JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, D07111, doi:10.1029/2008JD010651, 2009

Vertical cascade structure of the atmosphere and multifractal dropsonde outages

S. Lovejoy,[1] A. F. Tuck,[2] S. J. Hovde,[2] and D. Schertzer[3,4]

Received 23 June 2008; revised 7 October 2008; accepted 12 December 2008; published 15 April 2009.

[1] We use 220 atmospheric profiles from state-of-the-art dropsondes to test the predictions of multiplicative cascade models of the atmosphere on the horizontal velocity, pressure, temperature, log potential temperature, log equivalent potential temperature, air density, humidity, and vertical sonde velocity. We found that the predictions were accurately verified (to within ±1 to ±2% over 10 m to 1 km for the statistical moments up to second order); the effective outer cascade scale L_{eff} was in the range 1–30 km. In order to perform the analyses and to correctly interpret the results, we needed to overcome technical difficulties caused by the sonde's highly intermittent sampling. This intermittency is the result of both data outages and variable sonde fall speeds; we (surprisingly) found that the outages also had a cascade structure. The wide-range scaling of the sampling rate implies a variable sonde resolution, so that interpolation onto regular grids should generally be avoided (e.g., it would give rise to serious artifacts in estimating the corresponding spectra). In earlier studies, before the cascade nature of the outages was understood, interpolation was avoided by studying the fluctuations using all the pairs of measurement points; this was adequate for fluctuation scaling exponents in the range $0 \leq H \leq 1$. However, determining the cascade structure involves systematically degrading the resolution of fluxes (not fluctuations) so that the variable resolution and their attendant biases could not be avoided. We therefore developed a new method of estimating the fluxes and theoretically determined the corrections necessary to estimate the unbiased exponents. The resulting sonde cascade picture was given further support by (much more straightforward) analysis of uniformly sampled vertical cross sections of the atmosphere obtained from airborne lidar. Using the turbulent fluxes obtained from these various sources, we determined the corresponding cascade regimes and the corresponding exponents as well as the small deviations from the theoretical behavior. In addition to the fluxes, we also studied the fluctuations. To do this we generalized the data point pair method (restricted to nonconservation parameters $0 \leq H \leq 1$) to data triplets (extending the method to $0 \leq H \leq 2$). The resulting fluctuations were analyzed using (generalized) structure functions. We found that while the scaling of the fluxes often broke down at scales greater than about 1 km, the scaling of the fluctuations extended over the entire range 10 m to 10 km.

Citation: Lovejoy, S., A. F. Tuck, S. J. Hovde, and D. Schertzer (2009), Vertical cascade structure of the atmosphere and multifractal dropsonde outages, *J. Geophys. Res.*, *114*, D07111, doi:10.1029/2008JD010651.

1. Introduction

[2] Over the past 25 years, technological advances have produced unprecedented quantities of high-quality atmospheric data spanning large ranges of space-time scales. In the horizontal, satellite data routinely span 4 orders of magnitude in scale (20,000 km/1 km), whereas in the vertical, lidars span over three (10 km/3 m). Similarly, campaigns using in situ aircraft data readily cover up to 4 orders of magnitude in the horizontal (7000 km/100 m) while dropsondes span over 3 in the vertical (12 km/5 m). Both remote sensing and in situ measurements have advantages and disadvantages: remote sensing determined radiances are only nontrivially related to the more physically significant dynamic and thermodynamic variables, whereas in situ measurements have nontrivial (and nonclassical) biases due to long-range correlations between the measuring device and the phenomena measured [*Lilley et al.*, 2008; *Lovejoy et al.*, 2004, 2009a], or due to the extreme intermittency of data outages (dropsondes; see below).

[1]Department of Physics, McGill University, Montreal, Quebec, Canada.
[2]Chemical Sciences Division, NOAA Earth System Research Laboratory, Boulder, Colorado, USA.
[3]CEREVE, Université Paris Est, Marne-la-Vallee, France.
[4]Météo France, Paris, France.

0148-0227/09/2008JD010651$09.00

[3] In the last 5 years, these state of the art data combined with modern scale by scale data analysis techniques have been exploited to firmly establish an atmospheric paradigm first proposed 25 years ago: that the statistical properties are scaling in both horizontal and vertical directions but, owing to scaling stratification, they have systematically different exponents. For example, the scaling of the radiances (visible, IR, passive microwave) holds to within about ±1% from planetary scales down to at least 8 km [*Lovejoy et al.*, 2008a, 2009a] while in the vertical it holds with comparable accuracy for the lidar backscatter, but also for the vertical scaling of the horizontal wind from 10 km down to at least 5 m (*Lilley et al.* [2008, 2004], *Lovejoy et al.* [2007, 2008c], and below). This scaling paradigm can be regarded as a generalization of the classical turbulence laws of Kolmogorov, Bolgiano, Corrsin and Obukhov. Whereas the classical laws were homogeneous and isotropic, the new paradigm shows that they are highly intermittent (multifractal) and anisotropic; in addition, they can be generalized to account for turbulence generation of waves; for a systematic overview, see the series [*Lovejoy et al.*, 2008b; *Lilley et al.*, 2008; *Radkevitch et al.*, 2008], the recent book [*Tuck*, 2008] and the review (S. Lovejoy and D. Schertzer, Towards a new synthesis for atmospheric dynamics: space-time cascades, submitted to *Atmospheric Research*, 2009).

[4] The finding of wide range atmospheric scaling, while significant in itself, is only a rather general consequence of the underlying scaling dynamics, which are a priori compatible with many different dynamical models. However, in the atmosphere these wide range anisotropic scaling laws arise because the dynamics repeat scale after scale in a cascade-like manner. Although cascades are (still) often invoked in a vague sense, starting in the 1960s, precise multiplicative cascade models have been developed as phenomenological models of fluid turbulence [*Novikov and Stewart*, 1964; *Yaglom*, 1966; *Mandelbrot*, 1974]. In the 1980s it became clear that multiplicative cascades were extremely general, being the generic multifractal process. If the dynamics are indeed dominated by this mechanism then one obtains the rather precise prediction that the turbulent fluxes (φ) satisfy the generic multiscaling/multifractal relation

$$M_q = \left(\frac{\lambda}{\lambda_{eff}}\right)^{K(q)}; \quad \lambda = L_{ref}/L; \quad \lambda_{eff} = L_{ref}/L_{eff}, \qquad (1)$$

where $M_q = \langle \varphi_\lambda^q \rangle / \langle \varphi_1 \rangle^q$ is the normalized (and nondimensionalized) qth moment, L_{eff} is the effective outer scale of the cascade and L is the resolution at which it is measured/averaged and "$\langle\ \rangle$" means statistical (ensemble) averaging, here over all our empirically available realizations. Here $\langle \varphi_1 \rangle$ is the ensemble large-scale mean (i.e., the climatological value). L_{ref} is a reference scale; in the horizontal it can conveniently be taken the largest great circle distance on the Earth (=20,000 km) whereas in the vertical (below) it can be taken as 10 km (roughly the thickness of the troposphere). The scale ratio λ_{eff} is determined empirically.

[5] It is important to realize that multiplicative cascade models and their prediction (equation (1)) follow from rather general considerations: essentially (1) that the over a wide range basic dynamical mechanism has no characteristic scale, (2) that there exists a scale by scale conserved flux, and (3) that the dynamics mostly couple structures which are not too different in scale. The first property is much more general than the usual "inertial range" (i.e., a range without sources or sinks of flux); it can apply when the sources and sinks are themselves scaling (this is presumably the relevant case in the atmosphere since the energy-containing short and long wave radiances are scaling as mentioned above). Furthermore, the scaling need not be isotropic. This is important if only because the vertical stratification of the atmosphere prevents it from obeying any wide range isotropic scaling laws. In contrast, if the scaling is anisotropic it can hold over huge ranges. In addition, the usual approaches identify significant fluxes a priori as informed by isotropic theories. The usual choices are: the energy flux, the enstrophy flux and the pseudopotential enstrophy flux; however the resulting predictions (typically of the value of spectral exponents) are often not obeyed. However, it turns out that even without knowing the physical nature of the anisotropic conserved flux φ we can still determine its statistics and test the cascade prediction (equation (1))! We may note that these cascades are highly intermittent in space and time; they involve only ensemble scale by scale conservation. Finally, cascades do not preclude "backscatter" of fluxes from small to large scales.

[6] The significance of these generalizations of the usual cascade picture has recently been underlined by the direct verification of equation (1) on numerical models and reanalyses of the atmosphere (J. Stolle et al., The stochastic cascade structure of deterministic numerical models of the atmosphere, submitted to *Nonlinear Processes in Geophysics*, 2009), and this to very high precision (typically with deviations $< \pm 1\%$) over virtually the entire range of simulated scales. In contrast, early classical scaling studies on essentially the same reanalyses [*Strauss and Ditlevsen*, 1999], found poor scaling with exponents and behavior different from those predicted by isotropic theories.

[7] Although all the mainstream turbulence theories are isotropic (or at least have the same exponents in the horizontal and vertical directions) it is striking that on the contrary all the mainstream empirical vertical spectra are anisotropic: experimentalists invariably find vertical exponents which are different from those in the horizontal. As reviewed in detail by *Lovejoy et al.* [2008b] and *Lilley et al.* [2008], for the horizontal wind, they almost unanimously favor roughly $k^{-5/3}$ spectra in the horizontal and k^{-3} spectra in the vertical (k is a wave number [see, e.g., *Gardner*, 1994; *Gardner et al.*, 1995, 1993; *Dewan*, 1997; *Dewan and Good*, 1986]). It is significant that the postulated vertical k^{-3} spectra are justified by dimensional analysis on the Brunt-Väissälä frequency. This is hardly satisfactory since the latter is not a turbulent flux: even its square is not a positive definite quantity!

[8] One of the reasons that the vertical exponent 3 seems at all plausible is that most of the corresponding studies were based on low ($\approx$100–150 m) resolution radiosonde data (the main exceptions were those of *Adelfang* [1971], *Van Zandt* [1982], *Lazarev et al.* [1994], and *Schertzer and Lovejoy* [1985], which already obtained exponents close to the more accurate value 2.4 [*Lovejoy et al.*, 2007, 2009b]). Indeed, while the advent of high vertical resolution drop-

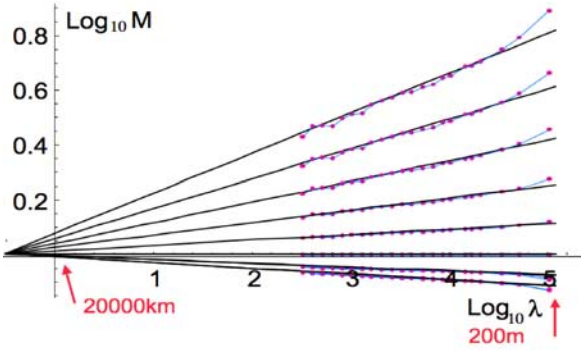

Figure 1a. Horizontal analysis of the moments of the normalized lidar backscatter ratio for 10 atmospheric vertical cross-sections (L_{ref} = 20,000 km corresponding to λ = 1). The curves are for the moments of order q = 0.2, 0.4, . . . 2. The largest directly accessible scale is ≈100 km, and the lines converge to an effective outer scale of $L_{eff} \approx$ 25,000 km.

sondes has demonstrated the existence of hitherto unimagined small-scale variability (e.g., "sheets" [*Dalaudier et al.*, 1994; *Muschinski and Wode*, 1998]), for the moment, it has not resulted in a scientific consensus about the statistical structure of the atmosphere. The purpose of this paper is to exploit the dropsonde revolution (10 to 30 times the resolution of radiosondes), to finally come to grips with the vertical stratification. We will see that a byproduct of this study is the discovery that sonde outages are a more serious problem than hitherto recognized (this may explain why they have not been exploited for this purpose earlier on), and which requires new data analysis techniques to handle. By combining these drop analyses with those of lidar backscatter from aerosols, this paper effectively extends the above cited horizontal scaling/cascade analyses to the vertical.

[9] The paper is structured as follows. Section 2 performs a direct and relatively unproblematic comparison of horizontal and vertical cascade structures using high-resolution lidar data, and section 3 describes the dropsonde data set and the outage problem. Section 4 discusses new data analysis techniques needed to overcome the outage problem and applies them to the data. In section 5 we conclude.

