response functions

A single box with its unique relaxation time is an inflexible model. Hasselmann *et al.* (1993; 1997) already noted that it was desirable to go beyond this to use the more general linear response function framework. In this context, the response functions are called "climate response functions" (CRFs), and following Hasselmann *et al.* (1993) we have a choice between the equivalent impulse and step CRFs, the former being the derivative of the latter (see below).

Unfortunately, without more assumptions or information, the linear framework with (nearly) arbitrary CRFs is unmanageably general. In order to make progress, Hasselmann *et al.* (1997) proposed a response function consisting of a CRF equal to the sum of $N$ exponentials corresponding to $N$ boxes that mutually exchange heat, each with its own exponential relaxation time and relative weight. But a "multibox" model with $N$ boxes requires $2N$ parameters – it is still too general – so that out of the practical necessity of fitting GCM outputs, Hasselmann *et al.* (1997) ultimately chose $N = 3$. Following the more usual procedure of deriving the impulse responses from integer-ordered linear differential equations (where CRFs are "Green's functions"), Li and Jarvis (2009) recalled that the Green's functions of the $n$th-order differential equations (with constant coefficients and with $n$ an integer) can quite generally be reduced to sums of exponentials. However, in the application part of their paper, they nevertheless also used the value $N = 3$. Similarly, Fredriksen and Rypdal (2017) argued that two boxes were not enough and used a three-box model to approximate a wider scale range (power-law) CRF over scales ranging from roughly 1 to 1,000 years.

Even the $N = 3$ multibox model is difficult to manage and most authors have settled for $N = 1$ or 2, including the IPCC AR5 (2013, Section 8.SM.11.2) who recommend the four-parameter $N = 2$ model. Yet problems remain; to be taken seriously, the boxes and their time constants must have physical interpretations. For example, the role of each box is to store heat, but what do the boxes physically represent? Using the atmosphere implies relaxation times of weeks, which is too short for most applications. For the more popular two-box model, there is a rough consensus that each represents part of the ocean: an upper "mixed layer" and a lower "deep" layer. By appropriately choosing the layer thicknesses, one can get relaxation times $\tau$ spanning years to millennia. For example, single-box models have $\tau$ ranging from 4 to 40 years (e.g. Schwartz, 2007; 2012; Held *et al.*, 2010; Zeng and Geil, 2016). Two-box models typically have one box with $\tau$ in the range 2–10 years and another with a $\tau$ in the range 20–800 years (Rypdal, 2012; Geoffroy *et al.*, 2013) with the IPCC AR5 favouring $\tau = 8.4$ and $\tau = 409.5$ years.

### 2.2 | Scaling storage

Rather than justifying a choice of CRF by invoking hypothetical homogeneous boxes, one can instead derive the CRF directly from numerical model outputs (e.g. Hansen *et al.*, 2011) or from physical considerations. The key is to exploit the wide-range spatial scale invariance of geo-processes such as the ocean and atmospheric dynamics associated with energy storage (see e.g. Lovejoy and Schertzer (2013) for a review). The idea is to use the fact that from small to large spatial scales, there is a whole scaling hierarchy of storage processes (e.g. atmospheric or oceanic eddies). Since the heat transfer times of each structure in the hierarchy depend on its spatial scale, it is reasonable to assume that the collective overall heat storage is also scaling.

As mentioned in the introduction, up until now the scaling principle has not been applied to the storage, but instead directly to the overall temperature response (Rypdal, 2012; Van Hateren, 2013; Rypdal and Rypdal, 2014; Hebert, 2017; Myrvoll-Nilsen *et al.*, 2020; Hébert *et al.*, 2021). Scaling impulse (Dirac δ-function) response CRFs are power laws:

$$G_\delta(t) \propto t^{H_F - 1}, \tag{1}$$

where $H_F$ is a scaling exponent. However, such pure power-law CRFs suffer from the "runaway Green's function effect" (Hébert and Lovejoy, 2015): the fact that they imply nonphysical divergences either at low or at high frequencies (depending on whether $H_F > 0$ or $< 0$). This can readily be seen by considering the physically important integral of $G_\delta$, the step response CRF: $G_\Theta(t) \propto t^{H_F}$. For example, Rypdal (2012) proposed a CRF with $H_F > 0$ which implies that $G_\Theta(t)$ diverges as $t \to \infty$; such CRFs would yield infinite temperature responses to a doubling in $CO_2$: an infinite ECS. To avoid the divergence, forcings must be carefully restricted so as to eventually return to zero. If low-frequency truncations on a power law with $H_F > 0$ are invoked (Rypdal, 2012; Rypdal and Rypdal, 2014), then the behaviour depends critically on the details of the truncation.

To avoid divergences, Van Hateren (2013) instead proposed a "fractal climate response" model that had $N = 6$ boxes but used the scaling principle to link the amplitudes and time constants by power laws yielding a four-parameter model. As mentioned in Hébert *et al.* (2021), his model featured unphysical logarithmic oscillations and an unnecessary low-frequency cut-off.

Recently, Myrvoll-Nilsen *et al.* (2020) have applied the original (Rypdal, 2012) infinite ECS model to twenty-first century warming scenarios claiming that their model with "not well-defined" ECS nevertheless "provides an accurate description of both forced and unforced surface temperature fluctuations". Be that as it may, an additional attempt to justify their model as "a cost of the reduction of model complexity" cannot be sustained. For example, Hebert (2017) and Hébert *et al.* (2021) proposed a scaling CRF but instead took $H_F < 0$ combined with an explicit high-frequency cut-off to avoid divergences:

$$G_\Theta(t) = 1 - \left(1 + \frac{t}{\tau}\right)^{H_F}; \quad H_F < 0, \quad t \geq 0. \tag{2}$$

At long times $t \gg \tau$, this step response CRF yields a power-law approach to a constant (corresponding to energy balance), and $G_\Theta(0) = 0$ so that there is no divergence at $t = 0$ either. The truncated scaling model (Equation (2)) or the FEBE (below) both show that it is easy to avoid models with "not well-defined" ECS while continuing to maintain model simplicity.

The fundamental cut-off scale $\tau$ was estimated as 2 years from the empirical coupling time of fluctuations in the ocean and atmosphere (Hebert, 2017; Lovejoy *et al.*, 2017). Using historical forcings and temperature responses since 1880, they empirically estimated the exponent $H_F$ ($\approx -0.5$) as well as the climate sensitivity parameter $s = 2.4$ K·per $CO_2$ doubling (with 90% confidence interval [1.8–3.7K]: Hébert *et al.*, 2021). Using the IPCC AR5 scenarios (the Representative Carbon Pathways, RCPs), the parameters were then used to make temperature projections to 2100. Since the source of the projection uncertainties were from the past data – not the large GCM "structural uncertainties" – they showed that the projections were compatible with – but much less uncertain – than the CMIP5 alternatives.

In spite of its success in making climate projections, due to their truncations, the scaling CRF models are only approximations; they apply the scaling principle directly to the CRF whereas in fact it should only be applied to the storage term in the energy balance equation (EBE). It turns out that this model change can be made with a surprisingly small tweak to the usual EBE; it suffices to change the order of differential equation from integer to fractional order, the Fractional Energy Balance Equation (FEBE) discussed below. We will show that this seemingly trivial generalization can account for responses to both externally forced and internal variability, effectively explaining them both as different aspects of energy conservation.

## 2.3 | The Earth's energy balance

The Earth's energy balance can be expressed as:

$$S(t) + E_{\text{tot},\uparrow}(t) = E_{\text{tot},\downarrow}(t), \tag{3}$$

where the upward arrow indicates energy emitted to outer space and the downward arrow the energy received from the Sun (and other sources external to the climate system); the difference is the energy stored, $S(t)$; the subscript "tot" indicates the total over a long time, starting at a convenient reference baseline.

Before continuing, it is worth making a few comments:

1. This equation directly expresses conservation of energy; it is not yet the more usual rate equation obtained by taking derivatives (see below).
2. Equation (3) is "zero-dimensional", that is, all the spatial degrees of freedom are averaged out, $S$, $E_{\text{tot},\uparrow}$, $E_{\text{tot},\downarrow}$ are averages over the whole Earth: energies per unit surface area (in standard units, $\text{J}\cdot\text{m}^{-2}$).

Now consider anomalies, that is, the deviations $E_\uparrow$, $E_\downarrow$ from the long-term values $E_{0,\uparrow}$, $E_{0,\downarrow}$:

$$\begin{aligned} E_{\text{tot},\uparrow}(t) &= E_{0,\uparrow}(t) + E_\uparrow(t) \\ E_{\text{tot},\downarrow}(t) &= E_{0,\downarrow}(t) + E_\downarrow(t). \end{aligned} \tag{4}$$

The long-term radiances $E_{0,\uparrow}$, $E_{0,\downarrow}$ could be taken as being proportional to long-term average rates, outgoing and incoming radiances $R_{0,\uparrow}$, $R_{0,\downarrow}$ so that $E_{0,\uparrow}(t) = tR_{0,\uparrow}$ $E_{0,\downarrow}(t) = tR_{0,\downarrow}$.

Since the Earth is nearly in radiative balance, $R_{0,\uparrow} = R_{0,\downarrow}$ hence $E_{0,\uparrow} = E_{0,\downarrow}$, so that the storage $S(t)$ is the difference between the long- and short-wave radiative anomalies:

$$S(t) + E_{\uparrow}(t) = E_{\downarrow}(t) \tag{5}$$

(subtract the two equations in (4) and use Equation (3)).

If the storage processes are scaling in space (i.e. hierarchical), and the characteristic storage time of each structure depends on its spatial extent in a power-law manner, then it is reasonable to assume that the storage at a time $t$ depends on the past in a power-law manner:

$$S_h(t) = \frac{C_h}{\Gamma(1-h)} \int_{t_0}^{t} T(u) \left( \frac{t-u}{\tau} \right)^{-h} \frac{du}{\tau}; \quad 0 \le h \le 1, \tag{6}$$

where $C_h$ is the (generalized) heat capacity per unit area of the Earth, which in standard units is $[C_h] = \frac{\mathrm{J}}{\mathrm{km}^2}$, $h$ is a fundamental scaling exponent and $\tau$ is the (power law) relaxation time (see below). The power-law kernel in Equation (6) relates the storage to the temperature via a Riemann–Liouville fractional integral of order $1-h$ with $0 \le h \le 1$; $\Gamma$ is the usual Gamma function. $t_0$ is the reference time from which the total storage is measured (here, taken to be in the distant past): $S_H(t_0) = 0$, later we will take $t_0 = -\infty$. We have added the subscript $h$ to emphasize that except in the classical $h = 1$ case, $C_h$ is not a usual heat capacity.

The storage defined in Equation (6) is a generalization of the standard "box model" storage to which it reduces when $h = 1$. To see this, we integrate by parts and take $T(t_0) = 0$, this is equivalent to fixing the reference of our anomalies:

$$S_h(t) = \frac{C_h \tau^{h-1}}{\Gamma(2-h)} \int_{t_0}^{t} T'(u)(t-u)^{1-h} du; \quad 0 \le h \le 1. \tag{7}$$

Taking $h = 1$, we now obtain the usual $h = 1$ box storage result:

$$S_1(t) = C_1 T(t); \quad h = 1. \tag{8}$$

A key difference between the $h < 1$ and $h = 1$ cases is that only in the latter (box model) does the storage depend instantaneously on the temperature.

# 3 | THE FRACTIONAL ENERGY BALANCE EQUATION

## 3.1 | Fractional storage

The usual energy balance equation is an equation for the rates, it is obtained by differentiation of Equation (5):

$$\frac{dS_h(t)}{dt} + s^{-1} T(t) = F(t); \; F(t) = \frac{dE_{\downarrow}}{dt}; \; s^{-1} T(t) = \frac{dE_{\uparrow}}{dt}, \tag{9}$$

where $F$ is the rate of anomalous energy input – the usual forcing (power per area), $s$ is the climate sensitivity and $T$ the temperature anomaly with respect to the long-term equilibrium temperature.

The sensitivity is the temperature rise per forcing; in standard units, $[s] = \frac{\mathrm{Kkm}^2}{\mathrm{W}}$. When combined with the generalized heat capacity, it defines a time-scale:

$$\tau = sC_h. \tag{10}$$

Since this dimensional combination is unique, $\tau$ is the characteristic time for the system to relax to equilibrium when it is perturbed by a step-function forcing. However, when $h \neq 1$, $C_h$ does not correspond to the usual thermodynamic specific heat where energy in a material is stored in internal thermodynamic (molecular) degrees of freedom. Our model is a macroweather model, that is, for the system averaged (roughly) monthly or longer and in space over tens or hundreds of kilometres. The energy can effectively be stored in internal weather degrees of freedom, it is really an average coefficient that expresses the collective heat transfer and storage characteristics of a hierarchy of storage mechanisms. Rather than treating $C_h$ as fundamental and then using Equation (10) to determine the relaxation time, it is better to treat $\tau$ as fundamental and $C_h$ as a derived quantity, $C_h = \tau/s$, so that the storage can be written:

$$S_h(t) = \frac{\tau^h}{s\Gamma(1-h)} \int_{t_0}^{t} T(u)(t-u)^{-h} du; \quad 0 \le h \le 1. \tag{11}$$

In this model, beyond the dimensionless exponent $h$, $\tau$ is the fundamental physical quantity characterizing the storage processes.

We now recall the Riemann–Liouville derivative of order $h$ that is defined as the ordinary derivative of the $1-h$ order fractional integral:

$${}_{t_0}D_t^h T = \frac{d}{dt}({}_{t_0}D_t^{h-1} T)$$

$$= \frac{1}{\Gamma(1-h)} \frac{d}{dt} \left[ \int_{t_0}^{t} (t-u)^{-h} T(u) du \right] \quad 0 \le h \le 1$$

$$= \frac{1}{\Gamma(1-h)} \int_{t_0}^{t} (t-u)^{-h} T'(u) du, \tag{12}$$

where the bottom equality follows by integrating by parts and using $T(t_0) = 0$, the reference temperature. Therefore,

$$\frac{dS_h}{dt} = \frac{\tau^h}{s} {}_{t_0}D_t^h T; \quad 0 \le h \le 1. \tag{13}$$

Putting this into Equation (9) and taking $t_0 \to -\infty$ we obtain:

$$\tau^h {}_{-\infty}D_t^h T + T = sF. \tag{14}$$

This is the fractional energy balance equation, FEBE. The physical meaning of the fractional derivative is that when the forcing changes, the change in temperature is not instantaneous but rather depends on the entire past history of temperature changes. At low frequencies, the linear term dominates, it corresponds to the anomalous blackbody radiation caused by the anomalous forcing. If $F$ is a step function, then the new equilibrium temperature is $sF$. Without this term – the pure scaling model discussed in the introduction (e.g. Rypdal, 2012) – the temperature would diverge.

