Scaling turbulent atmospheric stratification. I: Turbulence and waves

S. Lovejoy,[a,b]* D. Schertzer,[c,f] M. Lilley,[a,e] K. B. Strawbridge[d] and A. Radkevich[a,g]

[a] *Department of Physics, McGill University, Montreal, Canada*
[b] *Centre GEOTOP UQAM/McGill, Université du Québec, Montréal, Canada*
[c] *Université Paris-Est ENPC/CEREVE, Marne-la-Vallée, France*
[d] *Centre for Atmospheric Research Experiments, Science and Technology Branch, Environment Canada, Egbert, Canada*
[e] *GreCo, Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014 Paris, France*
[f] *Météo France, 1 Quai Branly, Paris 75007, France*
[g] *Canadian Centre for Remote Sensing, 588 Booth Street, Ottawa, Ontario, Canada*

ABSTRACT: In this first of a three-part series, we argue that the dynamics of turbulence in a stratified atmosphere should depend on the buoyancy over a wide range of vertical scales and on energy flux over a wide range of horizontal scales; it should be scaling, but anisotropic, not isotropic. We compare the leading statistical theories of atmospheric stratification which are conveniently distinguished by the elliptical dimension D_s which quantifies their degree of spatial stratification. This includes the mainstream isotropic 2-D (large scales), isotropic 3-D (small scales) theory but also the more recent linear gravity wave theories ($D_s = 7/3$) and the classical fractionally integrated flux (FIF) 23/9-D unified scaling model. In the latter, the horizontal wind has a $k^{-5/3}$ spectrum as a function of horizontal wavenumber determined by the energy flux and a $k^{-11/5}$ energy spectrum as a function of vertical wavenumber determined by the buoyancy force variance flux. In this model, the physically important notion of scale is determined by the turbulent dynamics, it is not given *a priori* (i.e. the by usual Euclidean distance). The 23/9-D FIF model is the most physically and empirically satisfying, being based on turbulent (spectral) fluxes. The FIF model as originally proposed by Schertzer and Lovejoy is actually a vast family of scaling models broadly compatible with turbulent phenomenology and with the classical turbulent laws of Kolmogorov, Corrsin and Obukov. However, until now it has mostly been developed on the basis of structures localized in space–time. In this paper, we show how to construct extreme FIF models with wave-like structures which are localized in space but unlocalized in space–time, as well as a continuous family of intermediate models which are akin to Lumley–Shur models in which some part of the localized turbulent energy 'leaks' into unlocalized waves.

The key point is that the FIF requires two propagators (space–time Green's functions) which can be somewhat different. The first determines the space–time structure of the cascade of fluxes; this must be localized in space–time in order to satisfy the usual turbulence phenomenology. In contrast, the second propagator relates the turbulent fluxes to the observables; although the spatial part of the propagator is localized as before, in space–time it can be unlocalized. (It is still localized in space, now in wave packets.) We display numerical simulations which demonstrate the requisite (anisotropic, multifractal) statistical properties as well as wave-like phenomenologies. In parts II and III we will examine the empirical evidence for the spatial and temporal parts, respectively, of the model using state-of-the-art lidar data of aerosol backscatter ratios (which we use as a surrogate for passive scalar concentration). Copyright © 2008 Royal Meteorological Society

KEY WORDS   Fractal; multifractal; turbulence; atmospheric spectrum; lidar; passive scalar

Received 9 December 2005; Revised 10 October 2007; Accepted 23 November 2007

1. Introduction

The modern view of the atmosphere is that of a turbulent hierarchy of interacting structures covering a wide range of scales. This picture largely originated in the pioneering work of Richardson (1922). Although primarily concerned with laying the foundations of numerical forecasting, Richardson's book also contains a note in the form of a poem where he proposes that the basic dynamic mechanism in the atmosphere is a cascade of eddies passing their energy from large to small scales eventually undergoing viscous dissipation. Since the atmosphere has a very large aspect ratio (20 000 km/10 km; horizontal/vertical), we see that Richardson's cascade picture immediately poses the question as to the nature of the stratification and how it should be incorporated into the cascade.