2. An Intercomparison of Horizontal and Vertical Atmospheric Cascade Structures Using Lidar Backscatter

[10] We start our investigation of the vertical cascade structure by exploiting a unique data set from an airborne lidar obtained courtesy of K. Strawbridge (Environment Canada). The data were taken over three afternoons in August 2002 near Vancouver, British Columbia (see *Radkevitch et al.* [2007] for more information on the lidar). The lidar backscatter is primarily from aerosols; *Lilley et al.* [2004] compared the first-order horizontal and vertical structure functions, and *Radkevitch et al.* [2007, 2008] studied the corresponding spectra, including a new anisotropic scaling analysis technique (ASAT) involving nonlinear coordinate transformations. *Lilley et al.* [2008] provided a literature review and additional anisotropy analyses including of the fluxes estimated from gradients of the backscatter. The conclusions were broadly that the backscatter statistics can be accurately described if the ratio of horizontal to vertical scaling exponents was $H_z \approx 0.55$. In addition, the scale at which horizontal and vertical fluctuations are of equal magnitude ("the sphero-scale") was directly estimated for the first time (it varied between about 10 and 80 cm). H_z was close to the value 5/9 which is theoretically predicted assuming that the horizontal statistics of the wind are dominated by energy fluxes (so that the Kolmogorov law holds in the horizontal) and that the vertical statistics are dominated by buoyancy variances fluxes (so that the Bolgiano-Obukhov law holds in the vertical [*Schertzer and Lovejoy*, 1985]).

[11] However a direct intercomparison of the normalized fluxes M_q (equation (1)) was not given, the horizontal and vertical external cascade scales were not estimated. Figures 1a and 1b show the normalized moments of the horizontal and vertical fluxes respectively. We see that the cascade structure predicted by equation (1) is well respected: not only are the lines quite straight, they also "point" to the effective outer scale of the process, i.e., the scale at which a multiplicative cascade would have to start in order to account for the statistics over the observed range. We see that in both cases L_{eff} is a little larger than the physical scales (≈25,000 km and 50 km for the horizontal and vertical, respectively) indicating that even at the largest scales there is residual variability, presumably due to the nonlinear interaction of the observed flux with other atmospheric fluxes at the largest scales. Table 1 shows some of the parameters characterizing $K(q)$ and shows that they are indeed quite different for the horizontal and vertical directions. The parameter $C_1 = K'(1)$ characterizes the intermittency of the mean. In addition, we give the multifractal index α from the "universal" multifractal form

$$K(q) = \frac{C_1}{\alpha - 1}(q^{\alpha} - q); \quad q \geq 0; \quad 0 \leq \alpha \leq 2. \qquad (2)$$

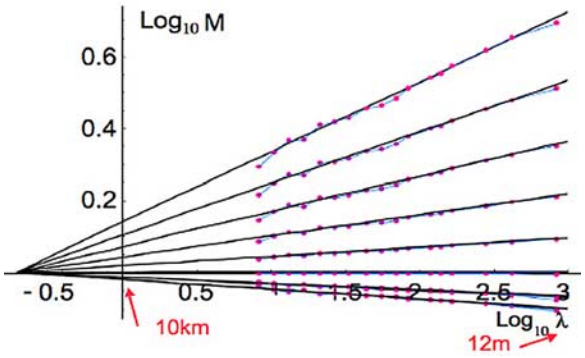

Figure 1b. The same cross sections as in Figure 1a but analyzed in the vertical direction, (L_{ref} = 10 km corresponding to λ = 1). The largest directly accessible scale is ≈3 km and the point of convergence is $L_{eff} \approx$ 50 km; see Table 1. Note that the vertical axis is not the same as for the horizontal analysis: the vertical slopes are different; this is a consequence of the scaling anisotropy; the exponents are roughly in a constant ratio.

Table 1. Ten Vertical Lidar Cross Sections Backscatter Ratio[a]

Field	Resolution (m)	$\overline{\delta}$ (%)	C_1	H	α	$Leff$ (km)
B vertical	12 × 192	0.4	0.11	0.60	1.82	50
B horizontal	12 × 192	0.5	0.076	0.33	1.83	25,000

[a]This shows the cascade parameters estimated here. H and α were estimated by *Lilley et al.* [2008, 2004].

To quantify the accuracy, we characterize the deviations by the mean absolute residuals for the statistical moments M_q of order q from 0 to 2 for all points between the scale of the grid and the overall scale (size) of the data set (about 100 times larger),

$$\Delta = \overline{\left|\log_{10}\left(M_q\right) - K(q)\log_{10}\left(\lambda/\lambda_{eff}\right)\right|}. \quad (3)$$

To convert Δ to a percent deviation, use $\delta = 100(10^{\Delta}-1)$, which we find is generally less than ±0.5% (see Table 1 for the mean $\overline{\delta}$).

[12] Before continuing, we should explain how the fluxes in Figure 1 were estimated. In a scaling regime, the fluctuations over a distance Δz in an observed quantity Δv (consider for example the horizontal component of the wind) are related to the fluxes φ by the following general type of relation:

$$\Delta v(\Delta z) = \varphi_{\Delta z}\Delta z^{H}. \quad (4)$$

If we take Δv to be the (absolute) fluctuation in the turbulent velocity (see below) and $H = 1/3$, $\varphi_{\Delta z} = \varepsilon_{\Delta z}^{1/3}$, where $\varepsilon_{\Delta z}$ is the energy flux at resolution Δz, then equation (4) is the real space expression of the Kolmogorov law, which is the prototypical fluctuation/flux relation. We therefore see that if we take $\Delta z = \eta$ to be the smallest available lag in the scaling regime at scale ratio $\Lambda = L/\eta$, then the flux is $\varphi_\eta = \Delta v(\eta)/\eta^{H}$ and the normalized flux is: $\varphi_\eta/\langle\varphi_\eta\rangle = \Delta v(\eta)/\langle\Delta v(\eta)\rangle$. In other words, even without knowing a priori the physical nature of φ, nor the value of the exponent H, we can nevertheless estimate the normalized flux at resolution η. In order to estimate the flux at a lower resolution $l > \eta$ we simply average it over a set B_l of scale l (for example an $l \times l$ square in 2D, or as below, $l \times \eta$ thin rectangles with $l >> \eta$),

$$\varphi_\lambda = \frac{1}{volB_l}\int_{B_l}\varphi_\Lambda d\underline{x}; \quad volB_l = \int_{B_l} d\underline{x}; \quad \lambda = L/l; \quad \Lambda = L/\eta, \quad (5)$$

where we have indexed the scale of the flux by the scale ratios λ and Λ rather than the sizes l and η. Note that there is no unique definition of the fluctuations Δv: they can be taken as absolute wavelet coefficients, differences or can be estimated by other means (see section 4.4).

3. Dropsondes

3.1. Description of the Data Set

[13] The lidar data analyzed in section 2 are relatively straightforward to analyze, being uniformly spaced in orthogonal directions with high signal-to-noise ratios. However, if we seek to study the usual dynamic or thermodynamic variables, we are forced to turn to in situ measurements. Traditionally over a substantial part of the troposphere, radiosondes have been the only way to get vertical information. However, they have numerous problems including payloads swinging into and out of the balloon's wake, low vertical resolutions (typically of the order of 100 m) and slow ascent speeds which, in areas of strong downdrafts, can even temporarily become descents. As mentioned earlier, these technical difficulties have contributed to the absence of consensus on the nature of the vertical stratification.

[14] In the last ten years, the development of GPS dropsondes has drastically changed this situation [*Hock and Franklin*, 1999]. Dropsondes are free of problems with swinging payloads and wakes and they have rapid descent times (about 15 min from the top of the troposphere) and, with the help of GPS tracking, they have high vertical resolutions (of the order of 5 m, although see the discussion below). The data discussed here were part of the Winter Storms 2004 experimental campaign, held in the western Pacific using the NOAA Gulfstream 4 aircraft. During a 2-week period, 10 flights each dropped 20–30 sondes (see Figure 2), a total of 262. Of these, 237 reasonably complete sets were analyzed by *Lovejoy et al.* [2007, 2008c]. An example of meteorological fields estimated by a single sonde dropped at about 12 km altitude is shown in Figure 3 (the data were interpolated to a regular 5 m grid).

[15] The main problems with the dropsonde data are from the frequent data outages (mostly caused by transmission problems and difficulties getting GPS locking although there are probably other factors). A lesser problem is the variable fall speed which when combined with the outages yield very irregular data sets. *Lovejoy et al.* [2007] used special (point pair based) analysis techniques (which avoided problematic interpolations) to estimate the gradients in the horizontal wind as functions of the thickness of the atmospheric layer and the altitude (see section 4.4). In the work by *Lovejoy et al.* [2008c], where traditional stability criteria were evaluated, the conclusions about a fractal hierarchy of unstable layers embedded within stable layers were substantiated by the frequent use of pairs of sondes dropped at 0.3 s intervals (corresponding to roughly 30 m apart). Intercomparison of such pairs showed for example that they agreed to within ±0.014 K in temperature and ±0.1 m/s in horizontal wind directly giving upper bounds on the measurement errors.

3.2. Intermittent Multifractal Sampling: The Problem of Outages

[16] The cascade structure equation (1) is the consequence of variability building up scale by scale over a potentially large-scale ratio λ. In order to uncover the cascade structure, i.e., to verify equation (1) and to estimate $K(q)$, we must attempt to invert the cascade process by systematically degrading the resolution of the fluxes by averaging. This is straightforward enough for data sampled at regular intervals (see equation (5)) but for data with highly irregular resolutions we must take into account the variability of the resolution. The resolution is variable for two reasons: first, even if the sampling was at the nominal 0.5 s, the variable vertical sonde fall speed would lead to variable vertical sampling intervals. This source of variability is not too large: the mean vertical sonde velocity decreases from about 18 m/s to about 9 m/s near the surface

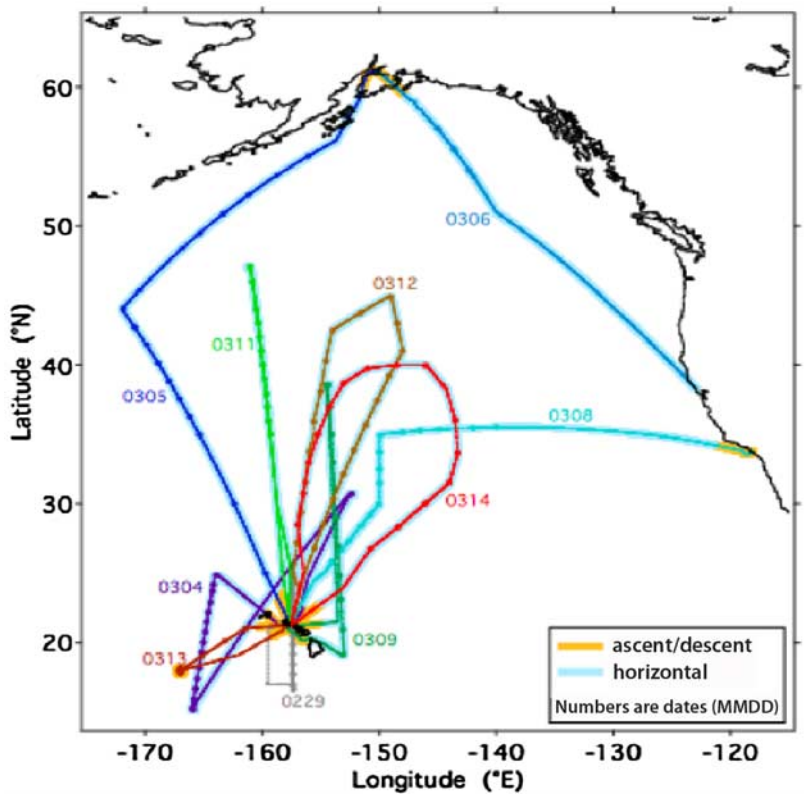

Figure 2. The position of the dropsondes used in this study. The dots indicate the locations of the drops, and the lines are the corresponding aircraft trajectories with dates (four digits: mmdd).

(owing to increased air resistance; see Figure 4) and turbulence induced fluctuations increase this range of resolution by about another factor of 2. However, the variability problem is made much worse because of the outages. To see how serious the problem is we refer the reader to Figure 4, which shows the comparison of vertical intervals for one of the simultaneous sonde pairs mentioned above. We see from Figure 4 that the mean intermeasurement vertical distance Δz is about 9 m (the data were sampled at a minimum interval of 0.5 s, on average, every 0.6 s) that there are several very large outages (the maximum in Figure 4 is 140 m) and that the outages are in fact highly clustered. This example is quite typical: for our study we chose a subsample of 220 of the 262 sondes which had over 1000 points and started above 10 km in altitude; the overall mean sampling intervals were 0.60 s, 9.29 m. Sonde by sonde, the mean minimum time interval is 0.5 s and mean minimum vertical interval is 3.70 m. In addition over all measurements of all of the 220 sondes, 90.5% were sampled at the design frequency of 0.5 s so that on average the outages affected less than 10% of the values, although each sonde had numerous outages. As mentioned below, we also examined a particularly low outage subsample of 44 sondes, but even this had essentially the same problems.