Starting with the continuum mechanics heat equation, L1 derives this with the value $h = 1/2$; the full FEBE (with $0 < h < 1$) can be analogously derived from the fractional heat equation, itself a generalization of the continuum mechanics heat equation. The use of half-order derivatives in heat problems has a long history going back to at least Meyer (1960), Oldham and Spanier (1972; 1974) and Oldham (1973), developed also by Babenko (1986), Podlubny (1999) and Sierociuk *et al.* (2013; 2015). Interestingly, Oldham (1973) derives an equation mathematically identical to the $h = 1/2$ special case of the FEBE as a short time approximation to electrolyte diffusion in a spherical geometry, and Oldham and Spanier (1974) anticipates our present application by noting that half-order derivatives can be applied to "not one but an entire class of boundary value problems … ".

The $t_0 = -\infty$ limit is convenient when the forcing is a stationary stochastic process (such as a white noise representing the internal variability) since the temperature response will then also be statistically stationary. Alternatively, it is also convenient if the forcing is periodic with mean zero. In the case of a deterministic forcing that starts at a finite time (e.g. $t = 0$), we need only assume that $T(t) = 0$, $F(t) = 0$ for $t \le 0$. Due to the infinite range of the fractional derivative, the FEBE Equation (16) is a "Weyl" fractional relaxation equation.

We derived the FEBE for temperatures and forcing that were taken as anomalies with respect to a loosely defined reference state. In fact, since the equation is linear, one can simply add or subtract reference forcings and temperatures for convenience. The same is true for deterministic and stochastic forcings whose solutions can be superposed when needed (Section 4.5). If the temperature and forcing are nonlinearly coupled (e.g. via the albedo), then this freedom can be used to derive a nonlinear FEBE that could be used to model temperature/albedo feedbacks as well as the effects of orbital forcings, glacial–interglacial transitions and other low-frequency behaviours. In the future, specific models and their stability may be investigated in this way.

## 3.2 | Newton's law of cooling and thermal impedance

To understand the FEBE, we integrate both sides of Equation (14) by order $h$ and rewrite it in the form:

$$T(t) = {}_{-\infty}D_t^{-h}(\tau^{-h}(sF - T))$$

$$= \frac{1}{\Gamma(h)} \int_{-\infty}^{t} (T_{eq}(u) - T(u)) \left( \frac{t-u}{\tau} \right)^{h-1} \frac{du}{\tau}; \quad T_{eq}(t) = sF(t)$$

$$0 < h < 1; \tag{15}$$

The temperature is thus the fractional integral of order $h$ of the difference between the temperature and the equilibrium temperature $T_{eq}$. If the forcing is constant at a value $F$, then at long enough times, the temperature will "relax" to the equilibrium value $T_{eq} = sF$. When $h \neq 1$, this occurs in a power-law way (due to the power-law weight $\left(\frac{t-u}{\tau}\right)^{h-1}$ term). In the special $h = 1$ case – the conventional EBE – the relaxation is instead exponential.

If we interpret the forcing $sF$ in terms of an effective external equilibrium temperature $T_{eq}$, then the FEBE satisfies Newton's law of cooling (NLC) that states that a body's rate of heat loss is directly proportional to the difference between its temperature and its environment. In these horizontally homogeneous models, it is the heat flux (energy rate / area = $Q_s$) across the surface into the body (Figure 1) that is important so that the NLC can be written:

$$Q_s = \frac{1}{Z}(T_{eq} - T), \tag{16}$$

where $Z$ is a transfer coefficient sometimes called the "thermal impedance" (units: $m^2 \cdot K \cdot W^{-1}$), its reciprocal $Y$ is the surface "thermal admittance". Since $Q_s$ equals the rate

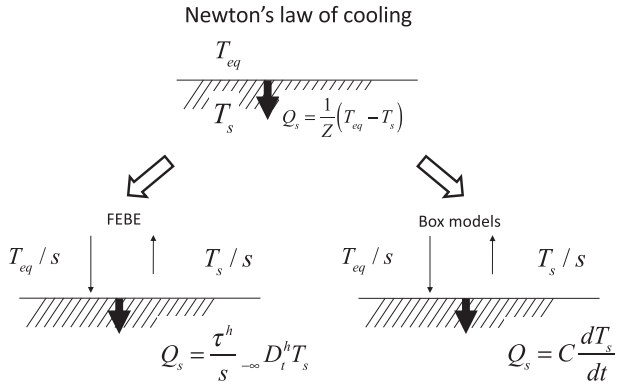


**FIGURE 1** A schematic showing Newton's law of cooling (NLC) that relates the temperature difference across a surface to the heat flux crossing the surface, $Q_s$ (into the surface). $T_{eq}$ is the fixed outside temperature, heat will flow across the surface as long as the surface temperature $T_s \neq T_{eq}$. $Z$ is the thermal impedance, equal here to the climate sensitivity $s$. To apply the NLC, we need to relate the heat flux to the surface temperature. The lower left shows the consequence of applying heat equation with conductive-radiative boundary conditions, the lower right shows the phenomenological assumption made by box models. The arrows represent heat fluxes, hence the factor $s$ in the denominators. The system is assumed to be horizontally homogeneous

of stored heat loss, we can rewrite Equation (9) as:

$$Q_s = \frac{dS_h}{dt} = \frac{\tau^h}{s}\,{}_{t_0}D_t^h T = \frac{1}{s}(T_{eq} - T). \tag{17}$$

By comparing this with Equation (16), we conclude (a) that the FEBE satisfies the NLC, and (b) that the climate sensitivity $s$ is a thermal impedance.

Before proceeding, a word about the energy balance, radiative equilibrium, thermal equilibrium and thermodynamic equilibrium. As long as we consider a "zero-dimensional" model representing a completely homogeneous Earth, energy balance is a state of radiative equilibrium, itself a state of thermal equilibrium. Thermal equilibrium is a necessary condition for thermodynamic equilibrium, but is generally not identical since thermodynamically irreversible processes generally occur. When the (fractional) time derivative is zero, the Earth is thus in a steady, energy balance state, equivalently in a state of radiative or thermal equilibrium.

## 3.3 | Solving the FEBE using Green's functions

Mathematically, when $0 < h < 1$ the FEBE is a "fractional relaxation equation"; the extension to $1 < h \leq 2$ is a "fractional oscillation equation". In the initial value problem, $F(t) = 0$ for $t \leq t_0$, the FEBE impulse Green's function – the impulse CRF – is given in terms of Mittag–Leffler functions:

$$G_\delta(t) = G_{0,h}(t) = \tau^{-1}\left(\frac{t}{\tau}\right)^{h-1} E_{h,h}\left(-\left(\frac{t}{\tau}\right)^h\right);$$
$$t \geq 0; \quad 0 \leq h \leq 2$$
$$G_{0,h}(t) = 0; \quad t < 0; \quad 0 \leq h \leq 2, \tag{18}$$

where:

$$E_{\alpha,\beta}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + \beta)} \tag{19}$$

is the α, β order Mittag–Leffler function (these and most of the following results are in the notation of Podlubny (1999)). The condition $G = 0$ for $t < 0$ is needed to respect causality. In the following this condition will be understood for all the $G$ functions. As can be seen directly from the series expansion (Equation (19)), the Mittag–Leffler functions are often called "generalized exponentials". In particular, the classical $h = 1$ box model is recovered with help of the (exceptional) ordinary exponential:

$$E_{1,1}(z) = e^z.$$

As usual, the Green's function gives us the particular solution to the inhomogeneous equation for a Dirac δ forcing. The full solution, including the solution to the homogeneous equation, is:

$$T(t) = s\int_{t_0}^{t} G_{0,h}(t-u)F(u)du + T_0 E_{h,1}\left(-\left(\frac{t-t_0}{\tau}\right)^h\right);$$
$$T_0 = T(t_0); \quad t \geq t_0 \tag{20}$$

(Cheng and Chu, 2011). Since $E_{h,1}(0) = 1$ we can confirm that this satisfies the initial condition $T(t_0) = T_0$.

Just as integer-ordered linear differential equations with constant coefficients can generally be solved with exponentials, the analogous fractionally-ordered differential equations can generally be solved with generalized exponentials. However, the latter are based on power laws so that the usual exponentials correspond to the special case where all the orders of derivatives are restricted to integer values; power-law Green's functions are thus more general.

A convenient feature of Mittag–Leffler functions is that they can be easily integrated by any positive order $\zeta$:

$$G_{\zeta,h}(t) = {}_0D_t^{-\zeta}(G_{0,h}(t)) = \tau^{\zeta-1}\left(\frac{t}{\tau}\right)^{h-1+\zeta}$$
$$E_{h,h+\zeta}\left(-\left(\frac{t}{\tau}\right)^h\right); \quad \zeta \geq 0; \quad 0 \leq h \leq 2 \tag{21}$$

(Podlubny, 1999). Since $G_{0,h}$ is the response to a Dirac function forcing, with $\zeta = 1$ we obtain the response to the first-order integral of the Dirac: the Heaviside (step function) forcing. Similarly, using $\zeta = 2$ we obtain the response to its integral, the "ramp" (=0 for $t<0$, $=t$ for $t\geq 0$, Section 4.5), using $\zeta = 3$ we have the response to a parabolic forcing, etc.:

$$ {}_{t_0}D_t^{-n}\delta(t) = \frac{t^{n-1}}{(n-1)!}\Theta(t);\ n \geq 1; \quad t_0 \leq 0, \tag{22}$$

where $n$ is an integer. Figure 2 shows the impulse, step and ramp responses for various $h$'s. The log–log plot is particularly revealing since it shows directly that for $h<1$, there are two power-law limits and that for the $h = 1$ case, we obtain the usual exponential CRF. Note for the step response (middle graphs, and Figure 3), the approaches to the asymptotic value 1 (left) corresponding to energy balance can be extremely slow.

The physically important step-function response – the step CRF – is thus obtained by taking $\zeta = 1$:

$$ G_\Theta(t) = G_{1,h}(t) = \int_0^t G_{0,h}(t')dt' = \left(\frac{t}{\tau}\right)^h E_{h,h+1}\left(-\left(\frac{t}{\tau}\right)^h\right); $$
$$ t \geq 0; \quad 0 \leq h \leq 2. \tag{23}$$

The $h = 1$ (box model) yields:

$$ G_{0,1}(t) = \tau^{-1}\mathrm{e}^{-t/\tau}; \quad G_{1,1}(t) = 1 - \mathrm{e}^{-t/\tau}, \tag{24}$$

where we have used $E_{1,2}(z) = \frac{\mathrm{e}^z - 1}{z}$.

By integrating Equation (20) by parts, and using $G_1(0) = 0$, and the condition $F(t_0) = 0$, we can express the temperature response in terms of the step CRF rather than the impulse CRF:

$$ T(t) = s\int_{t_0}^t G_{1,h}(t-u)F'(u)du; \quad F'(t) = \frac{dF}{dt} \tag{25}$$

(we have taken $T(t_0) = 0$ for simplicity). Mathematically, the equivalence between Equations 20 and 25 arises because $\delta(t)$ is the derivative of $\Theta(t)$, and $G_{0,h}(t)$ is the derivative of $G_{1,h}(t)$. Expressing the solution in terms of the step CRF rather than the impulse CRF is advantageous. For example, $G_{1,h}(t)$ is dimensionless so that the sensitivity parameter $s$ has the usual dimensions of degrees per power per area. In addition, following Hansen *et al.* (2011) and emphasized by Marshall *et al.* (2014; 2017), the step CRF directly expresses temperature in terms of the results of a $CO_2$ doubling experiment.

## 3.4 | Complex climate sensitivities and the annual cycle

The essential FEBE behaviour can be understood by taking its Fourier transform ("F.T."). Using ${}_{-\infty}D_t^h \overset{F.T.}{\leftrightarrow} (\mathrm{i}\omega)^h$ (e.g. Podlubny, 1999) and taking $sF(t) = \delta(t)$ we obtain the Fourier transform of the Green's function $\widetilde{G}_{0,H}(\omega)$ and solution $\widetilde{T}(\omega)$:

$$ ((\mathrm{i}\omega\tau)^h + 1)\widetilde{G}_{0,h}(\omega) = 1; \quad \widetilde{G}_{0,h}(\omega) = \frac{1}{1+(\mathrm{i}\omega\tau)^h} $$
$$ \widetilde{T}(\omega) = s\widetilde{G}_{0,h}(\omega)\widetilde{F}(\omega) = \frac{s\widetilde{F}(\omega)}{1+(\mathrm{i}\omega\tau)^h} \tag{26}$$

where the tilde indicates Fourier transform with conjugate variable $\omega$, the frequency. This shows that if the forcing is purely sinusoidal $F = F_s\mathrm{e}^{\mathrm{i}\omega t}$ that the temperature response is also sinusoidal with the same frequency but different phase: $T = T_s\mathrm{e}^{\mathrm{i}\omega t}$. Using the notion of thermal impedance identified with the climate sensitivity, we can follow engineering practice (useful for estimating diurnal temperature – heating lags in buildings and structures) and use complex thermal impedances, that is, complex climate sensitivities $s(\omega)$:

$$ T_s = s(\omega)F_s; \quad s(\omega) = Z(\omega) = \frac{s_0}{1+(\mathrm{i}\omega\tau)^h}, \tag{27}$$

where we have used the notation $s_0 = s(0)$ for the static climate sensitivity (i.e. $s_0 = s$ in the previous notation). In the context of the Earth's energy balance, it is more useful to think in terms of sensitivities than impedances so that when convenient we use $s(\omega)$. We could mention that in dynamical systems theory, the Fourier filter between the forcing and the response (the impedance $s(\omega) = Z(\omega)$) is also called the "susceptibility", see for example Lucarini *et al.* (2017) for applications to climate models.

Deterministic scale-dependent climate sensitivities have been used in scaling contexts in order to account for the temperature fluctuations and spectra at multicentennial/multimillennial scales, for example, Rypdal (2012) proposed frequency-dependent scaling sensitivities to account for the solar cycle and Rypdal and Rypdal (2014) for the temperature response to Gaussian white-noise internal variability. Anticipating possible nonlinear feedbacks at these scales, Lovejoy and Schertzer (2012b) used scale-dependent but stochastic sensitivities to account for responses to stochastic (and scaling) external natural forcings. In this framework, at multicentennial/ multimillennial scales, the deterministic temperature response to external forcing could be nonlinear while the statistics of the corresponding fluctuations could still be linear.