[17] To better understand the distribution of outages, we therefore determined the scaling properties of the time and distance intervals as function of the number of measurements. In fact, we used the same scaling form as equation (1),

$$\begin{aligned} \langle \Delta t_\lambda^q \rangle &= \lambda^{K_t(q)} \langle \Delta t_\lambda \rangle^q; \\ \langle \Delta z_\lambda^q \rangle &= \lambda^{K_z(q)} \langle \Delta z_\lambda \rangle^q; \end{aligned} \qquad \lambda_n = \frac{N}{\Delta n}; \quad \lambda = \frac{\lambda_n}{\lambda_{eff}}, \qquad (6)$$

where $\Delta t(\Delta n)$, $\Delta z(\Delta n)$ are the temporal and vertical distances between measurements (indexed with the integers n) separated by Δn (out of a total of N measurements in all, e.g., $\Delta t(\Delta n) = t(n + \Delta n) - t(n)$. Equation (6) is simply the same behavior as equation (1) except that the intervals Δt, Δz are considered to be the cascade quantities. In this case, there is no compelling theoretical reason to expect this form, except that it is presumably the consequence of a hierarchical scaling outage mechanism and Figures 5a and 5b show this scaling is remarkably well respected in both time and space (with outer scales corresponding to ≈200 s and 3.2 km respectively). Note that these outer scales are presumably underestimates of the true outer scale since the sondes with the most extreme outages (42 out of 262) were not analyzed (too few measurements or over too thin an atmospheric layer). We could also note that the intermittency as characterized by the codimension of the mean (equation (2)) is very high: $C_{1t} \approx 0.26$, $C_{1z} \approx 0.23$ (indeed they are much higher than those of the fields, see section 4!). Also, using fluctuation analyses (section 4.4) we found that $H \approx 0.03 \pm 0.04$ in both time and space so that the assumption (equation (6)) that Δt, Δz are conservative cascade quantities ($H = 0$) is reasonably well obeyed. (Recall that from equation (4) that if $H = 0$, then the mean fluctuation $\langle \Delta v \rangle = \langle \varphi_{\Delta z} \rangle$ = constant, it is independent of scale Δz, it is "conserved" from one scale to another).

[18] The direct examination of the probabilities will be useful in what follows. Consider the probability distribution of the temporal and spatial increments: $\Pr(\Delta t > s)$, $\Pr(\Delta z > s)$ at the finest resolution (s is a threshold and "Pr" indicates "probability." The tails on the distributions are very long,

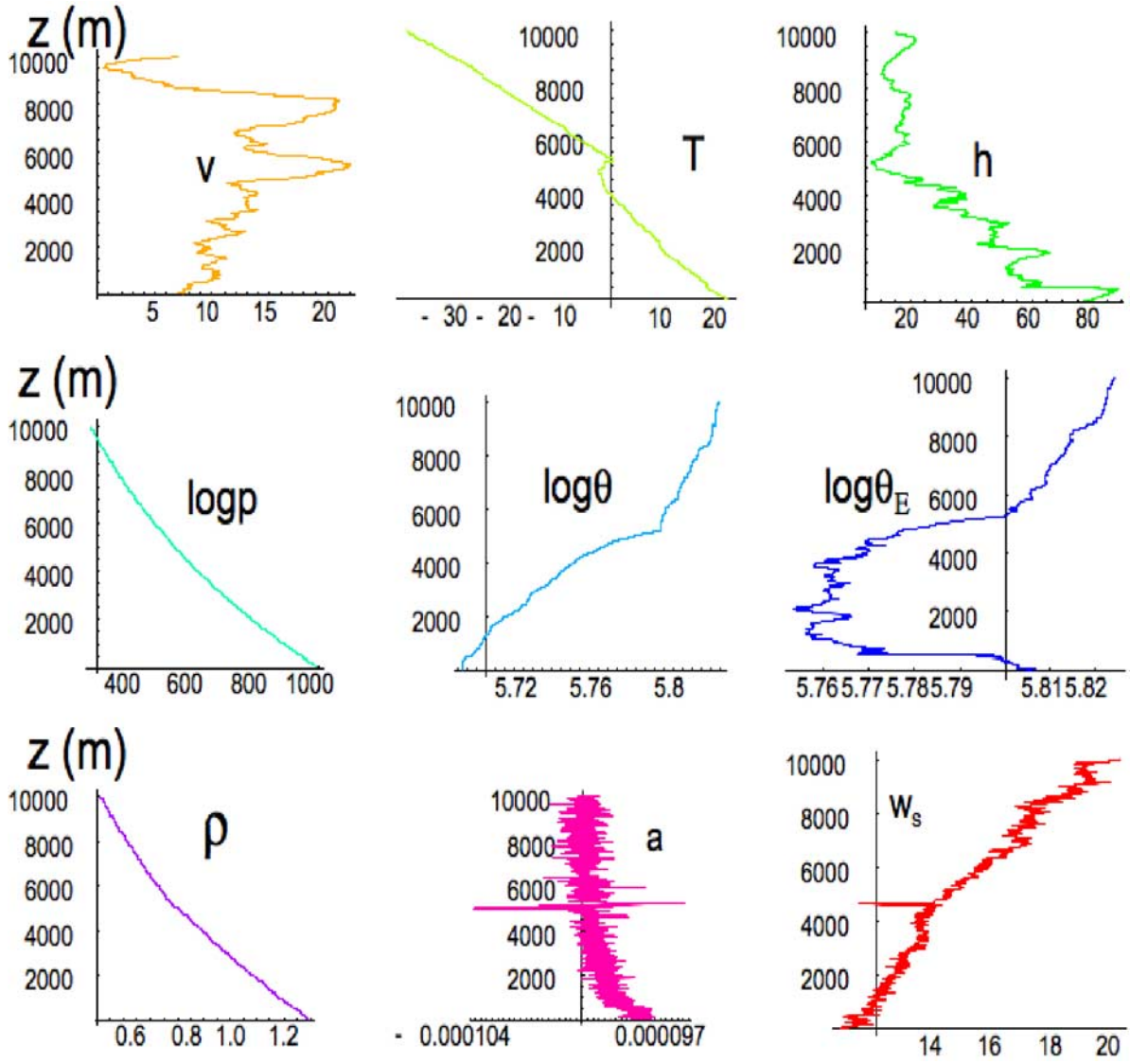

Figure 3. Representative outputs of various fields for sonde 1, 20040229 (yyyymmdd). The units are velocity (v, m/s), temperature (T, °C), humidity (h, percent), pressure (p, millibars), log potential temperature (log θ, dimensionless), log equivalent potential temperature (log θ_E, dimensionless), total air density (ρ, Kg/m^3), vertical air acceleration (a, dimensionless normalized by g, estimated from pressure gradients), and sonde vertical velocity (w_s, m/s).

this can be quantified by comparing the distributions with the algebraic form $\Pr(\Delta > s) \approx s^{-q_D}$ where according to Figure 6a, the exponent $q_D \approx 2$ which would imply that the variance does not converge (the multiscale estimate in Figure 6b indicates that q_D is probably a little higher ≈2.2). Note that, as indicated, the straight lines in Figures 6a and 6b are just for reference, they are not regressions.

[19] However, Figure 6a only characterizes the probabilities at a single resolution; in order to obtain a multiscale characterization, we note that the analysis the probabilities/histograms is equivalent to the analysis of the statistical moments of all orders (equation (6)). The corresponding scaling form of the probabilities for fields obeying the moment equation (1) is

$$\Pr(\gamma' > \gamma) \approx \lambda^{-c(\gamma)}; \quad \gamma = \frac{\log \varphi_\lambda}{\log \lambda}, \tag{7}$$

where "Pr" means "probability" and c is the codimension function and γ is the "order of singularity" and we have normalized the flux φ_λ so that $\langle \varphi_\lambda \rangle = 1$. The probability exponent $c(\gamma)$ is related to the moment scaling exponent $K(q)$ by the Legendre transform [*Parisi and Frisch*, 1985],

$$\begin{aligned} c(\gamma) &= \max_q (q\gamma - K(q)) \\ K(q) &= \max_\gamma (q\gamma - c(\gamma)) \end{aligned} . \tag{8}$$

This establishes a one to one relationship between singularities and statistical moments q: $c'(\gamma) = q$, $\gamma = K'(q)$.

[20] To analyze the biases induced by the outages, it turns out we will need a characterization of the low orders of singularities which for multifractal Δt, Δz corresponds to the finest resolutions. Since the low-order singularities are fairly tightly bounded below (measurements are no closer

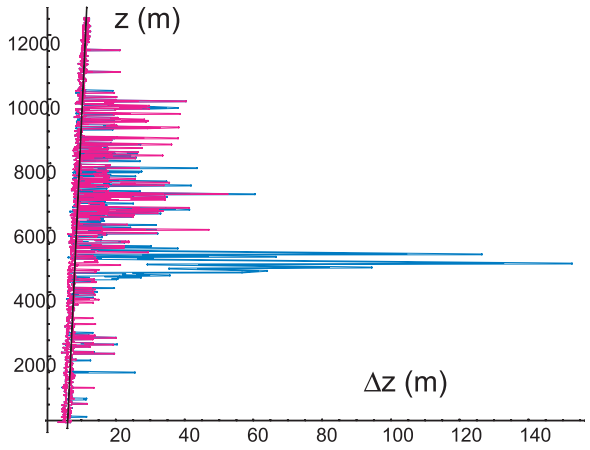

Figure 4. Intercomparison of vertical sampling intervals of two (near) simultaneous sondes (red and blue). Notice the strong (and typical) clustering of the outages. The mean Δz is larger at high altitudes owing to the lower air resistance.

than 0.5 s apart), the minimum orders of singularity are actually quite near zero. We need an estimate of $\gamma_{\min}$ since this will turn out to be the dominant singularity in the sampling. Using the finest scale resolution only (as in Figure 6a), we obtain $\gamma_{\min,t} \approx \log(\Delta t_{\min}/\langle\Delta t\rangle)/\log N = -0.026$, $\gamma_{\min,z} \approx \log(\Delta z_{\min}/\langle\Delta z\rangle)/\log N \approx -0.126$ (where we have used the normalized cascade quantities $\varphi_\lambda = \Delta t_{\min}/\langle\Delta t\rangle$ and $\Delta z_{\min}/\langle\Delta z\rangle$). The maximum scale ratio = N is the total number of measurements; since this varied somewhat from sonde to sonde, we used the average. In order to get a multiscale estimate of $\gamma_{\min}$ which is not too sensitive to the smallest scale, we can use the Legendre relation $\gamma = K'(q)$ and from Figure 5c graphically find the line which has the minimum slope: $\gamma_{\min} = \mathrm{Min}(K'(q))$. Doing this, we obtain Figure 6b and for the vertical distances $\gamma_{\min,z} = -0.197$ which is smaller than, but close to the previous estimate. In the time domain, the intervals with no outages (i.e., 0.5 s temporal resolutions) we have the same $\gamma_{\min,t}$ as above: -0.026. For the minimum of the intervals with outages we find $\gamma_{\min,t} = 0.187$, on a fractal set with $c(\gamma_{\min,t}) = 0.196$ (i.e., the point at $\gamma = -0.026$ is an isolated discontinuity in the $c(\gamma)$ function). We can see from Figure 6b that the $c(\gamma)$ are almost identical (except for the lowest singularities): this is another confirmation that the main intermittency is caused by the outages and not by the variable fall speeds. In Figure 6b we added the bisectrix (the line at 45°) because this line is tangent to the $c(\gamma)$ at the point C_1 (i.e., $c(C_1) = C_1$ and $c'(C_1) = 1$; see Table 1). Since universal multifractals have $\gamma_{\min} \leq -C_1$, we see that these cannot be particularly well approximated by universal multifractals so that we did not estimate the universal parameter α.

4. Data Analysis

4.1. Estimating the Fluxes From Highly Intermittent Data

[21] In spite of the high intermittency of the outages, we mentioned that over 90% of the data were sampled at the high design rate of 0.5 s so that we anticipate the corrections may not be too large. Before describing how to statistically correct for the multifractality of the outages, we first demonstrate a robust technique for estimating the fluxes. To do this, we recall the discussion in section 2 that in a scaling regime for a field v, the fluxes are related to the fluctuations as $\Delta v(l) = \varphi_l l^H$. With regularly spaced data, the usual way to estimate the flux is to degrade it starting from the highest resolution in the scaling regime (so that the above law holds), and estimate the fluctuations at the finest scale (e.g., by absolute first differences or by absolute wavelet coefficients). The result is the flux at the finest resolution; one then degrades the result by averaging over larger and larger scales (smaller and smaller λ, equation (5)). However, we can also estimate the fluctuations from the local derivatives, for example (for uniformly spaced data) for the nth flux estimate, we can use the data at the $(n-1)$th and $(n+1)$th points to yield a "centered difference" estimate

$$\Delta v_l(z_n) \approx \left|\frac{dv(z_n)}{dz}\right|_l l \approx s_n l; \quad s_n = \left|\frac{v(z_{n+1}) - v(z_{n-1})}{z_{n+1} - z_{n-1}}\right|, \qquad (9)$$

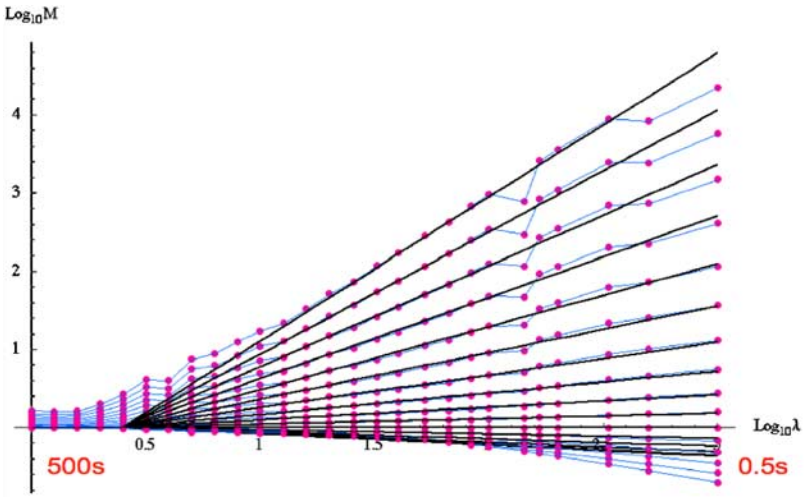

Figure 5a. Intermittency of the temporal sampling intervals: for the 220 sondes with $N > 1000$ which were dropped from $z > 10$ km. The lines correspond to the moments $q = 0, 0.2, \ldots 2$. The lines converge to the outer scale of about 200 s and the scaling is well respected up to ≈50 s.

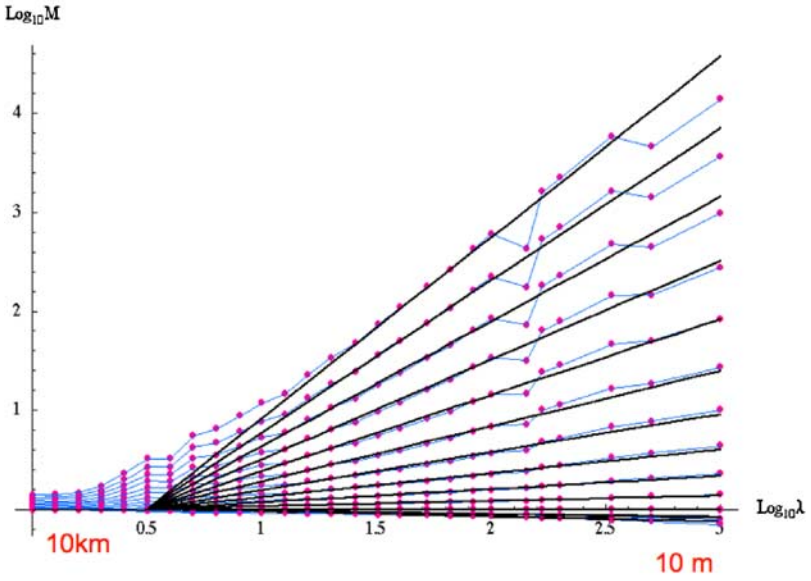

Figure 5b. The intermittency of the vertical sampling interval corresponding to Figure 5a. The lines correspond to the moments, the outer scale is about 3 km, and the scaling is well respected up to ≈1 km.

where s_n is the local absolute slope. Since we saw that in the scaling regime the normalized flux is $\varphi_l = \Delta v(l)/\langle \Delta v(l) \rangle$ we therefore have the estimate $\varphi_l = s/\langle s \rangle$. In order to generalize this to intermittent locations z_n, we can simply include the point z_n and estimate s_n by a linear regression of the point triplets: $\{(z_{n-1}, v(z_{n-1})), (z_n, v(z_n)), (z_{n+1}, v(z_{n+1}))\}$. This method has the advantage of not being too sensitive to the noise in either the data (v) nor the position estimates (z) (these factors were found to make interpolations unreliable for derivative estimates; similarly estimates of spectra obtained from interpolated series contain serious artifacts of the interpolation technique). The resulting series of s_n values determined at the irregular locations z_n can then be interpolated at uniform intervals. The (normalized) slopes themselves can be used as estimates of the normalized fluxes ($\varphi_l = s/\langle s \rangle$) as long as one recognizes that their resolutions are not fixed so that the regularly sampled values obtained from the interpolated s series can be systematically degraded as discussed in section 2 to obtain a series of lower and lower resolution flux estimates. The last step discussed below is to statistically correct the result for this variable resolution effect.

[22] To see how this method works we determined the fluxes for the simultaneous sonde pair analyzed in Figure 4. The results for the main dynamic and thermodynamic fields are shown in Figure 7a. We can see that the estimates for the two sondes are very similar (the curves are mostly indistinguishable) even though as Figure 4 shows, the outages were significantly different. In Figure 7b, we blow up a particularly intermittent section, which shows both the enormous variability of the fluxes (especially the humidity and equivalent potential temperature) and that with this method it is well reproduced by the two sondes. Finally, in Figure 7c we examine a section of the data that was particularly poorly sampled by the sondes (see Figure 4): we see that even here the two sondes give fairly similar flux estimates although at different resolutions.

4.2. Statistically Correcting for the Outages

[23] We now develop some simple theory to show how the statistics of the fluxes with variable resolutions can be statistically corrected for the multifractal intermittency of the resolutions. Since we will see that the corrections do not qualitatively change our empirical results (and are not so large), those who are interested in results can skip to the next subsection.

[24] Let us assume that equation (6) for the outages is correct (i.e., that they are indeed multifractal), we will only consider the results with respect to the scaling in the vertical, but the same formalism would apply for the temporal scaling. In this case, applying equation (7) we find that the probability distributions at resolutions λ_n, λ_z are

$$\begin{matrix} \Pr(\gamma_n > s) \approx \lambda_n^{-c_z(\gamma_z)} \\ \Pr(\gamma > s) \approx \lambda_z^{-c(\gamma)} \end{matrix} , \qquad (10)$$

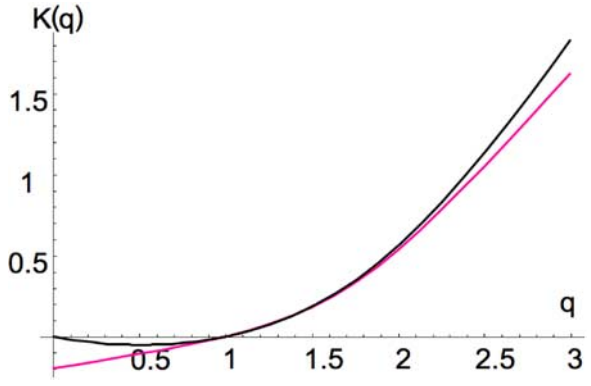

Figure 5c. The exponents $K(q)$ of the sampling intervals (from the slopes in Figure 5a and 5b). Black is $K(q)$ for z, and pink is for t.

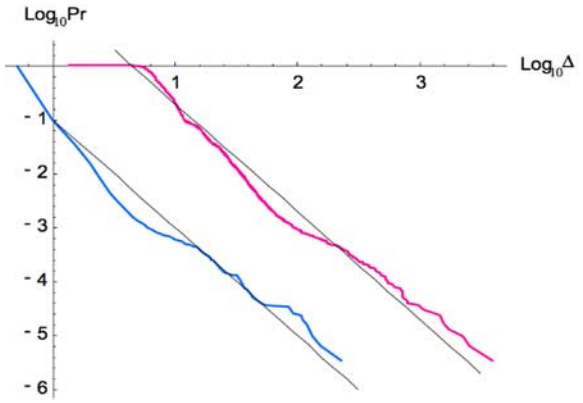

Figure 6a. The cumulative probability distribution of the temporal increments Δt (blue, units, second), and layer thickness Δz (pink) for 220 sondes at the finest measurement scale (i.e., the Δt, Δz between consecutive measurements). The reference lines have slopes of -2.

where γ_z is the random singularity of the intermeasurement distance Δz with respect to the number of measurements, and γ is the random singularity of the flux of interest with respect to the vertical distance Δz and the c_z, c are the corresponding codimension functions

$$\begin{array}{c} \Delta z = \lambda_n^{\gamma_n}\langle \Delta z\rangle; \quad \lambda_n = \frac{N}{\Delta n} = \frac{L}{\langle \Delta z\rangle} \\ \varphi_{\lambda_z} = \lambda_z^{\gamma}; \quad \lambda_z = \frac{L}{\Delta z} \end{array}, \tag{11}$$

where we have assumed that the flux φ has been normalized: $\langle\varphi\rangle = 1$. $c_z(\gamma_z)$ is the codimension function (given by Legendre transform from $K_z(q)$ in equation (8)), N is the total number of measurements in the series and Δn is the (integer) number of measurements taken over a given layer thickness Δz. Given that the empirical flux follows the multiscaling equation (1), then we must take into account the variable vertical resolution (λ_z) via

$$\lambda_z = \frac{L}{\Delta z} = \frac{L}{\lambda_n^{\gamma_z}\langle \Delta z\rangle} = \lambda_n^{1-\gamma_z}. \tag{12}$$

This shows how the random outage singularity γ_z affects the vertical resolution λ_z. Substituting λ_z and $\varphi_{\lambda z}$ into equation (1), and averaging over the flux singularities γ at a fixed γ_z,

$$\left\langle \varphi_{\lambda_z}^q \right\rangle_{\gamma_z} = \left\langle \lambda_n^{(1-\gamma_z)q\gamma} \right\rangle_{\gamma_z} = \lambda_n^{K_\varphi((1-\gamma_z)q)} \tag{13}$$

(the subscript "γ_z" on the statistical average means that the average is only over γ, i.e., conditional on γ_z). If we now average over the γ_z variability, we can define the "effective" moment scaling function $K_{\varphi,eff}(q)$,

$$\begin{aligned} \left\langle \varphi_{\lambda_z}^q \right\rangle = \lambda_z^{K_{\varphi,eff}(q)} &= \int \left\langle \varphi_{\lambda_z}^q \right\rangle_{\gamma_z} p(\gamma_z) d\gamma_z \\ &= \int \lambda_n^{K_\varphi((1-\gamma_z)q)} \lambda_n^{-c_z(\gamma_z)} d\gamma_z, \end{aligned} \tag{14}$$

where we have used the fact that the probability density of γ_z (obtained by differentiation of the distribution Pr, equation (10)) is $p(\gamma_z) \approx \lambda_n^{-c_z(\gamma_z)}$ where "$\approx$" indicates "to within slowly varying factors" (i.e., logarithms) and $\langle \varphi_{\lambda_z}{}^q \rangle = \lambda_z^{K_\varphi(q)}$. Using the standard saddle point method for evaluating asymptotic approximations to integrals of the above type (valid for large log λ_z), we obtain

$$K_{\varphi,eff}(q) = \operatorname*{Max}_{\gamma_z} \left(K_\varphi(q(1-\gamma_z)) - c_z(\gamma_z) \right). \tag{15}$$

To find the maximum of the expression in equation (15) we now use the fact that c is an increasing function with Min(c) = 0, and for all $q > q_0$ with $q_0 < 1$, $K_\varphi(q)$ is also increasing (actually for all the universal multifractals $q_0 < 1/2$). Start by considering large q; decreasing γ_z will therefore increase K and decrease c hence to maximize equation (15), γ_z must be decreased to its minimum possible value (which here is slightly below 0); this is valid at least for all $q > q_0/(1-\gamma_{\min,z})$ and is a good approximation even for $q < q_0/(1-\gamma_{\min,z})$. With $\gamma_{\min,z} < 0$ and $q_0 < 1/2$, this argument is valid for most of the interesting range of q. It implies

$$K_{\varphi,eff}(q) = K_\varphi\left(q\left(1-\gamma_{\min,z}\right)\right); \quad q > q_0/\left(1-\gamma_{\min,z}\right). \tag{16}$$

Since the minimum of K is usually small, this argument may actually be valid over the entire range of q, although the range of equation (16) is adequate for our purposes. Since $c_n(\gamma_{\min,z}) = 0$, the dominance of $\gamma_{\min,z}$ corresponds to the most probable line filling singularity. The other singularities, which have $c_z(\gamma_z) > 0$, correspond to sparse fractal subsets of the sampling space.