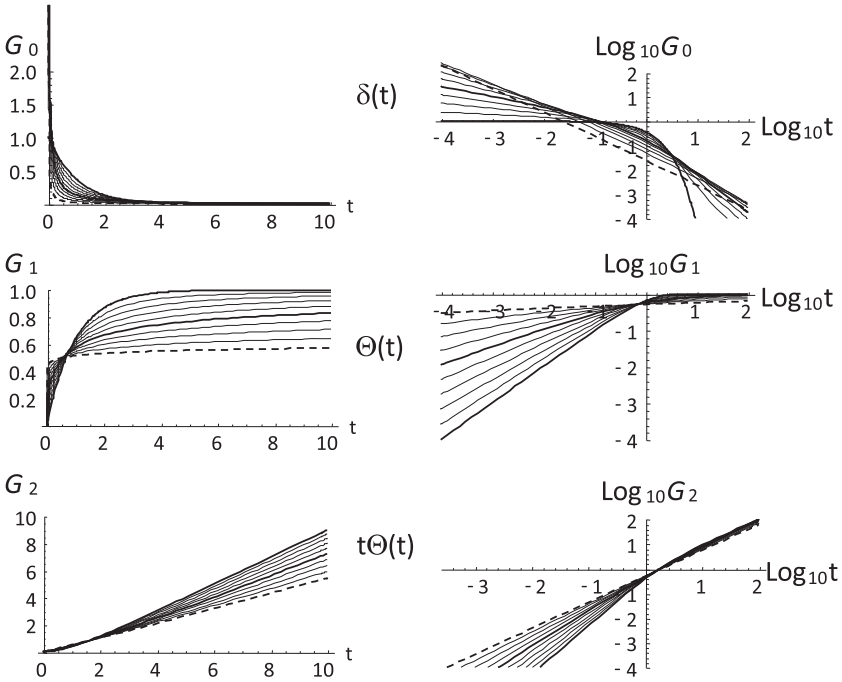


**FIGURE 2** The various response functions linear–linear (left column) and log–log (right) for $h = 1/10$ to 1 (exponential) in steps of 1/10. The lines bounding the envelope are $h = 1/10$ (dashed), $h = 1$ (thick). The middle value ($h = 1/2$) is also thick. The top row shows the Dirac ($\delta(t)$) response functions ($G_0$), the middle row the step ($\Theta(t)$) responses ($G_1$), and the bottom row, the ramp ($t\Theta(t)$) responses ($G_2$)

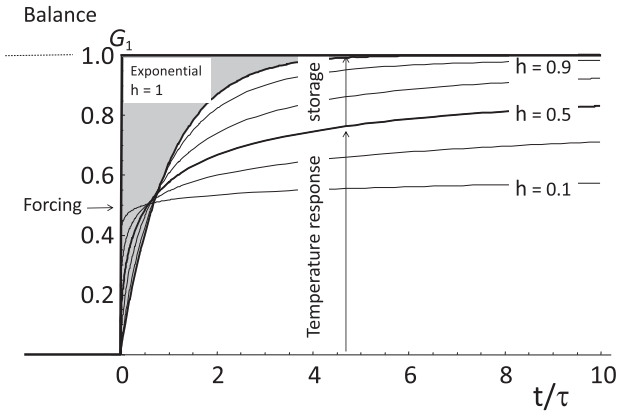


**FIGURE 3** The FEBE response to a step-function forcing ($G_1 = G_\Theta$ shown in thick) for various values of $h$, using $s = 1$ with the response for $h = 0.1, 0.3, 0.5, 0.7, 0.9$ (bottom to top). The thick middle line is for $h = 0.5$, the value close to the empirical exponent; the blue is for the box model value $h = 1$. The storage is the difference between the response and the forcing; both are shown with arrows for the case $h = 0.5$

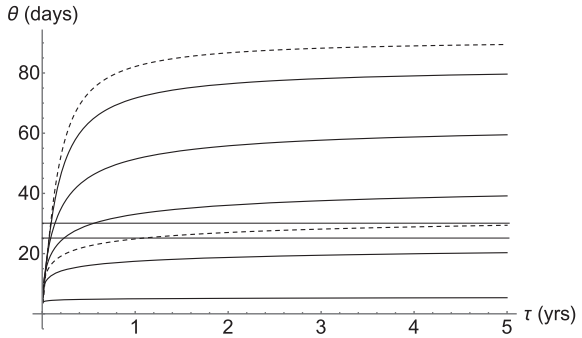


**FIGURE 4** The (negative) phase lag of the temperature with respect to the forcing (the negative of the argument of $s$) lag in days for $h = 1/10$, 3/10, ½, 7/10, 9/10 (black), 4/10 (dashed middle), 1 (dashed top) as a function of $\tau$ measured in years. The horizontal lines are for 25 and 30 days, we see that the dashed $h = 4/10$ falls in this range for between one and 5 years. This range of lags is close to the typical extratropical lags found in Donohoe *et al.* (2020)

Due to the low-frequency cut-off in Equation (27), FEBE sensitivities are complex climate sensitivities but with different high- and low-frequency scaling regimes. Complex sensitivities are useful in understanding the responses to periodic forcings, in particular when the forcing is due to the annual cycle. Figures 4 and 5 compare the phases and amplitudes of $s(\omega)$ as functions of the relaxation time $\tau$ (with $\omega = 2\pi$ rad·year$^{-1}$) for various values of $h$ including the HEBE value ($h = 1/2$) and the EBE value $h = 1$. From Figure 4, the thin horizontal lines show that taking the empirical value of $\tau$ in the range 1–5 years, and the approximate empirical value $h \approx 0.4$ (dashed), that the lag (the phase of the sensitivity) is in the range 25–30 days, which is in the observed range for the lag between the summer forcing maximum and maximum temperatures over most land areas.

If the forcing was only due to conductive heating (rather than conductive-radiative forcing), then, as pointed out in L1, one obtains the classical ($\pi/4$) lag (corresponding to $h = 1/2$, $\tau = \infty$) obtained by Brunt (1932) for the diurnal lag. This pure heating lag corresponds to 46 days and is already too long for most of the globe. Indeed, from the detailed maps in Donohoe *et al.* (2020) we estimate that in the extratropical regions, over land,

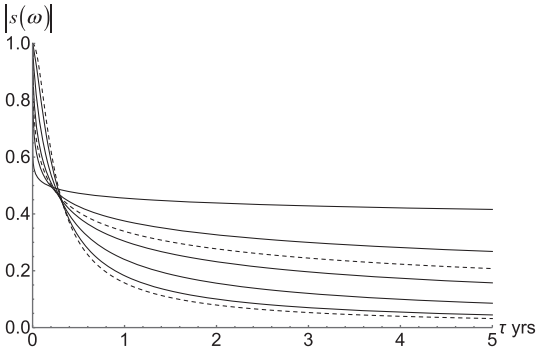


**FIGURE 5** The modulus of the climate sensitivity as a function of the relaxation time $\tau$ for $h$ ranging from 1/10, 3/10, 5/10, 7/10, 9/10 (solid, top to bottom) and $h = 4/10$ (middle, dashed) and $h = 1$ (bottom, dashed)

the summer temperature maximum is typically 30–40 days after the solstice, but only 20–30 days after the maximum forcing (insolation) and for ocean, 60–70 days after the solstice but only 30–40 days after the maximum insolation. Similarly, the $h = 1$ EBE lag is in the range 82 to 91 days (for $\tau > 1$ year), also much too long. While it is true that, over the ocean, the lag is typically longer than over land, this is probably because of the strong albedo periodicity associated with seasonal ocean cloud cover (Stubenrauch *et al.*, 2006; Donohoe *et al.*, 2020). This delays the summer solstice forcing maximum over the ocean, potentially explaining the extra lag.

L1 tested out the implications of the HEBE ($h = 1/2$) model using a latitudinally averaged model and parameters taken from North *et al.* (1981). The model uses a second-order Legendre polynomial to take into account the latitudinal variations and a sinusoidal annual cycle with empirically fitted parameters. Since it only models latitudinal variations, it effectively averages over land and ocean. In work in progress, we use modern satellite data to test the model at 2° spatial resolution. Future work will use these new results coupled with modern higher-resolution data of outgoing long-wave radiation to investigate this further and obtain more accurate local climate sensitivities.

## 4 | ESTIMATING $h$, $\tau$

### 4.1 | The low-frequency response and climate projections

The FEBE has distinct high- and low-frequency behaviours that can be understood from the complex climate sensitivity:

$$\widetilde{T}(\omega) = s(\omega)\widetilde{F}(\omega) \approx \begin{matrix} s_0(\mathrm{i}\omega\tau)^{-h}\widetilde{F}(\omega); & \omega >> \tau^{-1} \\ s_0(1-(\mathrm{i}\omega\tau)^{h})\widetilde{F}(\omega); & \omega << \tau^{-1} \end{matrix},$$

where the approximations were obtained as usual by expanding the denominator of Equation (27). In the high-frequency limit, $(\mathrm{i}\omega\tau)^{-h}\widetilde{F}(\omega)$ is the Fourier Transform of the $h$-order fractional integral of $F$, therefore at high frequencies the temperature integrates (smooths) the forcing whereas at low frequencies, the lowest-order term (unity) corresponds to a Dirac function (associated with the equilibrium temperature) while the next order $(\mathrm{i}\omega\tau)^{h}\widetilde{F}(\omega)$ term corresponds to a fractional differentiation of order $h$ that determines the rate at which equilibrium is approached.

Alternatively and equivalently, we can use small and large $t$ expansions of the Green's functions (Equation (21)). Let us first consider the large $t$ asymptotic expansions:

$$G_{\zeta,h}(t) = \tau^{\zeta-1}\sum_{n=0}^{\infty}\frac{(-1)^n}{\Gamma(\zeta-nh)}\left(\frac{t}{\tau}\right)^{\zeta-1-nh}; \qquad \zeta \geq 0 \tag{29}$$

(Podlubny, 1999). The most important cases are for $\zeta = 0$, 1, 2, corresponding to impulse, step and ramp responses:

$$\begin{aligned} G_{0,h}(t) &= -\frac{\tau^{-1}}{\Gamma(-h)}\left(\frac{t}{\tau}\right)^{-1-h} + \frac{\tau^{-1}}{\Gamma(-2h)} \\ &\quad \times \left(\frac{t}{\tau}\right)^{-1-2h} - \mathrm{O}\left(\frac{t}{\tau}\right)^{-1-3h} \\ G_{1,h}(t) &= \left(1 - \frac{1}{\Gamma(1-h)}\left(\frac{t}{\tau}\right)^{-h} + \frac{1}{\Gamma(1-2h)}\left(\frac{t}{\tau}\right)^{-2h}\right) \\ &\quad + \mathrm{O}(t^{-3h}) \;\; t >> \tau; \;\; 0 < h < 2 \\ G_{2,h}(t) &= t - \frac{\tau}{\Gamma(2-h)}\left(\frac{t}{\tau}\right)^{1-h} + \frac{\tau}{\Gamma(2-2h)}\left(\frac{t}{\tau}\right)^{1-2h} \\ &\quad + \mathrm{O}(t^{1-3h}). \end{aligned} \tag{30}$$

We have used $\Gamma(0) = \infty$ for $\zeta = 0$, note that $\Gamma(-h) < 0$ for $0 < h < 1$. From these, we see that, as required for the step response, $G_{1,h}(0) = 0$ and $G_{1,h}(\infty) = 1$ so that at large $t$, the FEBE implies a power-law relaxation to the new equilibrium temperature.

From either the low-frequency Fourier approximation (Equation (28)) or the asymptotic expansion (Equation (30)), we see that the low-frequency response $T_{\mathrm{low}}$ can be approximated by the first term: $G_{\mathrm{low},h}(t) \approx G_{0,h}(t)$:

$$T_{\mathrm{low},h}(t) \approx sG_{\mathrm{low},h} * F; \;\; G_{\mathrm{low},h}(t) = -\frac{\tau^{-1}}{\Gamma(-h)}\left(\frac{t}{\tau}\right)^{-1-h}. \tag{31}$$

$T_{\mathrm{low}}$ is expected to be dominated by the forced response to external forcings, appropriate for making multi-decadal temperature projections from anthropogenic forcings (in Equation (31), we assumed that the initial temperature is zero, otherwise the extra term in Equation (20) can be used). Comparing this with the Hébert *et al.* (2021)

truncated scaling CRF (Equation (1)) we see that the FEBE has identical low-frequency behaviour with:

$$h = -H_F. \tag{32}$$

Using the empirical value $H_F \approx -0.5$ that Hébert *et al.* (2021) obtained by (Bayesian) fitting the global temperature series since 1880 to the IPCC $CO_{2eq}$ forcing, we find $h \approx 0.5$. Similarly, Procyk *et al.* (2020) fitted the FEBE (using the exact $G_{0,h}$ Green's function with the same centennial-scale climate forcing) to estimate $h = 0.38 \pm 0.06$. In the next section we consider the high-frequency behaviour and use the internal variability (assumed to be forced by a Gaussian white noise) that gives the more precise estimate $h \approx 0.42 \pm 0.02$ (del Rio Amador and Lovejoy, 2019).

We have seen (Equation (15)) that $\tau$ quantifies the rate that the temperature relaxes to equilibrium following a perturbation. At the same time, according to the above analysis, $\tau$ represents the transition time between two power-law regimes $(t/\tau)^{h-1}$ to $(t/\tau)^{-h-1}$; using the empirical value $h \approx 0.4$ this is a transition from $(t/\tau)^{-0.6}$ ($t \ll \tau$) to $(t/\tau)^{-1.4}$ ($t \gg \tau$) behaviour thus with a relatively small change in the scaling. The exact FEBE solution (Equation (20), Figure 2, right-hand column) confirms that the transition is very slow – that is, it occurs over a wide range of scales. This very slow transition makes $\tau$ difficult to accurately estimate from observations of the temperature responses to forcings (see Appendix).

Based on the truncated power-law CRF (Equations 1 and 2), Hébert *et al.* (2021) used physical arguments to determine $\tau$. At short enough time-scales, different parts of the globe have temperature fluctuations that vary more or less independently of each other; they are primarily responses to "internal" forcings, small-scale storage mechanisms. At long enough scales, the variability is expected to be primarily due to the responses from external forcings. Hébert *et al.* (2021) argued that $\tau$ could be estimated from the scale at which the whole global surface temperature starts to fluctuate as a unit. In particular, the atmosphere and ocean should have temperatures that fluctuate synchronously; $\tau$ should be the atmosphere–ocean coupling time-scale.