[25] If in addition, φ is a universal multifractal, index α then the normalized $K_{\varphi,eff}$ (obtained by $\varphi \rightarrow \varphi/\langle\varphi\rangle$ hence $K(q) \rightarrow K(q) - qK(1)$) will be given by

$$K_{\varphi,eff}(q) = \left(1-\gamma_{\min,z}\right)^{\alpha} K_\varphi(q) \approx \left(1-\alpha\gamma_{\min,z}\right) K_\varphi(q) \tag{17}$$

(the binomial approximation on the far right is valid for small $|\gamma_{\min,z}|$). In the above we have seen empirically that

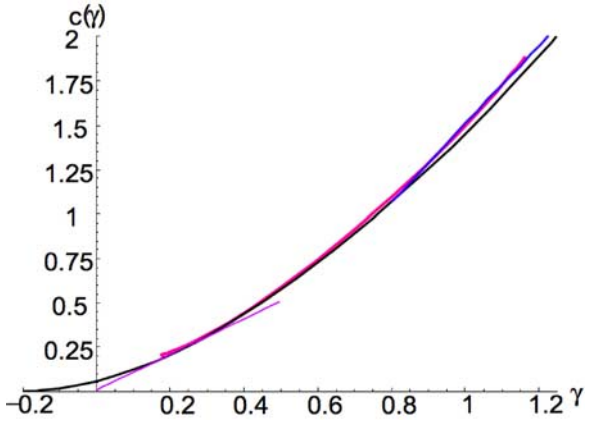

Figure 6b. The codimension function $c(\gamma)$ of the sampling intervals obtained by Legendre transform of Figure 5c. Black is for z, and pink for t. The reference lines are $x = y$ (purple), and the asymptote (blue) with slope 2.2. The former is tangent to $c(\gamma)$ at $\gamma = c(\gamma) = C_1$, and the latter is an estimate of the critical moment of divergence q_D.

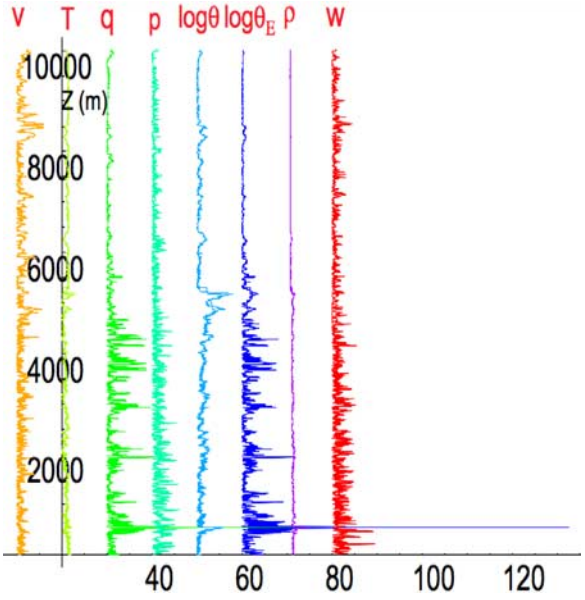

Figure 7a. Intercomparison of normalized (and hence nondimensional) fluxes from a simultaneous sonde pair. For each field, the profiles are so close as to be nearly indistinguishable. The acceleration was not analyzed since the noise was too large.

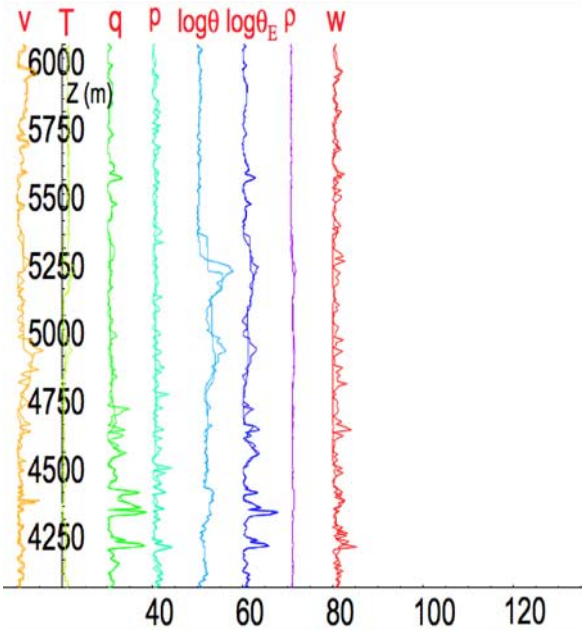

Figure 7c. Intercomparison of normalized flux estimates from the two sondes for a particularly poorly sampled section from Figure 7a.

$\gamma_{\min,z} = -0.197$, so that we get a boost for K_φ (and hence C_1) of a factor ≈ 1.39 (assuming $\alpha = 1.8$) implying (as expected) an increase in the intermittency.

[26] In order to test this correction factor, we made 220 multifractal simulations of the horizontal wind with the parameters $C_1 = 0.05$, $\alpha = 1.8$, $H = 0.6$, $\lambda = 2^{14}$ (corresponding for example to an outer scale of about 16 km and 1 m resolution). We then used the empirically observed z coordinates from the 220 sondes to intermittently resample these simulated wind fields and applied the regression based algorithm discussed above to estimate the fluxes. The result is shown in Figure 8. One can see that the theoretically predicted cascade behavior is well respected except at the smallest scales (deviations are mostly over roughly a range of factor 2–3). The outer scale is close to 16 km as in the model and the C_1 is 0.067 which is very close to that predicted by equation (17): $0.05\ (1 + 0.197)^{1.8} = 0.069$ (i.e., using $\gamma_{\min,z} = -0.197$).

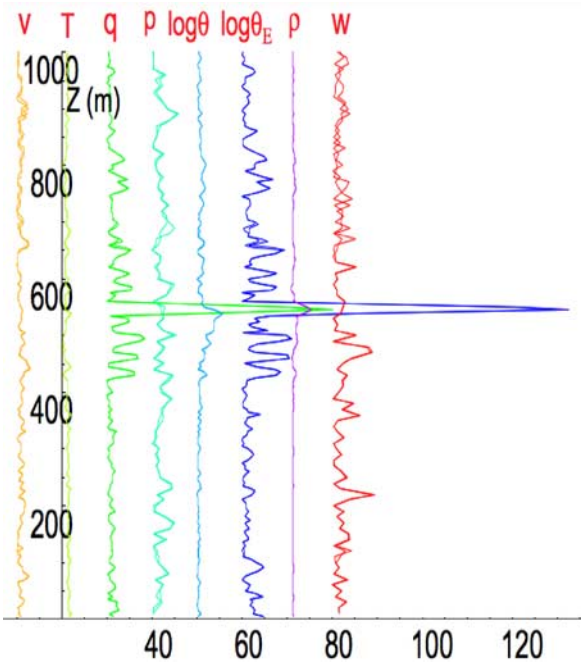

Figure 7b. Detail of Figure 7a showing that fluxes even from very intermittent layers can be well reproduced from sonde to sonde.

4.3. Flux Analyses

[27] We now turn to the analysis of the sondes. The quantities that we analyzed can be roughly grouped into two categories: dynamical and thermodynamical variables. The dynamic variables (Figure 9a) were the modulus of the horizontal wind v, the pressure p, the total air density (ρ, including that due to humidity), and the sonde vertical velocity w_s. We also separately analyzed the north–south and east–west components of the horizontal wind but the results were not much different and we will not discuss them further. For the vertical sonde velocity, the fluctuations around a quadratic fit (corresponding to a constant deceleration from 18 m/s to 9 m/s) were used. Owing to the parachute drag, the fluctuations in w_s depends on both the vertical and horizontal wind so that it should not be used as a surrogate for the vertical wind.

[28] The thermodynamic variables are temperature (T), log potential temperature ($\log\theta$), log equivalent potential temperature ($\log\theta_E$) and humidity (h); see Figure 9b. The $\log\theta$ and $\log\theta_E$ are proportional to the entropy densities of the dry and humid air respectively. In addition, their structure is important for the overall atmospheric (static)

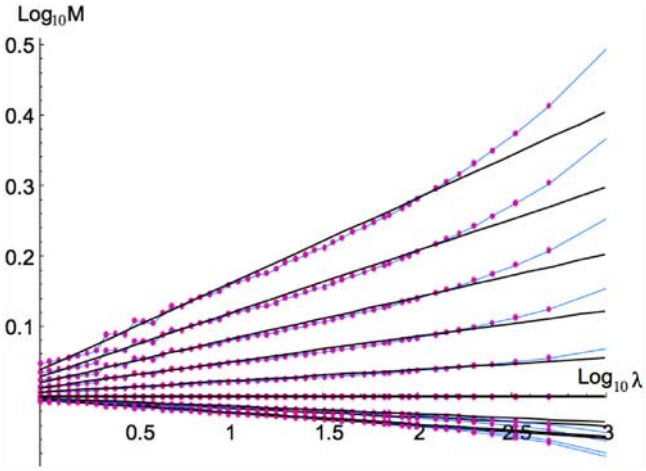

Figure 8. Analysis of a simulated multifractal field with parameters (H = 0.6, C_1 = 0.05, α = 1.8, outer scale ≈ 16 km) using empirically observed intermittent z coordinates (the moments of the interpolated fields were estimated using the regression method). The empirical C_1 = 0.067 is close to the theoretically predicted value 0.05 $(1 + 0.198)^{1.8} = 0.069$. The moments are q = 0, 0.2, 0.4 . . . 2.

stabilities, for example $gd\log\theta/dz$ is the square of the Brunt-Väisälä frequency so that where the latter is negative, the atmosphere is considered conditionally unstable, it is convectively unstable when $gd\log\theta_E/dz$ is negative.

[29] From Figures 9a and 9b we can see that with small deviations (see Tables 2a and 2b) all the fields have small residuals with respect to the predictions of cascade theories (equation (1); the residuals were averaged over the range 1 km to 10 m).

[30] The parameter H was estimated from ("generalized," qth order) structure functions $S_q(\Delta z)$,

$$S_q(\Delta z) = \langle \Delta v(\Delta z)^q \rangle \approx \Delta z^{\xi(q)}; \quad \xi(q) = qH - K(q), \qquad (18)$$

where the fluctuation Δv was estimated using pairs ($H < 1$) and triplets ($H > 1$) of measurements described in section 4.4; $\xi(q)$ is the structure function exponent; the relation equation (18) with the flux exponent $K(q)$ and the nonconservation parameter H is obtained using equations 1, 4; and α was estimated to the nearest 0.05 using the "double trace moment technique" which involves repeating the cascade analysis but with the flux at the finest resolution raised to a series of different powers [*Lavallée et al.*, 1993]. All the regressions were taken over range $1.5 < \log_{10}\lambda < 3$ with λ defined as the ratio of the reference-scale 10 km to the resolution scale (i.e., corresponding to 300 m to 10 m). Also shown are estimates of the parameters C_1 and λ_{eff} taken from the lowest 4 km and the highest 4 km only.

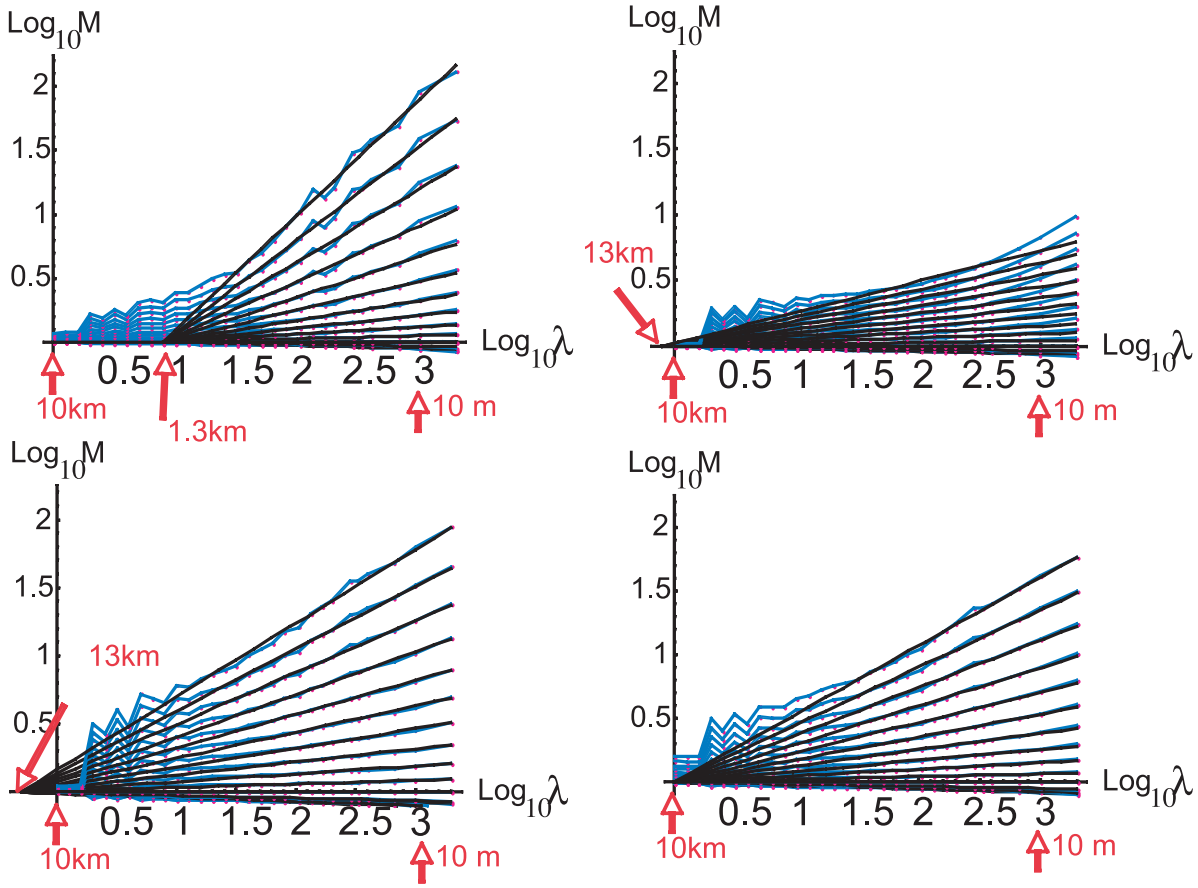

Figure 9a. The dynamical fields (left to right, top to bottom) v, p, ρ, and w_s, for q = 0.2, 0.4, . . . 2.