$\tau$ was therefore estimated by considering the correlations between fluctuations in the atmospheric temperature averaged over land, and fluctuations in the sea-surface temperature (SST) averaged over the oceans. At scales of months, there is practically no correlation; the land and ocean fluctuate nearly independently of each other. However, at long time-scales, they are on the contrary highly correlated; they fluctuate together. The transition between independence and dependence occurs over a short interval: between about 6 months and 2 years. Hébert *et al.* (2021) used this to infer that $\tau \approx 2$ years, a value that gave reasonable climate projections through to 2100. This time-scale is essentially the same as the lifetime of planetary-scale ocean gyres: the ocean weather – ocean macroweather transition scale (Lovejoy *et al.*, 2017).

More recently, Procyk *et al.* (2020) used a Bayesian approach and the FEBE CRF to find the optimum parameters $\lambda$, $\tau$, $h$ constrained by the estimated historic $CO_{2eq}$ forcings and the responses as estimated by five observational series of global average temperatures since 1880; they obtain a very similar value $\tau \approx 4.7$ years (the 90% confidence interval was 2.4–7.0 years). Since the FEBE is in the same family as the EBE, as a Bayesian prior distribution for $\tau$, they used the Geoffroy *et al.* (2013) estimate of $4.1 \pm 1.1$ years for the box ($h = 1$) model based on a dozen different GCM model outputs.

## 4.2 | The high-frequency response: Internal variability and macroweather forecasts

In Section 4.1 we saw that if we take $h = -H_F$, that to leading order, the low-frequency FEBE and truncated model impulse CRFs are both proportional to $t^{-h-1}$, these give reasonable approximations to the low frequencies. But with the FEBE, we are no longer thinking in terms of a small number of homogeneous boxes; rather the FEBE is a phenomenological model of a hierarchy of storage processes from large to small. It is therefore logical to assume that the FEBE applies to both external and internal forcings. Mathematically this implies that both variabilities have the same Green's function, opening up the possibility of estimating climate parameters directly from the internal variability. Basing himself on the weather–macroweather scale separation and assuming that the internal forcing from the weather regime was a Gaussian white noise, Hasselmann (1976) proposed a stochastic climate model that reduces to the stochastic box-model ($h = 1$); in the zero-dimensional case, the latter is therefore sometimes called the "Hasselmann equation" (see Watkins *et al.* (2020) for a useful discussion, review as well as various fractional extensions).

The idea of using the internal variability to infer the response to external forcing goes back to at least Leith (1975), who proposed that the Fluctuation Dissipation Theorem (FDT) be used for this purpose. The FDT says that the impulse response CRF of a nonlinear dynamical system to (small) external forcings is equal to the normalized covariance function. The FDT is attractive since – if it is valid – it is independent of specific climate models. While the original FDT only applied

at thermodynamic equilibrium, it has been extended using linear response theory to nonequilibrium systems (Kubo, 1966), although the extension remains controversial (Gottwald *et al.*, 2016); it is not at all obvious that the climate system is close enough to equilibrium for the FDT to apply. Nevertheless, starting with Leith, numerous such attempts have been made (especially North *et al.* (1993), Cionni *et al.* (2004) and Schwartz (2007), and recently Cox *et al.* (2018)). As we show in a future publication, it turns out that the FEBE *does* obey the FDT, but only in the low-frequency regime where $t \gg \tau$ where the scaling storage term is small enough that the system is indeed close to equilibrium and the asymptotic form Equation (44) holds (in this limit the normalized correlation function and Green's function $G_{0,h}$ are equal). In contrast, the EBE ($h = 1$) special case is exceptional: the FDT always holds, and the temperature is an Ornstein–Uhlenbeck process (North *et al.*, 1993; Cox *et al.*, 2018).

Although at high frequencies the FDT is not valid, if there is information about the internal forcing, then the FEBE can still be used to infer the CRF and hence the system response to external forcing. For example, if we make the usual assumption that the internal forcing is a Gaussian white noise, then we can use the temperature statistics to deduce the exponent $h$ (see Table 1, discussed here and in the next section). If the noise amplitude is also known, we can deduce $s$ (see Table 1, bottom row). Before discussing the full stochastically forced case, in this section we discuss the high-frequency stochastic response dominated by scaling storage processes.

In order to investigate the high-frequency behaviour, use the first term of the Green's function expansion (Equations 18–20) to obtain:

$$T_{\text{high}} \approx sG_{\text{high}} * F(t);\ G_{\text{high}} = \frac{1}{\tau\Gamma(h)}\left(\frac{t}{\tau}\right)^{h-1};\quad t << \tau. \tag{33}$$

This is the singular high-frequency FEBE limit corresponding to a fractional integration of order $h$ (Equation (28)).

Since we associate the high-frequency behaviour primarily with internal variability, let us consider the case where its source $f(t)$ is a white noise. Ignoring external forcings so that $T_{\text{high}}$ is its stochastic response:

$$T_{\text{high}}(t) \approx \frac{s}{\tau^h\Gamma(h)}\int_{-\infty}^{t}(t-u)^{h-1}f(u)du. \tag{34}$$

To ensure convergence, we assume that $f(t)$ is a statistically stationary process representing the "innovations" with mean $<f(t)> = 0$ and we take $T(-\infty) = 0$.

To understand this high-frequency behaviour, it is useful to compare it with the definition of a fractional Gaussian noise process (fGn):

$$g_H(t) = \frac{K_H}{\Gamma(H+1/2)}\int_{-\infty}^{t}(t-u)^{H-1/2}\gamma(u)du;\quad -1 < H < 0;$$
$$K_H = \left(\frac{\pi}{2\cos(\pi H)\Gamma(-2-2H)}\right)^{1/2}, \tag{35}$$

where $\gamma(t)$ is a "δ-correlated" unit white-noise process, the innovations satisfying $\langle\gamma(t)\gamma(t')\rangle = \delta(t-t')$ and $<\gamma> = 0$. $g_H(t)$ is a generalized function – a "noise" – that is properly defined over finite integrals (see e.g. Biagini *et al.* (2008), who terms this "fractional white noise"; it is commonly defined as a derivative of fractional Brownian motion, see e.g. Kou and Xie (2004)). Often, fGn is defined in terms of a Wiener process $W$ with $dW = \gamma(t)dt$. $K_H$ is a standard normalization constant chosen for convenience. The fGn process goes back to Mandelbrot and Van Ness (1968) and Mandelbrot (1971) who used it as a hydrology model (see also Molz *et al.* (1997) for a review including hydrology applications). Explicit use of fGn processes as macroweather temperature models was proposed by Rypdal and Rypdal (2014) and Lovejoy (2015).

Comparing Equations 34 and 35, we see that if $f(t)$ is a white noise (i.e. proportional to $\gamma(t)$), and if we take: $H = h - 1/2$, then the high-frequency temperature response is an fGn process with exponent $H$. This result is easy to understand since at high frequencies (Equation (28)), the derivative term in the FEBE is dominant (i.e. ${}_{-\infty}D_t^h T >> T$) and $g_H$ is indeed the solution of the resulting fractional differential equation:

$${}_{-\infty}D_t^{H+1/2}g_H(t) = K_H\gamma(t);\ -1 < H < 0. \tag{36}$$

The standard assumption about internal variability is that it is forced by a Gaussian white noise. With this, we can use this high-frequency behaviour to estimate $h$ from the empirical values of $H$. As mentioned above, according to the global-scale Haar fluctuation analyses (see e.g. Lovejoy *et al.*, 2015; del Rio Amador and Lovejoy, 2019), $H \approx -0.08 \pm 0.02$ so that $h = H + 1/2 \approx 0.4$ which is essentially the same as the Hébert *et al.* (2021) and Procyk *et al.* (2020) values estimated from the low-frequency response to external forcings (Equation (31)). These are also close to estimates of $H$ in the range 0 to −0.3 for various ocean and land areas (Fredriksen and Rypdal, 2017) and with corresponding spectral exponents estimated in Lovejoy and Schertzer (2013).

**TABLE 1** Above the thick horizontal line: dimensional formulae summarizing auto- and cross-correlation functions of the white-noise forced FEBE; the autocorrelations are dimensionalized versions of those in L3

| General | $\Delta t \ll \tau$ | $\Delta t \gg \tau$ | Range |
|---|---|---|---|
| Temperature Autocorrelation function: $\langle T_{\tau_r}(t) T_{\tau_r}(t-\Delta t)\rangle = R_{TT,\tau_r}(\Delta t)$ | $\frac{s^2\sigma^2_{f_{\tau r}}}{K_h^2}\left(\frac{\tau_r}{\tau}\right)\left[h(1+2h)\left(\frac{\Delta t}{\tau}\right)^{2h-1}+\ldots\right]; \tau_r << \Delta t$<br>$R_{T,\tau_r}(0)=\sigma^2_{T,\tau_r}=\frac{s^2\sigma^2_{f_{\tau r}}}{K^2}\left(\frac{\tau_r}{\tau}\right)^{2h}+\ldots$ | $-\lambda^2\sigma^2_{f,\tau_r}\left(\frac{\tau_r}{\tau}\right)\frac{(\Delta t/\tau)^{-h-1}}{\Gamma(-h)}$ | $0<h<1/2$ |
| | $\frac{s^2\sigma^2_{f,\tau_r}}{c^2}\left(\frac{\tau_r}{\tau}\right)\left[1-\frac{\lvert\Gamma(1-2h)\rvert\sin(\pi h)}{\pi}c^2\left(\frac{\Delta t}{\tau}\right)^{2h-1}+\ldots\right]$ | $-\lambda^2\sigma^2_{f,\tau_r}\left(\frac{\tau_r}{\tau}\right)\frac{(\Delta t/\tau)^{-h-1}}{\Gamma(-h)}$ | $1/2<h<1$<br>$1<h<3/2$ |
| | $\frac{s^2\sigma^2_{f,\tau_r}}{c^2}\left(\frac{\tau_r}{\tau}\right)\left[1-\left(\frac{\Delta t}{\tau}\right)^2\frac{c^2}{2}\int_0^\infty G'_{0,h}(u)^2 du+\ldots\right]$ | $-\lambda^2\sigma^2_{f,\tau_r}\left(\frac{\tau_r}{\tau}\right)\frac{(\Delta t/\tau)^{-h-1}}{\Gamma(-h)}$ | $3/2<h<2$ |
| | $\frac{s^2\sigma^2_{f,\tau_r}}{2}\left(\frac{\tau_r}{\tau}\right)e^{-\Delta t/\tau}$ | | $h=1$ |
| Temperature-forcing cross-correlation function: $\langle T_{\tau_r}(t) f_{\tau_r}(t-\Delta t)\rangle = R_{Tf,\tau_r}(\Delta t)$ | $s\sigma^2_{f,\tau_r}\left(\frac{\tau_r}{\tau}\right)G_{0,h}\left(\frac{\Delta t}{\tau}\right); \tau_r << \tau;\ \tau_r << \Delta t$<br>$R_{Tf,\tau_r}(0)=\langle T_{\tau_r} f_{\tau_r}\rangle=\frac{s\sigma^2_{f,\tau_r}}{\Gamma(2+h)}\left(\frac{\tau_r}{\tau}\right)^h;\ \tau_r << \tau$<br>$\rho_{Tf}=\frac{\langle T_{\tau_r} f_{\tau_r}\rangle}{\sigma_{T,\tau_r}\sigma_{f,\tau_r}}=\frac{K}{\Gamma(2+h)}$ | | $0<H<2$ |
| Variance ratio of temperature versus forcing: $\langle T^2_{\tau_r}\rangle/\langle f^2_{\tau_r}\rangle$ | $s^2K^{-2}\left(\frac{\tau_r}{\tau}\right)^{2h}$<br>$s^2c^{-2}\frac{\tau_r}{\tau}$ | | $0<h<1/2$<br>$1/2<h<2$ |

*Note:* $\tau$ is the relaxation time, $\tau_r$ is the resolution. $T_{\tau_r}(t)$ is a resolution $\tau_r$ fRn process, the special $h=1$ (EBE) is the Ornstein–Uhlenbeck process; in the $h=1/2$ (HEBE) special case, there are logarithmic corrections to $R_{TT,\tau r}$ see appendix B of L3. To ensure small-scale convergence, all quantities are taken at resolution $\tau_r$, so that $\langle f^2_{\tau_r}\rangle = \sigma^2/\tau_r$ (for $R_{TT,\tau r}$, this is only important for $h \le 1/2$, for $R_{Tf,\tau r}$, it is important for $h<1$). Also shown (bottom row) is the temperature to the forcing variance ratio. The numerical constant $K$ is from Equation (35) with $H=h-1/2$ and $c$ is from Equation (42). Note the exact value $c^2=2$ for $h=1$ and that $\Gamma(-h)<0$ for $0<h<1$. When $h=1/2$ there are small-scale logarithmic divergences.

We could note that Cox *et al.* (2018) recently used an analogous method of exploiting the internal variability to estimate the ECS. However, they made the unrealistic Ornstein–Uhlenbeck ($h=1$) assumptions that the autocorrelation was exponential and that the ratio of temperature variance to forcing variance follows the (strongly biased) $h=1$ result (Table 1, bottom row).

fGn processes are power-law smoothed white noises (Equation (35)), but when $-1/2 < H \leq 0$, ($0 < h \leq 1/2$), the smoothing is not enough to eliminate the small-scale divergences (i.e. mathematically, they remain "noises"). For this reason, finite-resolution fGn processes are defined as the increments of fractional Brownian motions (fBm). This is equivalent to defining the "$\tau_r$ resolution fGn" by averaging over a time interval $\tau_r$:

$$g_{H,\tau_r}(t) = \frac{1}{\tau_r} \int_{t-\tau_r}^{t} g_H(t')dt'. \quad (37)$$

Then, with the above choice of constant $K$ we have:

$$\langle g_{H,\tau_r}(t)^2 \rangle = {\tau_r}^{2H}; -1 < H < 0. \quad (38)$$

If $h < 1/2$, then $H = h - 1/2 < 0$, implying a small-scale divergence as the resolution $\tau_r$ is reduced; fGn processes thus have strong resolution dependencies. An important but regularly overlooked practical consequence of such divergences is that in macroweather, this leads to "space–time reduction factors" that multiplicatively bias macroweather anomalies and are important for example in accurately estimating the temperature of the Earth (Lovejoy, 2017a).

This is perhaps an opportune moment to comment on the relation between $h$ and the Hurst-inspired exponent $H$. Although Hurst's original exponent (Hurst, 1951) was for the "rescaled range" of a process, at least for many statistically stationary processes (such as fGn), it was equal to the scaling exponent of fluctuations of running sums. At first, Mandelbrot and Van Ness (1968) considered Gaussian processes that had average (absolute) fluctuations growing with scale ($\Delta t^H$ with $H > 0$): fractional Brownian motion (fBm). If we define fluctuations using differences (increments) then the fBm fluctuation exponent $H$ is directly related to the order ($H + 1/2$) of the fractional integration of a white noise.

Later, in the 1980's, wavelets allowed fluctuations to be defined much more generally than just as differences, so that fluctuations with exponents $H < 0$ could be obtained. For example, we can define the anomaly fluctuation at scale $\Delta t$ as the mean over the interval $\Delta t$ of the process with its long-term average removed – or with zero ensemble average (Lovejoy and Schertzer, 2012a). In this case, Equation (37) shows that for fGn with $\Delta t = \tau_r$, that we can obtain fluctuation exponents in the range $-1 < H < 0$, see Equation (38) (for fGn, the exponent of the average absolute anomaly and RMS anomaly are equal).

Also in the 1980s, it was realized that scaling processes in general, and atmospheric processes in particular, are more generally multifractal so that they could not be characterized by a single exponent. However, Schertzer and Lovejoy (1987) showed that multifractal processes could nevertheless also be characterized by fluctuation exponents and used the Hurst-inspired symbol "$H$" for this. Beyond sharing a fluctuation exponent $H$ that could take on any real value, both multifractal and Gaussian processes could be generated by fractional integration: for multifractals, by fractionally integrating a "conservative" multifractal cascade process by order $H$, while for quasi-Gaussian processes by order $H + 1/2$ fractional integration of a Gaussian white noise (the extra $1/2$ is needed because the white noise itself has a fluctuation exponent equal to $-1/2$).

In summary, if we use appropriate wavelets to define the fluctuations and define $H$ as a fluctuation exponent, then it applies equally well to both quasi-Gaussian and multifractal processes.

In the FEBE, the order $h$ plays numerous roles including several that relate it to an $H$. For example, at high frequencies (when the storage term dominates), $h$ is the order of fractional integration needed to obtain $T$ from the forcing, so that when $F$ is a Gaussian white noise, at high frequencies $H = h - 1/2$. However, $F$ can also be deterministic in which case $h$ characterizes the responses to deterministic step forcings for short times ($t^h$), while for long times, equilibrium is approached as $t^{-h}$.

Before leaving the topic, two more comments are in order. First, although the domain of the FEBE discussed here – macroweather in time – has low intermittency (it is not far from being quasi-Gaussian, we model it as such), it is in fact the only dynamical atmospheric regime where the multifractality is low. Yet in the spatial domain, macroweather multifractality is on the contrary very high (Lovejoy, 2018), so that for regional (space–time FEBE: Lovejoy, 2021a; 2021b) multifractality is important. Second, it is worth pointing out that there is also a large mathematical literature on fBm and fGn processes that characterize them parametrically. Adding primes to avoid confusion, they discuss "$H'$ parameter fBm" and the "$H'$ parameter fGn" processes. For reference, for fBm, we have $H = H'$ whereas for fGn, $H = H' - 1$; the difference (unity) is due to differencing/summing of order one needed to obtain fGn from fBm (and vice versa).

### 4.3 | The stochastic FEBE, fractional Relaxation motion (fRm), fractional Relaxation noise (fRn)

In the previous section we considered the high-frequency stochastic response, fGn, which is the solution of the FEBE at high frequencies where the (fractional) storage term dominates. We noted that since the FEBE is a linear equation, we could separately model the deterministic and stochastic (internal) variability. In this section we give more information about this pure stochastic case driven by white noise innovations, and we briefly summarize and expand upon some of the results in Lovejoy (2019b) (hereafter L3). We note that the stationary process generated by the stochastic FEBE – fractional Relaxation noise (fRn) – generalizes Ornstein–Uhlenbeck processes from $h = 1$ to $0 < h < 2$.

Whereas the deterministic fractional relaxation equation has been well studied (see e.g. Miller and Ross, 1993; Podlubny, 1999), the stochastic version has enjoyed much less attention, although West *et al.* (2003) reviews some applications of the Riemann–Liouville case to nonstationary random walks. In contrast, the stochastic infinite range (Weyl version, when $t_0 \to -\infty$) leading to stationary solutions has only recently been considered (L3) including the optimum predictor problem (L3). Below we discuss the main points and give the dimensional form of the main statistical properties.

Although the stochastic fractional relaxation equation has received little attention, a closely related equation, the Fractional Langevin Equation (FLE), has received somewhat more study due to its applications in the physics of random walks, and diffusion (Mainardi and Pironi, 1996; Coffey *et al.*, 2012). Whereas the highest-order term in the fractional relaxation equation is fractional, in the FLE the fractional term is of lowest order. Watkins *et al.* (2020) provides a useful review that compares and contrasts various stochastic equations that appear in climate models and points out the relationship between the FLE and the FEBE.

The physical problem that we wish to solve is the noise-driven FEBE with noise amplitude $\sigma_f$ and sensitivity $s$:

$$\tau^h {}_{-\infty}D_t^h T + T = sf(t); f(t) = \sigma\gamma_f(t), \tag{39}$$

with initial conditions $T(-\infty) = 0$, $0 \le h \le 1$ and $\gamma(t)$ a unit Gaussian white noise so that $\langle f^2\rangle^{1/2} = \sigma_f$; $\langle f\rangle = 0$ , $\sigma$ is the amplitude of the noise, Equation (39) is a fractional Langevin equation (see L3). To understand the statistical properties of the internal variability response, it suffices to study the non-dimensional equation:

$${}_{-\infty}D_t^h U_h + U_h = N_h\gamma(t), \tag{40}$$

where for $0 < h < 1/2$, $N_h = K_h$ is the normalization constant in Equation (35) (for $1/2 < h < 2$, see Equation (41) and the nondimensional $U_h(t)$ function is called a fractional Relaxation noise (fRn) since it generalizes fGn (see L4). It also generalizes the $h = 1$ Ornstein–Uhlenbeck process.

Using $U_h$ and noting that $\gamma(t) \overset{d}{=} \tau^{-1/2}\gamma\left(\frac{t}{\tau}\right)$, we can obtain the solution to the dimensional Equation (39) using:

$$T(t) = \frac{s\sigma_f}{N_h\tau^{1/2}} U_h\left(\frac{t}{\tau}\right). \tag{41}$$

"$\overset{d}{=}$" means equality in a probability sense.

Since the fRn process is the solution of the fractional relaxation equation with a stationary, Gaussian, zero mean, white-noise forcing, it is also stationary, Gaussian with zero mean. Its statistics are therefore fully characterized by its autocorrelation function. A complication in the calculation is that when $0 < h \le 1/2$, in the small $t$ limit, the fractional term of Equation (39) dominates so that we obtain the fGn limit. The solution of Equation (40) is therefore – like $\gamma(t)$ – a generalized function; to obtain solutions with finite variances, we must take averages over finite resolutions $\tau_r$.

The resulting $\tau_r$ resolution autocorrelation function at lag $\Delta t$ is:

$$\begin{aligned}
&R_{h,\tau_r}(\Delta t) = \langle U_{h,\tau_r}(t)U_{h,\tau_r}(t-\Delta t)\rangle \\
&\qquad = N_h^2 \int_0^\infty G_{0,h}(\Delta t + u)G_{0,h}(u)du; \\
&R_{h,\tau_r}(0) = \langle U_{h,\tau_r}(t)^2\rangle = \tau_r^{2h-1}; \quad 0 < h < 1/2 \\
&R_{h,\tau_r}(0) = 1; \quad 1/2 < h < 2 \\
&N_h = K_h; \quad 0 < h < 1/2 \\
&N_h = c; \quad c^{-2} = \int_0^\infty G_{0,h}(u)^2 du; \quad 1/2 < h < 2
\end{aligned} \tag{42}$$

Note that when $0 < h < 1/2$, the equation for $R_{h,\tau_r}(\Delta t)$ is only valid for $\Delta t \ge \tau_r$, also recall the classical Ornstein–Uhlenbeck process ($h = 1$) where $G_{0,1}(t) = e^{-t}$ and the autocorrelation is: $R_1(\Delta t) = e^{-\Delta t}/2$. The above formulae were given for $\Delta t > 0$ but since $G_{0h}$ is causal, $G_{0h}(t) = 0$ for $t \le 0$, which implies $R_h(\Delta t) = R_h(-\Delta t)$. Table 1 gives the explicit dimensional formulae taken from L3 for the temperature autocorrelation. Also given in the Table are the corresponding formulae for the temperature-forcing ($f = \sigma_f\gamma$) and temperature storage cross-correlations ($\dot{S} = dS_h/dt = \tau^h {}_{-\infty}D_t^h T = sf - T$). Physically, the model is justified as long as the resolution $\tau_r$ is greater than the weather–macroweather transition scale $\tau_w$ (≈10 days). $\tau_w$ is the inner (smallest) scale over which the FEBE may expected to be valid. From this, we can obtain the high-frequency fGn approximation (valid

for $\Delta t \ll 1$ corresponding to $\Delta t \ll \tau$ in the dimensional equation):

$$R_{h,\tau_r}(\Delta t) \approx h(2h+1)\Delta t^{2h-1};\ \tau_r << \Delta t << 1; \quad 0 < h < \tfrac{1}{2} \tag{43}$$

(see Table 1 for $1/2 < h < 2$). At low frequencies, for $\Delta t \gg 1$, we obtain:

$$R_h(\Delta t) = -\frac{N_h^2}{\Gamma(-h)}\Delta t^{-1-h} + O(\Delta t^{-1-2h}) : \begin{matrix} 0 < h < 1 \\ 1 < h < 2 \end{matrix}; \quad \Delta t >> 1,$$

valid for all (non-negative) $\Delta t$, see L3 and Table 1 for the dimensional expressions and note that $\Gamma(-h) < 0$ for $0 < h < 1$. This large $\Delta t$ result is independent of the resolution $\tau_r$ and the fact that it holds over a wider range of $h$ values ($h = 1$ is the exponential exception, see Table 1). A technical point is that, although when $0 < h < 1/2$, the high-frequency limit is the fGn with a huge memory due to the slow $\Delta t^{2h-1}$ fall-off in $R_h(\Delta t)$ (Equation (43)), the corresponding fRn is effectively truncated with a memory of the order of $\tau$ (see L3).

Table 1 summarizes the (dimensionalized) versions of these relations along with the corresponding cross-correlation functions. Of particular note are parameter ranges where the resolution is important ($0 < h \leq 1/2$ for $R_{TT,\tau r}$ and $0 < h < 1$ for the temperature – forcing cross-correlation $R_{Tf,\tau r}$) as well as the exceptionally particularly rapid fall-off at large $\Delta t$ for the temperature–storage cross-correlation $R_{TS,\tau r}$. Other properties of solutions to the stochastic Equation (40) are given in L3.

A final interesting property is the predictability of fRn, also discussed in L3, since for times $< \tau$, fRn is close to the long-memory fGn process, the two can both be well predicted (for fGn, the skill – with infinite past data – becomes perfect in the limit $h \to 1/2$ [$H \to 0$]); however, beyond the relaxation time $\tau$, fRn cannot be well predicted, it approaches a white noise.

## 4.4 | ECS and TCR

In order to make accurate climate projections throughout the twenty-first century, the single most important parameter is the climate sensitivity $s$ since for any given forcing, this determines the ultimate (long-time) temperature change. It is conventional to express $s$ in terms of the change in temperature following a $CO_2$ doubling, this $s$ is the "Equilibrium Climate Sensitivity" (ECS). This is effectively a change in units and can be effected using the canonical value 3.71 W·m$^{-2}$ for a $CO_2$ doubling. However, in order to make accurate projections through the twenty-first century, it is also important to have accurate estimates of the memory, the "warming in the pipe" (Hansen *et al.*, 2011). The standard way of quantifying this is by comparing the ECS to the Transient Climate Response (TCR) that is defined as the temperature change following a $CO_2$ doubling with forcing linearly increasing over 70 years; since the $CO_2$ forcing is logarithmic in the $CO_2$ concentration, this is nearly exactly a 1% increase per year. The smaller the TCR/ECS ratio, the larger the memory. Note that in the following, we only consider static climate sensitivities that we indicate by $s$ (not $s_0$).

The dimensionless ratio TCR/ECS is independent of $s$; for the FEBE model, it is only a function of $\tau$, $h$, and theoretically – since the TCR is defined in terms of a "ramp" – the ratio is:

$$TCR/ECS = G_{2,h}(\Delta t)/\Delta t = 1 - \frac{1}{\Gamma(2-h)}\left(\frac{\Delta t}{\tau}\right)^{-h} + \frac{1}{\Gamma(2-2h)}\left(\frac{\Delta t}{\tau}\right)^{-2h} + \ldots; \quad \Delta t >> \tau, \tag{45}$$

where $\Delta t$ is the period over which the doubling occurs (conventionally 70 years).

Figure 6 shows a contour plot of the ratio as a function of exponent $h$ and relaxation times $\tau$. The FEBE parameters (black circle) estimated from the historical data and IPCC AR5 forcing by Procyk *et al.* (2020) with 90% confidence intervals are shown (dashed rectangle). One can see that the ratio is in the range ≈0.65–0.75 with mean ≈0.7. One can also see that the single-box ($h = 1$) model has ratios near 1 unless the relaxation time is 20 years or longer ($\Delta t/\tau < \approx 3.5$).

As $h$ approaches zero, we see that the transition from the ramp response to the equilibrium response takes a longer and longer time. In the limit $h \to 0^+$; the ratio TCR/ECS = 1/2, an exact result as can be easily seen by using $\lim_{h\to 0^+} D^h = 1$ in the FEBE (Equation (14)).

The $h = 1$ (exponential) case has too short a memory to be realistic so that most GCM outputs have been fitted to double exponential (2-box) models; for example, the IPCC AR5 suggests box relaxation times of 8.5 and 409.5 years with TCR/ECS = 0.58. Similarly, the IPCC AR5 (chapter 9, see also Yoshimori *et al.*, 2016), obtained TCR/ECS = 0.56 ± 0.19 (90% confidence limits) from 23 CMIP5 GCMs. Figure 7 shows these model-based estimates compared to the FEBE estimates. We can see that the two are compatible although the FEBE data-based estimate has a much smaller uncertainty than the model-based one.