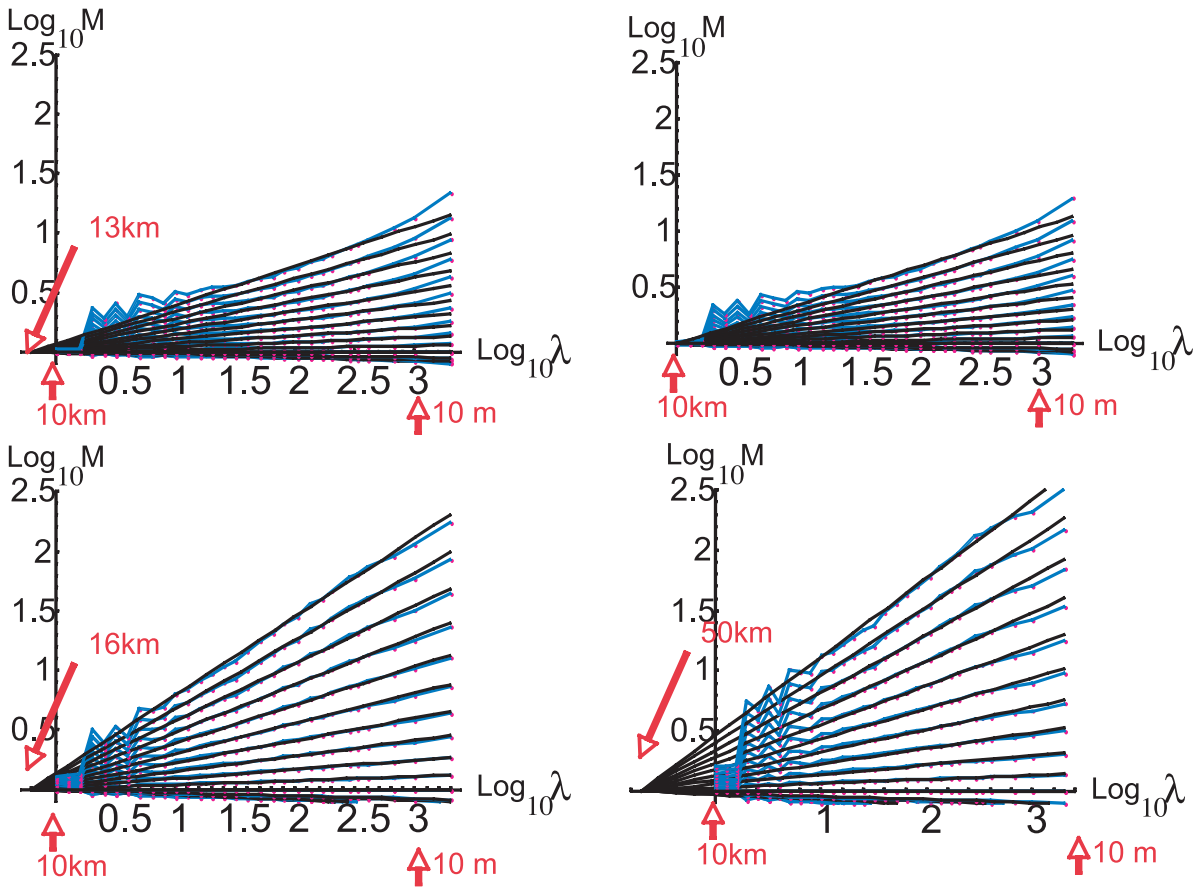

Figure 9b. The same as Figure 9a except for the thermodynamic fields (left to right, top to bottom) T, log θ, h, and log θ_E.

[31] The difference in C_1 between high- and low-altitude regions is about 0.04 in all cases except for h and log θ_E (where they are about same). In all cases the outer scale of the high-altitude data is larger than the outer scale of the low-altitude data by about a factor 2 (except for log θ_E and ρ). Also the outer scales of the 4-km-thick subsamples are typically smaller than those of the total (10 km) thick layer (which combines the variability of both). For a fixed scale, and for fixed exponents, a larger outer scale means larger λ, hence larger variability so that the overall increase of variability with increasing altitude is a combination of increased exponents (larger C_1 hence $K(q)$), and larger outer scales. In the case of the horizontal wind whose fluctuations were studied in detail by *Lovejoy et al.* [2007] at systematically increasingly altitudes, we can compare them with the current results (see also Figure 10a). It was found that the value of $C_1 \approx 0.04 \pm 0.02$ and was nearly altitude independent (the standard deviation is the typical sonde to sonde spread in the estimates) with H increasing from 0.60 near the surface to 0.77 at 12 km. The value of α was roughly 1.7 ± 0.2 with a slight tend to decrease with altitude. We can see that these values are fairly close to those in Tables 2a and 2b (see in particular the corrected $C_1 \approx 0.07$).

[32] Two comments are in order. First, we see that in some cases (notably the pressure, Figure 9a, top right), the scaling of the fluxes is not so good. However, there are several factors that may contribute to this. For example, in the case of pressure, the variability is quite small (all the axes in Figures 9a and 9b are identical in order to facilitate intercomparisons) so that the relative importance of statistical noise (due to inadequate sample size, data quality) is more important relative to the signal. Also, the deviations from the scaling are actually still fairly small (see Tables 2a

Table 2a. The 220 Vertical Sondes: Characteristics of the Dynamical Fields[a]

Field	$\overline{\delta}$(%)	$C_{1,eff}$	$C_{1,eff}$ < 4 km	$C_{1,eff}$ > 6 km	C_1	$C_{1,fluc}$	H	α	$Log_{10}\lambda_{eff}$	$Log_{10}\lambda_{eff}$ < 4 km	$Log_{10}\lambda_{eff}$ > 6 km
v	**2.3**	**0.100**	0.078	0.121	0.071	0.023 ± 0.016	0.75 ± 0.05	1.90	**0.9**	1.1	0.9
p	**1.1**	**0.045**	0.050	0.099	0.032	0.043 ± 0.032	1.95 ± 0.02	1.85	**−0.1**	1.5	1.4
ρ	**1.4**	**0.093**	0.072	0.105	0.065	0.123 ± 0.103	1.31 ± 0.12	1.95	**−0.1**	0.9	1.2
w_s	**1.6**	**0.092**	0.081	0.099	0.064	0.041 ± 0.027	0.68 ± 0.01	2.00	**0.5**	0.1	0.5

[a]The estimates of the basic cascade parameters, for the dynamic variables; those in bold are directly estimated from the graphs in Figure 9a. The $C_{1,fluc}$ and H are from the fluctuation analysis (section 4.4) and are the means of the fits from 30–300 m and 300–3000 m; the spread is half the difference. The column C_1 is estimated from the equation (17), which corrects for the multifractal outages. The scale ratio λ was defined with respect to 10 km.

Table 2b. The 220 Vertical Sondes: Characteristics of the Thermodynamic Fields[a]

Field	$\bar{\delta}$ (%)	$C_{1,eff}$	$C_{1,eff} < 4$ km	$C_{1,eff} > 6$ km	C_1	$C_{1,fluc}$	H	α	$Log_{10}\lambda_{eff}$	$Log_{10}\lambda_{eff} < 4$ km	$Log_{10}\lambda_{eff} > 6$ km
T	**1.4**	**0.067**	0.0565	0.099	0.049	0.066 ± 0.038	1.07 ± 0.18	1.70	**0.3**	1.2	1.2
h	**1.4**	**0.144**	0.130	0.125	0.103	0.144 ± 0.028	0.78 ± 0.07	1.85	**−0.2**	0.5	0.
Logθ	**1.2**	**0.065**	0.057	0.097	0.046	0.051 ± 0.027	1.07 ± 0.18	1.90	**0.4**	1.2	1.2
Logθ_E	**1.9**	**0.145**	0.115	0.091	0.102	0.140 ± 0.101	0.87 ± 0.10	1.95	**−0.4**	0.4	0.8

[a]The estimates corresponding to Table 2a but for the thermodynamic variables (Figure 9b).

and 2b; ±1.1% here); in section 4.4 we see that the scaling of the fluctuations (rather than the fluxes) of the same quantity is in fact excellent, and over a significantly wider range. This brings us to our second point: in many cases the outer scale of the fluxes is quite a bit less than the atmospheric-scale height/troposphere thickness (roughly 8 and 12 km, respectively). However, even in cases such as the horizontal wind where the outer scale of the fluxes is ≈1.3 km, we see that the scaling of the fluctuations (Figures 10a and 10b) is excellent right through the largest available scales (10 km).

[33] Finally, we should also mention that an attempt was made to avoid the outage problem by selecting only sondes whose maximum outage was less than 3 s long; there were 44 of the previous 220 sondes which satisfied this low-outage criterion. However, although shorter in duration, the outages were still very frequent and highly clustered so that the measurements were still very irregular. When the analyses were performed, it was found that they were not very different from the full analyses above and generally the parameter estimates were not too different (the C_1 values were all within ±0.02 of the values in Tables 2a and 2b except for the wind which was 0.05 smaller). The differences can probably be accounted for by the much smaller sample size.

4.4. Fluctuation Analyses: Data Point Pairs and Triplets

[34] The cascades analyses are important because they give direct evidence on the cascade structure of the atmosphere in the vertical, they yield direct estimates of the outer cascade scale, and they give relatively reliable estimates of the statistical exponents characterizing the fluxes, in particular the mean intermittency exponent C_1. However, they do not give a full characterization of the field since they do not estimate the nonconservation exponent H which characterizes how far the fluctuations are from the scale by scale conserved fluxes analyzed in the previous section. *Lovejoy*

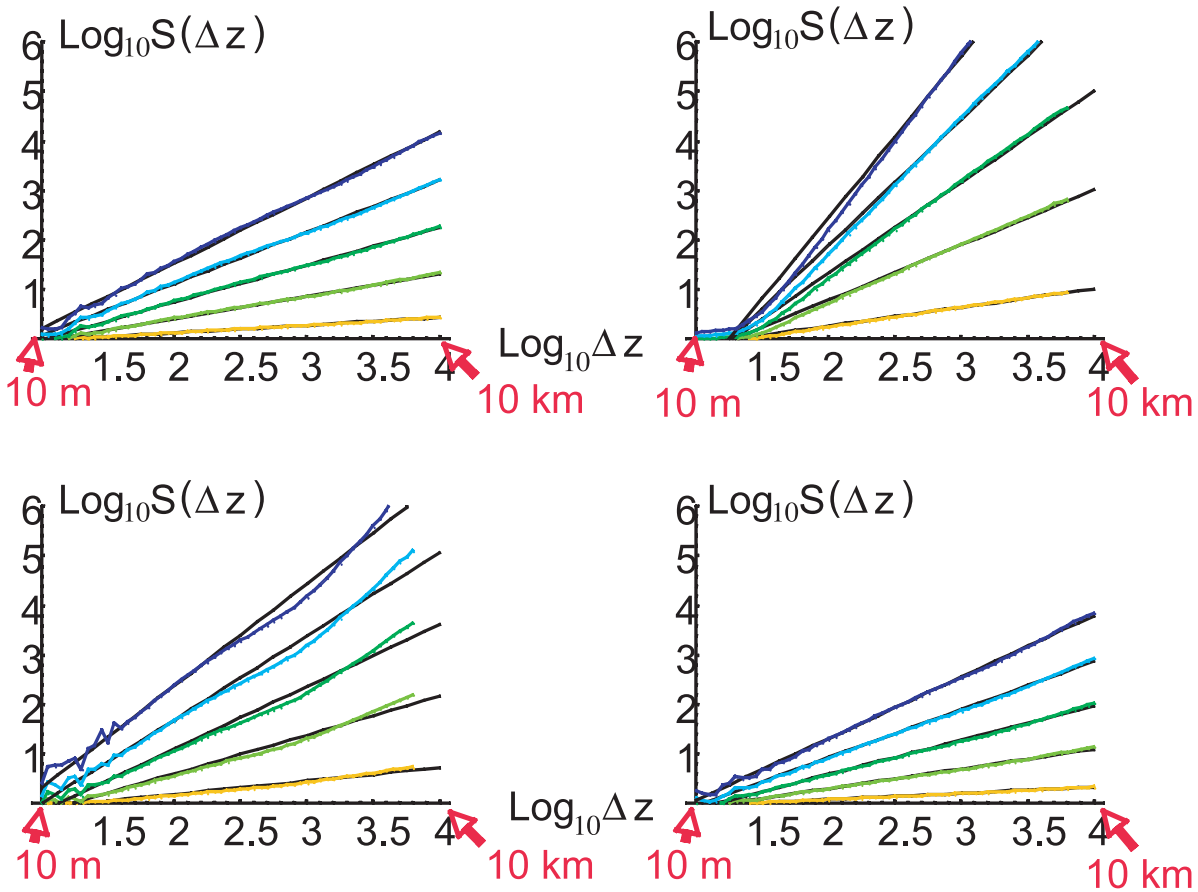

Figure 10a. Nondimensionalized dynamic variables, moments of the fluctuations Δv; $S(\Delta z) = <\Delta v(\Delta z)^q>$. Top to bottom, left to right: v, p, ρ, w_s. Moments q = 0.2, 0.6, 1, 1.4, 1.8.