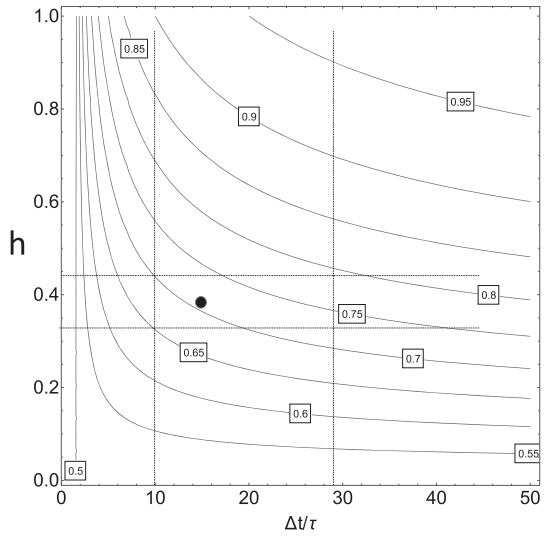


**FIGURE 6** A contour plot of TCR/ECS for various $h$ values as well as the dimensionless ratio of the ramp time $\Delta t$ to the relaxation time $\tau$; the canonical value of $\Delta t$ is 70 years. The dashed lines show the 90% confidence intervals from the analysis in Procyk *et al.* (2020) for both $\tau$ and $h$; (using $\Delta t = 70$ years) the black circle represents the median values

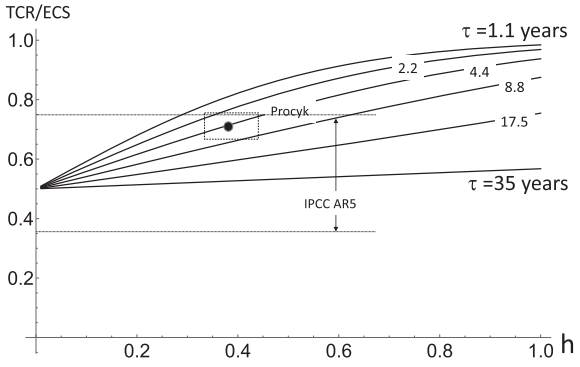


**FIGURE 7** The TCR/ECS ratio for the FEBE model as a function of $h$ for various relaxation times using the canonical ramp time constant $\Delta t = 70$ years with relaxation time $\tau$. The black circle and rectangle shows the mean and 90% confidence interval values found by Procyk *et al.* (2020). The horizontal dashed lines show the 90% confidence limits (IPCC AR5, chapter 9) found using CMIP5 GCMs

## 4.5 | Modelling the internal and externally forced variability: A simple FEBE model from 1825–2100

We have seen that with a single exponent $h$, the FEBE correctly predicts both high- and low-frequency scaling regimes, allows it to convincingly model the internally and externally forced variability; the two are linked by the Earth's multiscale (and scaling) storage mechanisms. Let us therefore treat the total forcing $\mathfrak{F}(t)$ as the sum of deterministic part $F(t)$ and a stochastic part $f(t)$ with $<f(t)> = 0$:

$$\mathfrak{F}(t) = F(t) + f(t), \tag{46}$$

and similarly for the temperature response:

$$T(t) = T_e(t) + T_i(t), \tag{47}$$

where the subscripts "e" and "i" are for externally forced and internal respectively. We propose that $\mathfrak{F}(t)$ satisfies the stochastic FEBE:

$$\tau^H {}_{-\infty}D_t^H T + T = s\mathfrak{F}. \tag{48}$$

If we identify the deterministic part with the externally forced response and the stochastic part with the internal variability, then the linearity of the FEBE implies that each satisfies the FEBE separately. First, take ensemble averages and use $<f(t)> = <T_i(t)> = 0$:

$$\tau^h({}_{-\infty}D_t^h T_e) + T_e = sF(t);\ F = \langle\mathfrak{F}\rangle;\quad T_e = \langle T\rangle \tag{49}$$

(since $T_i(-\infty) = 0$). Now, subtract this from Equation (48) to yield:

$$\tau^h({}_{-\infty}D_t^h T_i) + T_i = sf(t). \tag{50}$$

Therefore, in the stochastic FEBE, the externally forced and internal variability both simply represent sources or sinks of heat energy that are external to the radiating surface layer. Whether heat comes from internal storage or from outside the system, all that matters is that it appears at the surface where it can participate in the radiative part of the energy balance.

To illustrate how the FEBE models both the deterministic external and stochastic internal variabilities, let us make a simple model that includes them both (see Figures 8 and 9 and details in Appendix). The model has a resolution of the weather–macroweather transition scale (taken as 10 days) and it solves the FEBE over 100,000 days starting in the year 1825, up to the year 2100. The deterministic part of the forcing is taken to be zero up until 1880; this is followed by a linearly increasing forcing – a "ramp" – up until the year 2020 (≈51,100 days) after which the forcing remains constant at that value until the year 2100. The stochastic part of the forcing is due to a Gaussian white noise whose amplitude is such that the RMS monthly anomaly is ±0.14 K (as observed, e.g. Lovejoy, 2017a). The forcing was adjusted so that the forced response was 1 K in 2020 (close to the observed anthropogenic warming); according to this model, there is another 0.24 K of warming "in the pipe". More details including a scale-by-scale statistical analysis and comparison with statistical analyses of global temperature series

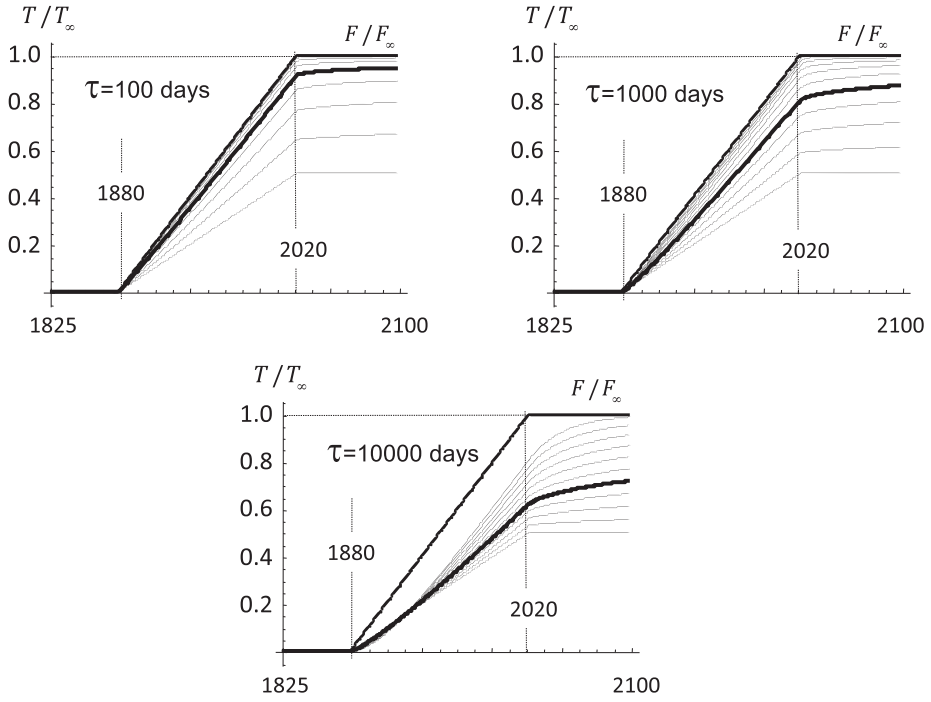


**FIGURE 8** The non-dimensional temperature $T/T_\infty$ response (thin lines) for non-dimensional forcing $F/F_\infty$ (thick line) with unit sensitivity ($s = 1$). Each graph shows the results for $h = 0$ (bottom) to $h = 1$ (top, exponential), at intervals of 0.1; the thick middle curve highlights the most realistic value $h = 0.4$. The plots are shown for the relaxation time $\tau = 10^2, 10^3, 10^4$ days (upper left to bottom as indicated). The resolution of the calculations was taken to be the weather–macroweather transition scale ($\tau_w = 10$ days), but this is only important for the stochastic internal variability; smooth deterministic forcings such as those here are insensitive to $\tau_w$ as long as it is much smaller than $\tau$

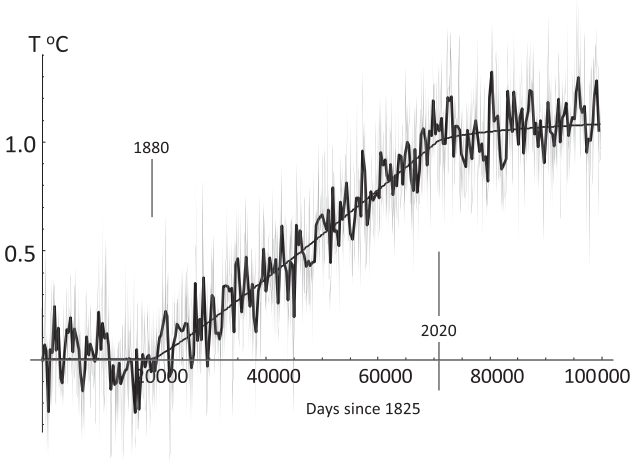


**FIGURE 9** Thin black line: a simulation of the monthly-resolution temperatures with resolution $\tau_w = 10$ days, the sum of the externally forced model with $h = 0.4$, $\tau = 1{,}000$ days (Figure 8, upper right) with the Gaussian white-noise driven stochastic internal variability model averaged over a factor of three to simulate a monthly resolution series. The amplitude of the internal variability was chosen to match observations of the monthly-resolution standard deviation of global temperatures $\langle T_{\tau_w}(t)^2 \rangle^{1/2} = 0.14$ C. Thick black line: at annual resolution obtained by averaging the previous over an additional factor of 12

are given in the Appendix. In place of this "toy model", Procyk *et al.* (2020) uses IPCC AR5 forcings and makes quite realistic hind-projections (1880–present) and projections (through to 2100) that are within the uncertainty limits of the IPCC AR5 (CMIP5) GCMs.

## 5 | CONCLUSIONS

Beyond the deterministic forecasting limit of about 10 days, GCMs are effectively stochastic. In addition, at least with the external forcing scenarios used in the IPCC AR5 – the Representative Carbon Pathways – each GCM response to external forcings is quite linear (Hébert and Lovejoy, 2018); to a good approximation, climate projections over the next century appear to be a linear stochastic problem (Hebert, 2017; Procyk *et al.*, 2020; Hébert *et al.*, 2021). In addition, stochastic, linear monthly, seasonal and interannual scale (macroweather) temperature forecasts have skill comparable to – or better than – those of GCMs (Lovejoy *et al.*, 2015; del Rio Amador and Lovejoy, 2019).

These facts – and the urgency of reducing forecast and (especially) projection uncertainties – militates for the rapid development of stochastic macroweather models constructed directly at macroweather scales. Up until now, such models have been based on the long memory of the Earth's atmosphere and climate that was exploited by using the scaling principle: the fact that over wide ranges of spatial scales, atmospheric dynamics are scaling (see the extensive review: Lovejoy and Schertzer, 2013). Predicting such a long-memory stochastic process requires a long series of historical data, therefore mathematically these methods are effectively "past value" rather than conventional "initial value" type problems (Lovejoy *et al.*, 2015, Del Rio amador and Lovejoy 2021).

At first, scaling was invoked to directly mediate the forcing and the response: a scaling Climate Response

Function (CRF: Rypdal, 2012; Van Hateren, 2013; Rypdal and Rypdal, 2014; Hebert, 2017; Myrvoll-Nilsen *et al.*, 2020; Hébert *et al.*, 2021). Recently, it was suggested by Lovejoy (2019a) that this phenomenological approach could be made more physical (and realistic) by suggesting that rather than applying it to the CRF directly, that the scaling principle should instead be applied to the energy storage mechanisms. It was pointed out that if this was done, then both high- and low-frequency memories and exponents would automatically be explained by a single more-fundamental exponent $h$, the (fractional) order of the resulting Fractional Energy Balance Equation (FEBE).

Since then, another approach starting with energy balance models has reached the same conclusion and has shown how the zero-dimensional (global) models can be generalized to full 2D models for temperature anomalies. Perhaps the most surprising result (Lovejoy, 2021a; 2021b) was that an apparently minor change to the classical Budyko–Sellers Energy Balance Models – the extension from 2D to the full (classical, integer-ordered) 3D continuum heat equation – generically implies that the surface temperature satisfies the half-order EBE (HEBE). Finally, a generalization to the (3D) Fractional Heat Equation was shown to imply that the (2D) surface temperatures satisfy the FEBE.

In this article, we pursued a more phenomenological approach by deriving the FEBE from the scaling principle applied to the energy storage processes. Mathematically, the FEBE is a fractional relaxation equation whose deterministic version has been studied for some while (Miller and Ross, 1993; Podlubny, 1999), and whose stationary stochastic version was studied in Lovejoy (2019b) in order to understand the response to internal forcing. We showed how the FEBE can be solved using Green's functions and we linked the high- and low-frequency limits of the Green's function to fractional Gaussian noise (high frequencies dominated by internal variability) and low frequencies to power-law CRFs for climate projections. The fact that the stochastic FEBE is a good model for the internal variability means that the internal variability can be used to determine the response of the system to external forcing, a goal that has been pursued ever since Leith (1975) proposed using the Fluctuation Dissipation Theorem for this purpose.

We showed how the response to periodic forcing can be conveniently handled by considering complex climate sensitivities, notably to take into account the phase lag between the annual maximum forcing and the temperature response. This opens the possibility of using the annual cycle to estimate the model parameters, exponents, relaxation times and usual (static) climate sensitivities. We showed that although the FEBE obeys Newton's law of cooling, that the heat flux crossing a surface nonetheless depends on the fractional temperature time derivative rather than the usual integer-ordered one (this is an extension of the half-order result that follows quite classically from the continuum mechanics heat equation). We also derived the theoretical FEBE ratio of transient to equilibrium climate sensitivity (TCR/ECS) and showed that it was ≈0.70, that is, within the 90% confidence interval of GCM estimates (IPCC AR5).

Finally, we put both the deterministic (external forcing) and stochastic (internal forcing) together into a simple 140-year ramp forcing experiment for the industrial epoch warming. The model assumed that between 1880 and 2020 there was 1 K of warming and predicted that if the current forcing was held constant from 2020 onwards, that there would be another 0.24 K "in the pipe line". We then statistically evaluated the model using Haar fluctuations and showed that the RMS fluctuations as functions of time-scale were already fairly realistic, but that precise parameter estimates are difficult.

Although in this article we only considered the zero-dimensional FEBE, various extensions to 2D have already been proposed (Lovejoy, 2021b). When forced by stochastic internal and deterministic external forcing, the FEBE thus constitutes a new class of low-frequency atmospheric model that uses historical data and the long memory to make macroweather forecasts and climate projections. There are indications (del Rio Amador and Lovejoy, 2019; Procyk *et al.*, 2020) that these are more skilful and are less uncertain than conventional approaches. The future challenge will be to improve this approach and extend it for predicting and projecting other atmospheric parameters – especially for precipitation.

## ACKNOWLEDGEMENTS

This project received no specific funding, but S. Lovejoy benefitted from a small grant from the National Science and Engineering research Council (Canada). R. Hébert benefited from funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 716092 and no. 772852). We acknowledge discussions with D. Clarke and C. Penland.