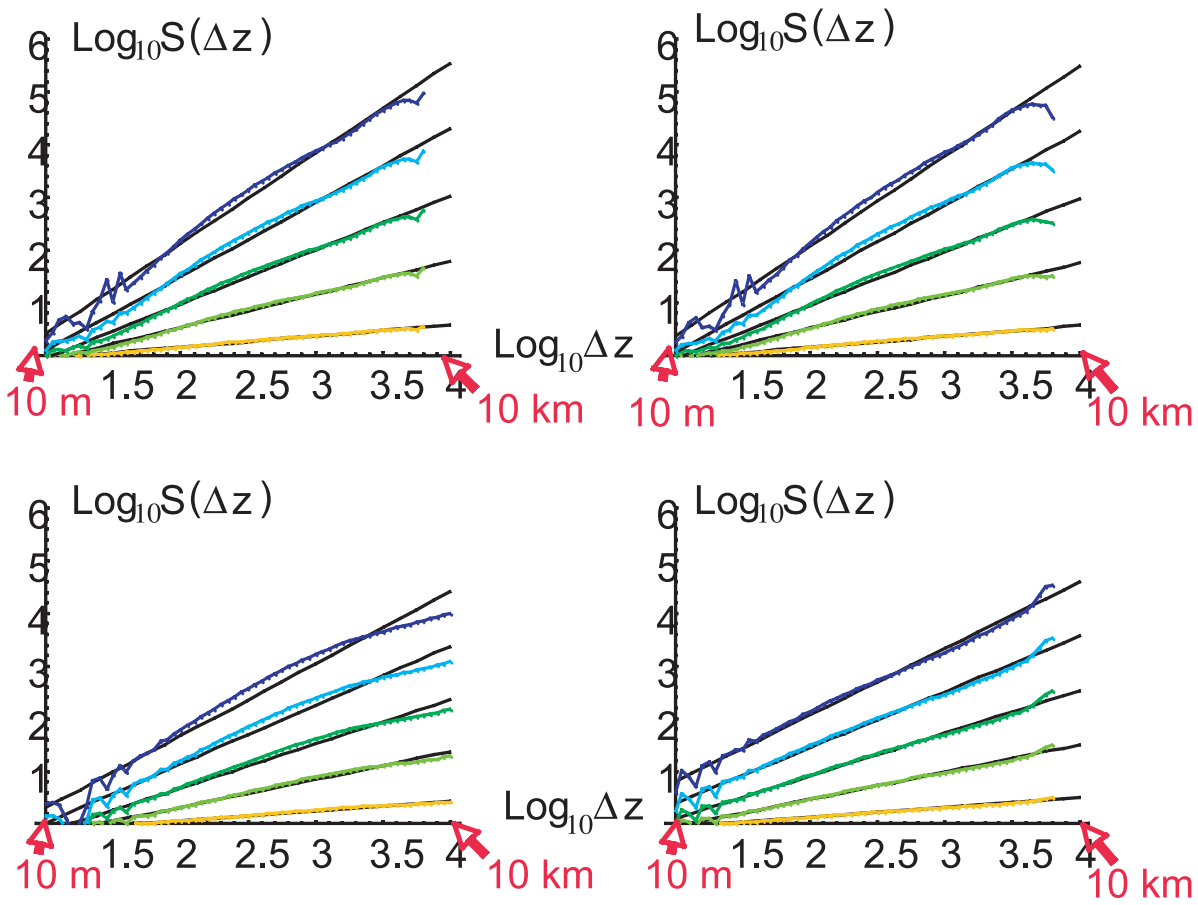

Figure 10b. Same as Figure 10a but for the thermodynamic variables, fluctuation analysis. Top to bottom, left to right: T, log θ, h, log θ_E. Moments q = 0.2, 0.6, 1, 1.4, 1.8.

et al. [2007] estimated the H parameter for the horizontal velocity field by defining the fluctuation $\Delta v(\Delta z)$ as the difference between the velocity at two vertical levels (indexed by n, m) separated by altitude Δz,

$$\Delta v(\Delta z) = |v(z_n) - v(z_m)|; \quad \Delta z = |z_n - z_m|. \tag{19}$$

By examining all the $N(N-1)/2$ observation pairs they avoided interpolations and obtained a robust estimate of the fluctuation statistics. To confirm that the estimates are reliable even in the presence of multifractal outages, we refer the reader to Figure 11a which uses multifractal simulations for a field with $C_1 = 0.05$, $H = 0.6$ combined with the observed z profiles as discussed in section 4.3. From Figure 11a, we see that the scaling is indeed well respected and we recover quite accurately the theoretical parameters: $C_1 = 0.052$, $H = 0.606$. We conclude that the pair method is quite robust even in the presence of large intermittency in the measurements.

[35] Before using the method on the sonde data, we must discuss a technical difficulty. Defining fluctuations as differences is equivalent to using a "poor man's wavelet" and is only valid for $0 < H < 1$. Since the temperature, pressure and potential temperatures are known to be nearly linear with Δz (for the temperature, this is the classical linear adiabatic temperature profile), we must define fluctuations in a way that is valid over a wider range of H values. The basic trick for doing this is to use the differences in the differences (i.e., the second differences). Consider first the case where the z_n are evenly spaced (the subscripts are the integer indexes of the measurements) the second centered difference is

$$\Delta v(\Delta z) = \left| \frac{v(z_{n+m}) + v(z_{n-m})}{2} - v(z_n) \right|; \quad \Delta z = z_{n+m} - z_n = z_n - z_{n-m} \tag{20}$$

(we have divided by 2 since we will only be interested in the scaling properties of the result). Graphically, this definition is equivalent to finding the distance between the point (z_n, $v(z_n)$) and the line joining the points (z_{n+m}, $v(z_{n+m})$), (z_{n-m}, $v(z_{n-m})$). To obtain a fluctuation estimate valid for intermittently spaced z_n, we can simply use the graphical definition (the distance from the line) to obtain the following estimate:

$$\Delta v(\Delta z) = |v(z_{n-m}) + (z_n - z_{n-m})s - v(z_n)|; \quad s = \frac{v(z_{n+m}) - v(z_{n-m})}{z_{n+m} - z_{n-m}}; \quad \Delta z = ((z_n - z_{n-m})(z_{n+m} - z_n))^{1/2}. \tag{21}$$

[36] Since now there are three points needed to define the fluctuation, there is not a unique choice of Δz with which to

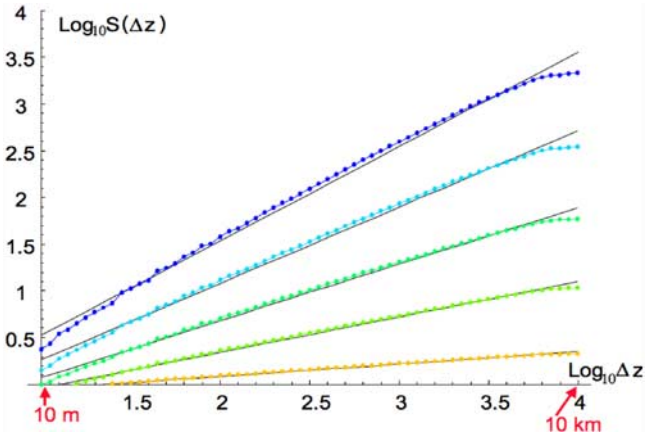

Figure 11a. Simulation with $C_1 = 0.05$, $H = 0.6$ pair method. The parameters recovered using the pair method to determine Δv is $C_1 = 0.052$, $H = 0.606$. Moments $q = 0.2$, 0.6, 1, 1.4, 1.8.

associate the fluctuation. The above choice (the geometric mean) evenly weights the logarithm of the scales and is appropriate for scaling fluctuations. In order to implement this definition it is usually not possible to consider all the triplets of points (this is of order N^3 for series of length N). Our choice was based on the fact that the outages affected only 9.5% of the points so that we considered only the triplets with altitudes z_{n-m}, z_n, z_{n+m} and for each scale we considered all the integer pairs n, m. Note that this choice along with the use of the geometric mean for Δz means that there will be few statistics for the largest factor of 2 in scale. Also, since most of the points are regularly spaced, $z_n - z_{n-m}$ is typically not so different from $z_n - z_{n+m}$ so that the definition does not "mix" different scales too much. To test the technique, we applied to multifractal simulations with $C_1 = 0.05$, $H = 1.5$ again using the observed and highly intermittent z coordinates. Figure 11b shows the result: again the scaling is excellent and the theoretical parameters are recovered quite accurately: the recovered parameters are: $C_1 = 0.056$, $H = 1.46$.

[37] We may now apply the fluctuation analyses to the atmospheric fields whose fluxes were analyzed above. Using definition equation (19) of the fluctuations the results for the qth order statistical moments (the "generalized structure functions" see equation (18)) for the fluctuations the dynamic fields, for the thermodynamic fields are shown in Figures 10a and Figure 10b, respectively. They have been nondimensionalized by dividing by the value $\Delta v(\Delta z)$ for $\Delta z = 10$ m although there is no expectation that the lines converge to a point as for the moments of the fluxes. In Figures 9a and 9b, we used definition equation (19) for the fluctuations of v, h, w_s (since $H < 1$) and definition equation (21) for p, $\log\theta$, $\log\theta_E$, ρ, T (since H is near 1 or larger). In the case of the pressure, we see that $H \approx 2$ so that we should perhaps generalize the definition of fluctuations further so as to obtain a result valid for $H < 3$.

[38] From Figures 10a and 10b, the linearity of the log fluctuation moments versus $\log \Delta z$ is quite striking; indeed, it apparently extends to somewhat larger scales than the scaling of the fluxes shown in Figures 9a and 9b. In order to assess both the quality of the scaling and to intercompare the scaling of the different fields we refer the reader to Tables 2a and 2b, where we have calculated the exponent $H = \xi(1)$ from the mean of the exponents calculated for $\Delta z <$ 300 m and $\Delta z >$ 300 m (300 m is the geometric mean of the observed range 10 m to 10 km). The range indicated by the "±" is half the difference. If the spread indicated in this way is small then the scaling is well respected over the whole range; we see that the H values for p, w_s are particularly well defined whereas for T, $\log\theta$ it is less so. Unlike the cascade analyses of the fluxes, the scaling seems to hold reasonably well over the entire range. In addition, here we made no attempt to investigate the altitude dependence of the exponents which although fairly small, can be systematic and are the subject of work by S. J. Hovde et al. (Vertical scaling of the atmosphere dropsondes from 13 km to the surface, submitted to *Quarterly Journal of the Royal Meteorological Society*, 2009). For example, *Lovejoy et al.* [2007], for the horizontal wind, found that the exponent H systematically increased from 0.60 near the surface (the theoretical Bolgiano-Obukhov value) to ≈0.77 at 12 km. The small altitude dependence of the exponents is partially responsible for the imperfect scaling since the thickest layers necessarily involve points at high altitudes.