## ORCID

*Shaun Lovejoy* https://orcid.org/0000-0002-9367-3137
*Raphael Hébert* https://orcid.org/0000-0002-9869-4658
*Lenin Del Rio Amador* https://orcid.org/0000-0003-4043-472X

## REFERENCES

Babenko, Y.I. (1986) *Heat and Mass Transfer*. Leningrad: Khimiya (in Russian).

Biagini, F., Hu, Y.Z., Øksendal, B. and Zhang, T.S. (2008) *Stochastic Calculus for Fractional Brownian Motion and Applications*. Springer-Verlag, London, UK.

Brunt, D. (1932) Notes on radiation in the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 58(247), 389–420.

Budyko, M.I. (1969) The effect of solar radiation variations on the climate of the Earth. *Tellus*, 21, 611–619.

Cheng, J.-F. and Chu, Y.-M. (2011) Solution to the linear fractional differential equation using Adomian decomposition method. *Mathematical Problems in Engineering*, 2011, 1–14. https://doi.org/10.1155/2011/587068.

Cionni, I., Visconti, G. and Sassi, F. (2004) Fluctuation dissipation theorem in a general circulation model. *Geophysical Research Letters*, 31(9), L09206. https://doi.org/10.1029/2004GL019739.

Coffey, W.T., Kalmykov, Y.P. and Titov, S.V. (2012) *Characteristic times of anomalous diffusion in a potential, in Fractional Dynamics: Recent Advances*, In: Klafter, J., Lim, S. and Metzler, R. (Eds.) pp. 51–76, World Scientific.

Cox, P.M., Huntingford, C. and Williamson, M.S. (2018) Emergent constraint on equilibrium climate sensitivity from global temperature variability. *Nature*, 553(7688), 319–322.

Del Rio Amador, L. and Lovejoy, S. (2019) Predicting the global temperature with the stochastic seasonal to interannual prediction system (StocSIPS). *Climate Dynamics*, 53(7), 4373–4411. https://doi.org/10.1007/s00382-019-04791-4.

Del Rio Amador, L., and Lovejoy, S. (2021) Long-range Forecasting as a Past Value Problem: Untangling Correlations and Causality with scaling. *Geophys. Res. Lett.*, under review.

Dijkstra, H.A. (2013) *Nonlinear Climate Dynamics*. Cambridge University Press, New York, USA.

Donohoe, A., Dawson, E., McMurdie, L., Battisti, D.S. and Rhines, A. (2020) Seasonal asymmetries in the lag between insolation and surface temperature. *Journal of Climate*, 33, 3921–3945. https://doi.org/10.1175/jcli-d-19-0329.1.

Dwyers, H.A. and Petersen, T. (1973) Time-dependent global energy modeling. *Journal of Applied Meteorology and Climatology*, 12, 36–42.

Franzke, C.L.E., Barbosa, S., Blender, R., Fredriksen, H.-.B., Laepple, T., Lambert, F., Nilsen, T., Rypdal, K., Rypdal, M., Scotto, M.G., Vannitsem, S., Watkins, N.W., Yang, L. and Yuan, N. (2020) The structure of climate variability across scales. *Reviews of Geophysics*, 58(2), e2019RG000657. https://doi.org/10.1029/2019rg000657.

Fredriksen, H.B. and Rypdal, M. (2017) Long-range persistence in global surface temperatures explained by linear multibox energy balance models. *Journal of Climate*, 30, 7157–7168. https://doi.org/10.1175/JCLI-D-16-0877.1.

Geoffroy, O., Saint-Martin, D., Olivié, D.J., Voldoire, A., Bellon, G. and Tytéca, S. (2013) Transient climate response in a two-layer energy-balance model. Part I: Analytical solution and parameter calibration using CMIP5 AOGCM experiments. *Journal of Climate*, 26, 1841–1857.

Gottwald, G.A., Wormell, J.P. and Wouters, J. (2016) On spurious detection of linear response and misuse of the fluctuation–dissipation theorem in finite time series. *Physica D*, 331, 89–101. https://doi.org/10.1016/j.physd.2016.05.010.

Hansen, J., Sato, M., Kharecha, P. and von Schuckmann, K. (2011) Earth's energy imbalance and implications. *Atmospheric Chemistry and Physics*, 11, 13421–13449. https://doi.org/10.5194/acp-11-13421-2011.

Hasselmann, K. (1976) Stochastic climate models. *Part I: Theory. Tellus*, 28, 473–485.

Hasselmann, K., Sausen, R., Maier-Reimer, E. and Voss, R. (1993) On the cold start problem in transient simulations with coupled atmosphere–ocean models. *Climate Dynamics*, 9, 53–61.

Hasselmann, K., Hasselmann, S., Giering, R., Ocana, V. and Storch, H.V. (1997) Sensitivity study of optimal $CO_2$ emission paths using a simplified structural integrated assessment model (SIAM). *Climatic Change*, 37, 345–386. https://doi.org/10.1023/A:1005339625015.

Hebert, R. (2017) *A scaling model for the forced climate variability in the anthropocene*. MSc thesis, McGill University, Montreal.

Hébert, R. and Lovejoy, S. (2015) The runaway Green's function effect: interactive comment on "Global warming projections derived from an observation-based minimal model" by K. Rypdal. *Earth System Dynamics Discussion*, 6, C944–C953.

Hébert, R., Casado, M and Laepple, T. (2020) Climate Variability Across Scales (CVAS): Phase Two. *Past Global Changes Magazine*, 28(2).

Hébert, R. and Lovejoy, S. (2018) Regional climate sensitivity- and historical-based projections to 2100. *Geophysical Research Letters*, 45, 4248–4254. https://doi.org/10.1002/2017GL076649.

Hébert, R., Lovejoy, S. and Tremblay, B. (2021) An observation-based scaling model for climate sensitivity estimates and global projections to 2100. *Climate Dynamics*, 56, 1105–1129. https://doi.org/10.1007/s00382-020-05521-x.

Held, I.M., Winton, M., Takahashi, K., Delworth, T., Zeng, F. and Vallis, G.K. (2010) Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *Journal of Climate*, 23, 2418–2427.

Hurst, H.E. (1951) Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 770–799.

Kou, C. and Xie, X.S. (2004) Generalized Langevin equation with fractional Gaussian noise: subdiffusion within a single protein molecule. *Physical Review Letters*, 93(18), 180603. https://doi.org/10.1103/PhysRevLett.93.180603.

Kubo, R. (1966) The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1), 255–284.

Leith, C.E. (1975) Climate response and fluctuation dissipation. *Journal of the Atmospheric Sciences*, 32, 2022–2026.

Li, S. and Jarvis, A. (2009) Long run surface temperature dynamics of an A-OGCM: the HadCM3 4x$CO_2$ forcing experiment revisited. *Climate Dynamics*, 33, 817–825.

Lovejoy, S. (2013) What is climate? *EOS*, 94(1), 1–2.

Lovejoy, S. (2015) Using scaling for macroweather forecasting including the pause. *Geophysical Research Letters*, 42, 7148–7155. https://doi.org/10.1002/2015GL065665.

Lovejoy, S. (2017a) How accurately do we know the temperature of the surface of the Earth? *Climate Dynamics*, 49, 4089–4106. https://doi.org/10.1007/s00382-017-3561-9.

Lovejoy, S. (2017b) How scaling fluctuation analysis transforms our view of the climate. *PAGES Magazine*, 25(3), 136–137. https://doi.org/10.22498/pages.25.3.136.

Lovejoy, S. (2018) Spectra, intermittency and extremes of weather, macroweather and climate. *Nature Scientific Reports*, 8, 1–13, article no. 12697. doi:. https://doi.org/10.1038/s41598-018-30829-4.

Lovejoy, S. (2019a) *Weather, Macroweather and the Climate: Our random yet predictable atmosphere*. Oxford University Press, New York, USA.

Lovejoy, S. (2019b) Fractional relaxation noises, motions and the fractional energy balance equation. *Nonlinear Processes in Geophysics Discussions*, in review. https://doi.org/10.5194/npg-2019-39.

Lovejoy, S. (2021a) The half-order energy balance equation. Part 1: The homogeneous HEBE and long memories. *Earth System Dynamics*, in press. https://doi.org/10.5194/esd-2020-12.

Lovejoy, S. (2021b) The half-order energy balance equation. Part 2: The inhomogeneous HEBE and 2D energy balance models. *Earth System Dynamics*, in press. https://doi.org/10.5194/esd-2020-13.

Lovejoy, S. and Schertzer, D. (2012a) Haar wavelets, fluctuations and structure functions: convenient choices for geophysics. *Nonlinear Processes in Geophysics*, 19, 513–527. https://doi.org/10.5194/npg-19-1-2012.

Lovejoy, S. and Schertzer, D. (2012b) Stochastic and scaling climate sensitivities: solar, volcanic and orbital forcings. *Geophysical Research Letters*, 39(11), L11702. https://doi.org/10.1029/2012GL051871.

Lovejoy, S. and Schertzer, D. (2013) *The Weather and Climate: Emergent laws and multifractal cascades*. Cambridge University Press, Cambridge, UK.

Lovejoy, S., del Rio Amador, L. and Hébert, R. (2017) Harnessing butterflies: theory and practice of the Stochastic Seasonal to Interannual Prediction System (StocSIPS). In: Tsonis, A.A. (Ed.) *Advances in Nonlinear Geosciences*. Springer Nature, Cham, Switzerland, pp. 305–355.

Lovejoy, S., del Rio Amador, L. and Hébert, R. (2015) The ScaLIng Macroweather Model (SLIMM): using scaling to forecast global-scale macroweather from months to decades. *Earth System Dynamics*, 6, 637–658. https://doi.org/10.5194/esd-6-1-2015.

Lucarini, V., Ragone, F. and Lunkeit, F. (2017) Predicting climate change using response theory: global averages and spatial patterns. *Journal of Statistical Physics*, 166, 1036–1064. https://doi.org/10.1007/s10955-016-1506-z.

Mainardi, F. and Pironi, P. (1996) The fractional Langevin equation: Brownian motion revisited. *Extracta Mathematicae*, 10, 140–154.

Mandelbrot, B.B. (1971) A fast fractional Gaussian noise generator. *Water Resources Research*, 7(3), 543–553. https://doi.org/10.1029/WR007i003p00543.

Mandelbrot, B.B. and Van Ness, J.W. (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10, 422–437.

Marshall, J., Scott, J. and Proshutinsky, A. (2017) "Climate response functions" for the Arctic Ocean: a proposed coordinated modelling experiment. *Geoscientific Model Development*, 10, 2833–2848. https://doi.org/10.5194/gmd-10-2833-2017.

Marshall, J., Armour, K.C., Scott, J.R., Kostov, Y., Hausmann, U., Ferreira, D., Shepherd, T.G. and Bitz, C.M. (2014) The ocean's role in polar climate change: asymmetric Arctic and Antarctic responses to greenhouse gas and ozone forcing. *Philosophical Transactions of the Royal Society A*, 372, 20130040. https://doi.org/10.1098/rsta.2013.0040.

Meyer, R.F. (1960) *A heat-flux-meter for use with thin film surface thermometers*. Aeronautical Report LR 279. Ottawa: National Research Council of Canada.

Miller, K.S. and Ross, B. (1993) *An Introduction to the Fractional Calculus and Fractional Differential Equations*. John Wiley and Sons, New York.

Molz, F.J., Liu, H.H. and Szulga, J. (1997) Fractional Brownian motion and fractional Gaussian noise in subsurface hydrology: a review, presentation of fundamental properties, and extensions. *Water Resources Research*, 33, 2273–2286.

Myrvoll-Nilsen, E., Sørbye, S.H., Fredriksen, H.-B., Rue, H. and Rypdal, M. (2020) Statistical estimation of global surface temperature response to forcing under the assumption of temporal scaling. *Earth System Dynamics*, 11, 329–345. https://doi.org/10.5194/esd-11-329-2020.

North, G.R., Cahalan, R.F. and Coakley, J.A., Jr. (1981) Energy balance climate models. *Reviews of Geophysics*, 19, 91–121.

North, G.R., Bell, R.E. and Hardin, J.W. (1993) Fluctuation dissipation in a general circulation model. *Climate Dynamics*, 8, 259–264.

North, G.R., Wang, J. and Genton, M.G. (2011) Correlation models for temperature fields. *Journal of Climate*, 24, 5850–5862. https://doi.org/10.1175/2011JCLI4199.1.

North, G.R. and Kim, K.-Y. (2017) *Energy Balance Climate Models*. Wiley-VCH, Weinheim, Germany.

Oldham, K.B. (1973) Diffusive transport to planar, cylindrical and spherical electrodes. *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, 41, 351–358.

Oldham, K.B. and Spanier, J. (1972) A general solution of the diffusion equation for semi infinite geometries. *Journal of Mathematical Analysis and Applications*, 39, 655–669.

Oldham, K.B. and Spanier, J. (1974) *The fractional calculus theory and applications of differentiation and integration to arbitrary order*. Academic Press, reprinted by Dover, New York, USA 2006.

Podlubny, I. (1999) *Fractional Differential Equations*. Academic Press.

Procyk, R., Lovejoy, S. and Hébert, R. (2020) The fractional energy balance equation for climate projections through 2100. *Earth System Dynamics, Discussion, under review.* https://doi.org/10.5194/esd-2020-48.

Rypdal, K. (2012) Global temperature response to radiative forcing: solar cycle versus volcanic eruptions. *Journal of Geophysical Research*, 117(D6), D06115. https://doi.org/10.1029/2011JD017283.

Rypdal, K. (2016) Global warming projections derived from an observation-based minimal model. *Earth System Dynamics*, 7, 51–70.

Rypdal, K., Rypdal, M. and Fredriksen, H.B. (2015) Spatiotemporal long-range persistence in Earth's temperature field: analysis of stochastic-diffusive energy balance models. *Journal of Climate*, 28, 8379–8395. https://doi.org/10.1175/JCLI-D-15-0183.1.

Rypdal, M. and Rypdal, K. (2014) Long-memory effects in linear response models of Earth's temperature and implications for future global warming. *Journal of Climate*, 27(14), 5240–5258. https://doi.org/10.1175/JCLI-D-13-00296.1.

Schertzer, D. and Lovejoy, S. (1987) Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *Journal of Geophysical Research*, 92(D8), 9693–9714.

Schwartz, S.E. (2007) Heat capacity, time constant, and sensitivity of Earth's climate system. *Journal of Geophysical Research*, 112(D24), D24S05. https://doi.org/10.1029/2007JD008746.

Schwartz, S.E. (2012) Determination of Earth's transient and equilibrium climate sensitivities from observations over the twentieth century: strong dependence on assumed forcing. *Surveys in Geophysics*, 33, 745–777.

Sellers, W.D. (1969) A global climatic model based on the energy balance of the Earth–atmosphere system. *Journal of Applied Meteorology and Climatology*, 8, 392–400.

Sierociuk, D., Dzieliński, A., Sarwas, G., Petras, I., Podlubny, I. and Skovranek, T. (2013) Modelling heat transfer in heterogeneous media using fractional calculus. *Philosophical Transactions of the Royal Society A*, 371, 20120146. https://doi.org/10.1098/rsta.2012.0146.

Sierociuk, D., Skovranek, T., Macias, M., Podlubny, I., Petras, I., Dzielinski, A. and Ziubinski, P. (2015) Diffusion process modeling by using fractional-order models. *Applied Mathematics and Computation*, 257, 2–11. https://doi.org/10.1016/j.amc.2014.11.028.

Stubenrauch, C.J., Chédin, A., Rädel, G., Scott, N.A. and Serrar, S. (2006) Cloud properties and their seasonal and diurnal variability from TOVS Path-B. *Journal of Climate*, 19, 5531–5553. https://doi.org/10.1175/JCLI3929.1.

Van Hateren, J.H. (2013) A fractal climate response function can simulate global average temperature trends of the modern era and the past millennium. *Climate Dynamics*, 40, 2651–2670. https://doi.org/10.1007/s00382-012-1375-3.

Watkins, N.W., Chapman, S.C., Chechkin, A., Ford, I., Klages, R. and Stainforth, D.A. (2020) On generalized Langevin dynamics and the modelling of global mean temperature. *arXiv*, arXiv:2007.06464v1.

West, B.J., Bologna, M. and Grigolini, P. (2003) *Physics of Fractal Operators*. Springer, New York, USA.

Yoshimori, M., Watanabe, M., Shiogama, H., Oka, A., Abe-Ouchi, A., Ohgaito, R. and Kamae, Y. (2016) A review of progress towards understanding the transient global mean surface temperature response to radiative perturbation. *Progress in Earth and Planetary Science*, 3, 21. https://doi.org/10.1186/s40645-016-0096-3.

Zeng, X. and Geil, K. (2016) Global warming projection in the 21st century based on an observational data-driven model. *Geophysical Research Letters*, 43, 10947–10954. https://doi.org/10.1002/2016GL071035.

**How to cite this article:** Lovejoy S, Procyk R, Hébert R, Del Rio Amador L. The fractional energy balance equation. *Q J R Meteorol Soc*. 2021;1–25. https://doi.org/10.1002/qj.4005

## APPENDIX A. EMPIRICAL VALIDATION OF THE RAMP MODEL

To illustrate how the FEBE models both the deterministic external and stochastic internal variability, in Section 4.5 we presented a simple model that includes them both (Figures 8 and 9). In this Appendix we give a few technical details and a statistical validation of the model using global temperature series.

First, consider the externally forced variability using a linearly increasing (ramp) model of the industrial epoch anthropogenic warming. Figure 8 shows the forcing model (thick black) non-dimensionalized with the long-time forcing $F_\infty$ with the response non-dimensionalized by the long-time (equilibrium) temperature $T_\infty$; since $s = T_\infty/F_\infty$, the curves are independent of $s$. The model has a resolution of the weather–macroweather transition scale (taken as 10 days) and starting in the year 1825, it solves the FEBE up to 100,000 days (to the year 2100). The forcing is taken to be zero up until 1880, followed by a linearly increasing forcing – a "ramp" – up until the year 2020 (≈51,100 days) after which the forcing remains constant at that value until the year 2100.

In the historical part, the forcing is indeed roughly of this form, the amplitude estimated for 2020 at close to 2.4 $W{\cdot}m^{-2}$ ($CO_{2eq}$ forcing, since the beginning of the industrial epoch). The resolution of the simulation was the typical weather–macroweather transition time $\tau_w \approx 10$ days, but as long as $\tau \gg \tau_w$ the results for this smooth, deterministic simulation will be insensitive to this. However, as explained in Section 4.2 (Equation (41), Table 1), the resolution is fundamentally important for stochastic, noise-driven simulations of internal variability. The future part of the model corresponds to a complete halt to all emissions (and other forcing) in 2020.

Before continuing, it is worth making a quick comment about the numerical techniques that were used for the simulations. Numerical convolution algorithms are very efficient, so that we used the FEBE solution in convolution form. Since the FEBE Green's functions have power-law behaviour at both high and low frequencies, care is required at small and large scales. To avoid high-frequency problems (especially sensitive since below, the FEBE is stochastically forced with a Gaussian white noise), we therefore used the step response $G_{1,h}$ Green's function that has a smoother small $t$ behaviour ($t^h$) rather than the impulse response $t^{-1+h}$ behaviour. After the convolution, the resulting series was then numerically differentiated to obtain the solution. The low frequencies also have some issues because of the long memory: the storage. For simulations of pure fGn this is a nontrivial issue. Here however, for scales $t \gg \tau$, the Green's function that determines the memory – although still power law (Equation 31, $t^{-1-h}$) – is much shorter than for pure fGn ($t^{-1+h}$) so that it suffices to make simulations of total length a few multiples of $\tau$. More details are given in Lovejoy (2019b).

Using these techniques, three plots are shown in Figure 8, one for each of the relaxation scales $\tau$ = 100, 1,000, 10,000 simulated days (upper left, upper right, bottom). The upper-right simulation with $\tau$ = 1,000 days (≈3 years) has a value close to the various relaxation-scale estimates and the curve corresponding to the realistic parameter $h$ = 0.4 (red) is fairly close to the actual

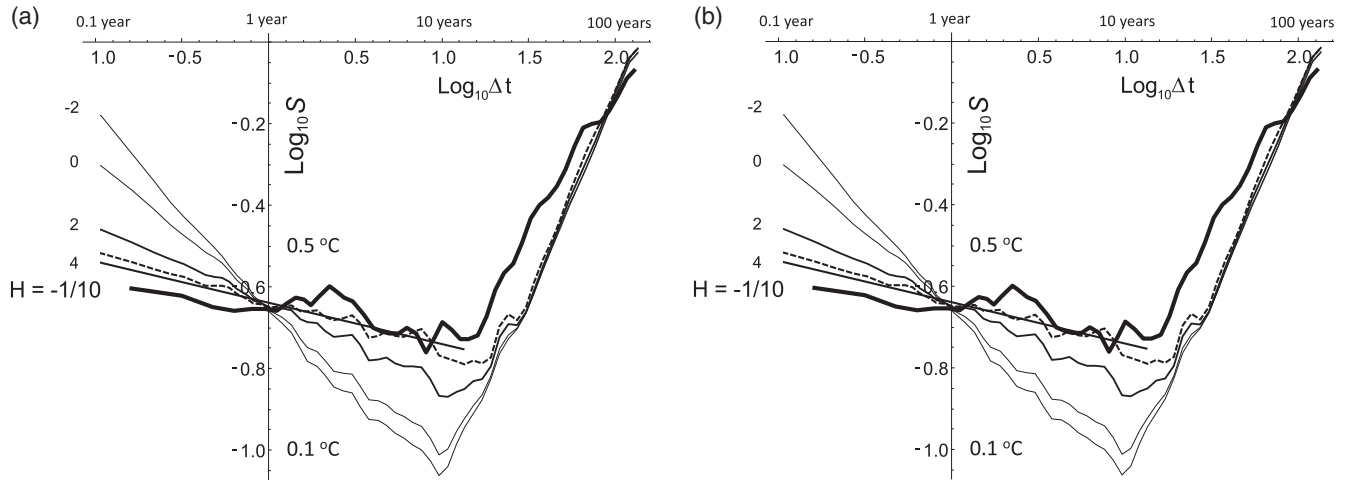


**FIGURE A1** (a) The average Haar fluctuation structure function $S(\Delta t) = \langle \Delta T(\Delta t)^2 \rangle^{1/2}$ for six globally averaged monthly-resolution temperature time series (thick) compared with ramp-model simulations with $h = 4/10$ ($H = h - 1/2 = -1/10$) analysed over the simulated period 1880–2012 (these are the series analysed in Lovejoy (2017a)). The simulations have been calibrated so that their RMS variability at 1 year is the same as the data. Simulation results are shown for relaxation times varying from $10^{-2}$ days (top left) to $10^4$ days (dashed, bottom left) each labelled at left by $\log_{10}\tau$. The reference (thick) line has the theoretical slope ($H$) for fGn, that is, infinite relaxation time. (b) The same as (a) except for $h = 1/2$, hence $H = 0$ (thick, flat reference line)

reconstructed warming over the period. These curves illustrate what might happen if emissions suddenly stopped; the parts beyond 2020 show the thermal "inertia" in the system. We see that with only some exceptions at the very long $\tau$ = 10,000 (≈30 years, bottom) that the response is very linear over the ramp region, so that the response at the end of the 140-year ramp is a little larger than the TCR (defined after a 70-year ramp, see Figures 6 and 7; the TCR/ECS contours in the figure are fairly horizontal, not very sensitive to $\Delta t/\tau$).

The $h = 0.4$ curve on the graph with the most realistic $\tau$ (=1,000 days, upper right) shows that $T(2020)/T(\infty) = 0.81$ (the TCR/ECS is ≈0.7, see Figures 6 and 7). If we take the current increase in temperature since pre-industrial epoch to be $T(2020) = 1$ K, then the equilibrium temperature $T(\infty)$ = ECS = 1/0.81 = 1.24 K. If we take the corresponding forcing since the pre-industrial epoch to be 2.4 $W{\cdot}m^{-2}$, then $s$ = 1.24/2.4 = 0.52 $K{\cdot}W^{-1}{\cdot}m^{-2}$; using the value 3.71 $W{\cdot}m^{-2}{\cdot}CO_2{}^{-1}$ doubling, we obtain $s = 1.92\ K{\cdot}CO_2{}^{-1}$ doubling. Comparing this with the IPCC AR5 range of 1.5–4.5 $K{\cdot}CO_2{}^{-1}$ doubling (90% confidence), we see that it is about 4% below the mean and easily within the range; it is close to the Procyk *et al.* (2020) estimate $2.0 \pm 0.4$ $K{\cdot}W^{-1}{\cdot}m^{-2}$. This is surprisingly good confirmation for a model based only on the scaling storage assumption.

We can now combine the externally forced part of the model (Figure 8) with a simulation of the monthly-resolution internally forced temperatures: the sum of using the $h$ = 0.4, $\tau$ = 1,000 day model (middle upper left) averaged over a factor of three to simulate a monthly resolution series. Figure 9 shows the result at monthly resolution (thin) and when this is averaged further to annual resolution (thick); the amplitude of the noise contribution was chosen to equal ±0.14 K at monthly resolution, close to the real global statistics. Due to the (singular) small-scale fGn behaviour, at annual resolution, the amplitude of the noise is thus reduced by the factor $12^H = 12^{-0.1} = 0.78$ (Equation (38), $H = h - 1/2$); that is, the annual anomaly fluctuation amplitude is $0.11 = 12^H$ 0.14 K.

In spite of its simplicity, Figure 9 shows that the model is apparently quite realistic, yet proper statistical analysis is needed to confirm this. A convenient way to compare the model statistics with those of globally averaged temperature series is to estimate their root-mean-square Haar fluctuations. For a series $T(t)$, the Haar fluctuation $\Delta T(\Delta t)$ over an interval $\Delta t$ is the difference between the average of the first and second halves of the interval. Figure A1a,b shows the root-mean-square (RMS) Haar fluctuations $S(\Delta t) = \langle \Delta T(\Delta t)^2 \rangle^{1/2}$ ("< >" indicates averaging over all the available series and disjoint intervals) for six globally averaged series from 1880 to 2012 (thick brown) compared to simulated series with various relaxation times $\tau$ ranging from $10^{-2}$ (top) to $10^4$ days (bottom) for $h$ = 4/10 (Figure A1a) and $h$ = 1/2 (Figure A1b, the empirical curve is repeated on each). As mentioned, the simulations were "calibrated" by adjusting the amplitude of the Gaussian noise part of the forcing so that the amplitude of the model and empirical fluctuations were equal

at annual scales. The monthly series are identical to those analysed in depth in Lovejoy (2017a); full details and references are given there.

The Haar fluctuations combine averaging and differencing so that their interpretation is straightforward. When the fluctuations increase with scale (i.e. regimes where the slopes in the plot are positive), and the raw fluctuations are multiplied by a factor 2 (as they are here, this is the canonical factor), the differencing dominates and the values are close to the RMS differences that quantify the typical change in temperature over a given time interval. For example, the figure shows that at centennial scales (far right), the typical fluctuations are of the order of 1 K. In scale ranges over which the RMS fluctuations decrease, the averaging dominates and the values are close to the anomaly fluctuations, that is, the temporal mean of the series over $\Delta t$ after its overall mean has been removed.

Over the range of 2 months (the smallest $\Delta t$) to about 10 years, their approximate slope (indicated in by the straight reference lines in Figure A1a,b) is about $H = -0.1$. In this decreasing macroweather regime, successive fluctuations tend to cancel each other out; the series appears to be stable. For $\Delta t \approx >10$ years, the fluctuations start increasing: the climate regime. Ten years is the typical scale at which the response to anthropogenic forcing exceeds typical responses to internal forcing; at these scales, fluctuations no longer tend to cancel each other out, the series appears to be unstable.

According to the high-frequency theory we developed in Section 4.2, for time-scales much smaller than the relaxation scale $\tau$, the stochastic part of the response is an fGn with RMS exponent $H = h-1/2$, so that when $\Delta t \ll \tau$ is large enough, $S(\Delta t) \approx \Delta t^H$ . Figure A1a shows the simulation results for the empirically estimated $h = 4/10$, and the theoretical (large $\tau$) $H = -1/10$ reference line is shown in red. Even when $\tau = 10^4$ days (≈30 years, dashed) the small $\Delta t$ slope of the simulation $S(\Delta t)$ is still a little high, although overall the fit is not bad. When we compare the simulation structure functions for decreasing $\tau$, we note a surprising feature: the $\log_{10} S(\Delta t)$ versus $\log_{10}\Delta t$ plots are remarkably linear over several orders of magnitude but with (absolute) slopes that systematically increase as $\tau$ decreases. Without the deterministic ramp that dominates the lower frequencies, this "pseudo-scaling" was investigated further in L3. Rather than being a true scaling regime, it is in reality a very slow transitional regime from high-frequency $\Delta t^H$ to $\Delta t^{-1/2}$ ($H$ <0, independent of $h$, $H$). Figure A1b shows the same simulations but with $h = 1/2$, $H = 0$. In this case, the effect is even more pronounced with the data being more compatible with $\tau \approx 10^3$ days (3 years). From these simulations, we see that when only a factor of 100 or 1,000 in scale is available for analysis, it may be difficult to accurately measure $h$ and $\tau$.