[39] In principle we can estimate C_1 and α from the fluctuation exponent $\xi(q)$ (as in work by *Lovejoy et al.* [2007]). To do this we can exploit the equation $\xi(q) = qH - K(q)$ so that $\xi'(1) = H - C_1$ hence $C_1 = \xi(1) - \xi'(1)$. However, for these fields, H is much larger than C_1; indeed, it is often the order of the error in the estimates of H, hence the C_1 estimates should not be too accurate. In work by *Lovejoy et al.* [2007], the C_1 estimates determined this way were more robust since the exponents were calculated altitude by altitude so that the scaling was better and the errors smaller. In contrast, the fluxes do not depend on H so that they can be used to directly estimate $K(q)$ and hence C_1; we therefore consider the flux-based estimate of C_1 to be more accurate.

5. Conclusions

[40] The nature of the vertical stratification is an outstanding problem in atmospheric science: its characteriza-

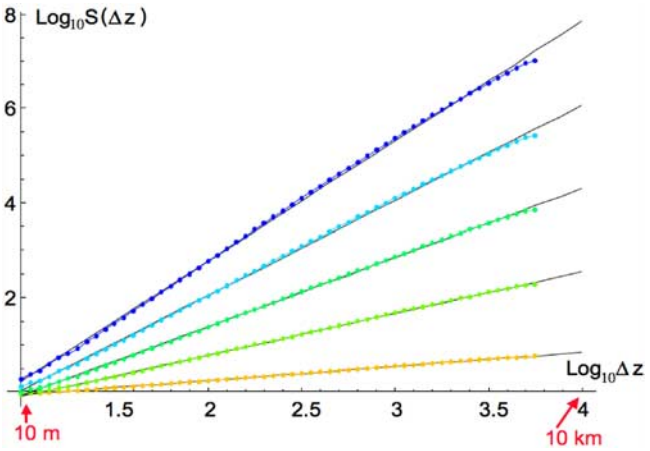

Figure 11b. Simulation with $C_1 = 0.05$, $H = 1.5$ showing triplet method for fluctuations with $H > 1$. The recovered parameters are: $C_1 = 0.056$, $H = 1.46$.

tion has been the subject of numerous theories of dynamical meteorology and its numerical modeling the subject of a series of problematic approximations. Since the 1980s we have argued that the fluctuations of each atmospheric field are scaling but with different exponents in the horizontal and vertical so that the stratification is differential (i.e., scale dependent); it emerges as a consequence of scaling. From the beginning the scaling was postulated to be the result of multiplicative cascade processes concentrating conserved turbulent fluxes scale by scale into smaller and smaller regions of space. However, it has only been recently that the quantitative predictions of these cascade theories have been directly verified in the horizontal at the largest (planetary) scales using large satellite radiance data sets. In addition, analysis of numerical models of the atmosphere (including reanalyses) show that they accurately follow a cascade structure with outer scale typically a bit larger than 20,000 km and scaling accurately holding up to about 5000 km (J. Stolle et al., The stochastic cascade structure of deterministic numerical models of the atmosphere, submitted to *Nonlinear Processes in Geophysics*, 2009). However, the models have very unevenly spaced vertical levels, with high resolutions only near the ground, so that it is nontrivial to study their vertical cascade structures (and these are likely to be quite dependent on the type of numerical approximation used: hydrostatic, anelastic etc.).

[41] In this paper we have attempted to directly confirm the existence of cascade structures in the vertical. We first demonstrated this with a reanalysis of vertical cross sections from lidar pollutant backscatter where we estimated the outer scale of the cascade in the horizontal to be $\approx$25,000 km and $\approx$50 km in the vertical and with different horizontal and vertical exponents. Unfortunately, the key dynamical and thermodynamical meteorological fields cannot be directly remotely sensed; in situ data are necessary. We therefore took advantage of advances in state of the art dropsondes and considered a set with over 200 sondes over the Pacific Ocean. Although the sondes have nominal resolutions of 0.5 s (corresponding to about 5–10 m in the vertical), they have such frequent outages that the intervals between measurements (in both time and in the vertical) were highly intermittent. To our surprise we found that the outages themselves were quite accurately multifractal. This means that the temporal and spatial resolutions of the series are not fixed but are highly variable. Therefore, special care is needed in analyzing the results, in particular, the use of interpolation to produce regularly spaced series can lead to large errors (particularly in estimating spectra).

[42] In previous publications, the fluctuations were analyzed without interpolation but using the differences between measurements at the observed locations, so there was no overall statistical bias due to the variable resolution. However, in order to study the cascade structure, estimates of the turbulent fluxes are needed over a wide range of resolutions and these are most easily estimated by systematically degrading high-resolution fluxes by integration. We therefore developed a way of estimating the fluxes at variable resolution and then statistically correcting for the resulting biases. The bias correction turned out to be quite simple: for universal multifractals, it amounted to multiplication of the moment scaling exponent $K(q)$ by a constant that depended on the minimum singularity present in the sampling intervals.

[43] The results convincingly showed the wide range scaling of the fluxes and displayed the characteristic cascade "signature" of lines converging to the effective outer cascade scale that we estimated in the range 1–50 km depending on the field: the outer scale was larger for fields related to the humidity and smallest for the horizontal wind ($\approx$1.3 km) with deviations from theoretical cascade behavior of the order of $\pm$1–2%. However, the scaling of the fluctuations was found to be better than that of the fluxes and was very good over the whole range up to 10 km although there were some systematic variations with altitude with both the outer scale and the intermittency parameter increasing with altitude.

[44] This study was motivated by the need to directly check the predictions of cascade models in the vertical direction; complete characterization of the atmosphere would involve a more systematic study of the variations of the cascades with altitude and will be the subject of a future paper. In addition, comparisons with the horizontal cascade structure are needed in order to quantify the degree of scale by scale stratification for each field (their "elliptical dimensions"). Unfortunately, for the fields examined here, the corresponding horizontal analyses have generally not yet been made, partly because of the difficulty in interpreting in situ aircraft measurements which are greatly affected by the nature of the turbulence that they measure [*Lovejoy et al.*, 2004, 2009b; *Lilley et al.*, 2008].

[45] **Acknowledgments.** We thank the NOAA G4 team. This research did not benefit from any specific funding.

References

Adelfang, S. I. (1971), On the relation between wind shears over various intervals, *J. Atmos. Sci.*, *10*, 138.

Dalaudier, F., C. Sidi, M. Crochet, and J. Vernin (1994), Direct evidence of "sheets" in the atmospheric temperature field, *J. Atmos. Sci.*, *51*, 237–248, doi:10.1175/1520-0469(1994)051<0237:DEOITA>2.0.CO;2.

Dewan, E. (1997), Saturated-cascade similitude theory of gravity wave spectra, *J. Geophys. Res.*, *102*, 29,799–29,817, doi:10.1029/97JD02151.

Dewan, E., and R. Good (1986), Saturation and the "universal" spectrum vertical profiles of horizontal scalar winds in the stratosphere, *J. Geophys. Res.*, *91*, 2742–2748, doi:10.1029/JD091iD02p02742.

Gardner, C. (1994), Diffusive filtering theory of gravity wave spectra in the atmosphere, *J. Geophys. Res.*, *99*, 20,601–20,622, doi:10.1029/94JD00819.

Gardner, C. S., C. A. Hostetler, and S. J. Franke (1993), Gravity wave models for the horizontal wave number spectra of atmospheric velocity and density fluctuations, *J. Geophys. Res.*, *98*, 1035–1049, doi:10.1029/92JD02051.

Gardner, C., X. Tao, and G. Papen (1995), Simultaneous lidar observations of vertical wind, temperature and density profiles in the upper atmosphere: Evidence of nonseparability of atmospheric perturbation spectra, *Geophys. Res. Lett.*, *22*, 2877–2880, doi:10.1029/95GL02783.

Hock, T., and J. Franklin (1999), The NCAR GPS dropsonde, *Bull. Am. Meteorol. Soc.*, *80*, 407–420.

Lavallée, D., S. Lovejoy, D. Schertzer, and P. Ladoy (1993), Nonlinear variability and landscape topography: Analysis and simulation, in *Fractals in Geography*, edited by L. De Cola and N. Lam, pp. 171–205, Prentice-Hall, Englewood, N. J.

Lazarev, A., D. Schertzer, S. Lovejoy, and Y. Chigirinskaya (1994), Unified multifractal atmospheric dynamics tested in the tropics: Part II. Vertical scaling and generalized scale invariance, *Nonlinear Processes Geophys.*, *1*, 115–123.

Lilley, M., S. Lovejoy, K. Strawbridge, and D. Schertzer (2004), 23/9 dimensional anisotropic scaling of passive admixtures using lidar aerosol data, *Phys. Rev. E*, *70*, 036307-1-7.

Lilley, M., S. Lovejoy, K. B. Strawbridge, D. Schertzer, and A. Radkevitch (2008), Scaling turbulent atmospheric stratification: II. Empirical study of

the stratification of the intermittency, *Q. J. R. Meteorol. Soc.*, *134*, 301–315, doi:10.1002/qj.202.

Lovejoy, S., D. Schertzer, and A. F. Tuck (2004), Fractal aircraft trajectories and nonclassical turbulent exponents, *Phys. Rev. E*, *70*, 036306-1-5.

Lovejoy, S., A. F. Tuck, S. J. Hovde, and D. Schertzer (2007), Is isotropic turbulence relevant in the atmosphere?, *Geophys. Res. Lett.*, *34*, L15802, doi:10.1029/2007GL029359.

Lovejoy, S., D. Schertzer, and V. Allaire (2008a), The remarkable wide range spatial scaling of TRMM precipitation, *Atmos. Res.*, *90*, 10–32, doi:10.1016/j.atmosres.2008.02.016.

Lovejoy, S., D. Schertzer, M. Lilley, K. B. Strawbridge, and A. Radkevitch (2008b), Scaling turbulent atmospheric stratification: I. Turbulence and waves, *Q. J. R. Meteorol. Soc.*, *134*, 277–300, doi:10.1002/qj.201.

Lovejoy, S., A. F. Tuck, S. J. Hovde, and D. Schertzer (2008c), Do stable atmospheric layers exist?, *Geophys. Res. Lett.*, *35*, L01802, doi:10.1029/2007GL032122.

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

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

Mandelbrot, B. B. (1974), Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier, *J. Fluid Mech.*, *62*, 331–350, doi:10.1017/S0022112074000711.

Muschinski, A., and C. Wode (1998), First in situ evidence for coexisting submeter temperature and humidity sheets in the lower free troposphere, *J. Atmos. Sci.*, *55*, 2893–2906, doi:10.1175/1520-0469(1998)055<2893:FISEFC>2.0.CO;2.

Novikov, E. A., and R. Stewart (1964), Intermittency of turbulence and spectrum of fluctuations in energy-dissipation, *Izv. Akad. Nauk. SSSR, Ser. Geofiz.*, *3*, 408–412.

Parisi, G., and U. Frisch (1985), A multifractal model of intermittency, in *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, edited by M. Ghil, R. Benzi, and G. Parisi, pp. 84–88, North Holland, Amsterdam.

Radkevitch, A., S. Lovejoy, K. Strawbridge, and D. Schertzer (2007), The elliptical dimension of space-time atmospheric stratification of passive admixtures using lidar data, *Physica A*, *382*, 597–615, doi:10.1016/j.physa.2007.03.046.

Radkevitch, A., S. Lovejoy, K. B. Strawbridge, D. Schertzer, and M. Lilley (2008), Scaling turbulent atmospheric stratification: III. Empirical study of space-time stratification of passive scalars using lidar data, *Q. J. R. Meteorol. Soc.*, *134*, 317–335, doi:10.1002/qj.203.

Schertzer, D., and S. Lovejoy (1985), The dimension and intermittency of atmospheric dynamics, in *Turbulent Shear Flow 4*, edited by B. Launder, pp. 7–33, Springer, New York.

Strauss, D. M., and P. Ditlevsen (1999), Two-dimensional turbulence properties of the ECMWF reanalyses, *Tellus, Ser. A*, *51*, 749–772.

Tuck, A. F. (2008), *Atmospheric Turbulence: A Molecular Dynamics Perspective*, Oxford Univ. Press, Oxford, U.K.

Van Zandt, T. E. (1982), A universal spectrum of buoyancy waves in the atmosphere, *Geophys. Res. Lett.*, *9*, 575–578, doi:10.1029/GL009i005p00575.

Yaglom, A. M. (1966), The influence on the fluctuation in energy dissipation on the shape of turbulent characteristics in the inertial interval, *Sov. Phys. Dokl., Engl. Transl.*, *2*, 26–30.

S. J. Hovde and A. F. Tuck, Chemical Sciences Division, NOAA Earth System Research Laboratory, 325 Broadway, Boulder, CO 80305-3337, USA. (dr.adrian.tuck@sciencespectrum.co.uk)

S. Lovejoy, Department of Physics, McGill University, 3600 University Street, Montreal, QC H3A 2T8, Canada. (lovejoy@physics.mcgill.ca)

D. Schertzer, CEREVE, Université Paris Est, 77455 Marne-la-Vallee CEDEX 2, France. (Daniel.Schertzer@cereve.enpc.fr)