Since Richardson, fluid stratification has been modelled in several ways. Perhaps the most familiar is the dynamical meteorological approach which considers

---

*Correspondence to: S. Lovejoy, Department of Physics, McGill University, 3600 University St., Montreal, Quebec, Canada.
E-mail: lovejoy@physics.mcgill.ca

a homogeneously stratified atmosphere with constant Brunt–Väisälä frequency (i.e. with a constant gradient of potential temperature). Although this homogeneity assumption is an unrealistic restriction on the dynamically significant range of scales, it is frequently used to give qualitative insight into the general effects of stratification, and it is sometimes used to interpret specific flows. In comparison, the 'Boussinesq' approximation is less restrictive: it postulates the existence of a well-defined 'reference' vertical density profile $\rho(z)$ so that the buoyancy forces on a fluid particle are determined by the difference between the particle's density and $\rho(z)$ rather than on the local difference with the surrounding fluid. This allows one to consider the small scales as isotropic fluctuations about an anisotropic large-scale reference state. Other approaches attempt to model only the large scales by assuming from the outset that they are completely stratified (two dimensional, flat). The shallow-water equations, the quasi-geostrophic and barotropic approximations are commonly used in large-scale models of this type.

These and other assumptions or approximations can be used as the basis for treating stratification in traditional turbulence approaches, i.e. statistical theories which attempt to consider the dynamics of structures spanning wide ranges of scale. Because the basic (e.g. Navier–Stokes) equations have no characteristic lengths (except the small – millimetric – dissipation scale, and the outer horizontal scale – the planet – and outer vertical scale, the atmospheric thickness), they classically admit isotropic scaling ('self-similar') solutions. Turbulence approaches to atmospheric dynamics thus commonly break up the dynamically active range (planetary to dissipation scales) into subranges in which various scaling (power law) behaviours are supposedly dominant.

To date, the great majority of turbulence theories have postulated *a priori* that all the relevant regimes are isotropic. They thus require at least two regimes to model the atmosphere: a (quasi) two-dimensional isotropic large scale and (quasi) three-dimensional isotropic small scale. Since the scale height H_s for the mean pressure is about 7.5 km, the 'dimensional transition' from isotropic 2-D to isotropic 3-D turbulence must occur somewhere in the mesoscale; this is the origin of the elusive 'mesoscale gap' in the energy spectrum which we discuss below. The main exceptions to isotropic scaling ('self-similar') theories are the quasi-linear gravity wave theories and the 23/9-D anisotropic 'unified scaling' model proposed by Schertzer and Lovejoy (1985a,b). Stochastic multifractal models following this symmetry and obeying the multifractal extensions of the Kolmogorov and Corrsin–Obukov statistics were proposed by Schertzer and Lovejoy (1987); the Fractionally Integrated Flux (FIF) model. These involve the notion of scaling stratification and are the main focus of this three-part paper.

To understand the meaning of scaling stratification, consider the dimension which characterizes the stratification. In a 2-D atmosphere, there is only variability in the horizontal direction. In a 3-D atmosphere,

the variability is isotropic – on average (indicated by $\langle \ldots \rangle$) then $\langle |\Delta\rho(\Delta x, 0, 0)| \rangle = \langle |\Delta\rho(0, 0, \Delta z)| \rangle$ if $\Delta x = \Delta z$ (x, y are horizontal, z is a vertical coordinate). In an intermediate ($2 < D_s < 3$) stratified but still scaling case, we need only go a distance $(\Delta z/l_s) = (\Delta x/l_s)^{H_z}$ with $0 < H_z < 1$ to find that $\langle |\Delta\rho(\Delta x, 0, 0)| \rangle = \langle |\Delta\rho(0, 0, \Delta z)| \rangle$. We see that $H_z = 1$ corresponds to 3-D isotropy and $H_z = 0$ to 2-D isotropy. Since the exponents refer to the way typical structures change their volumes with scale, $H_z = 0$, $D_s = 2$ refers to structures whose vertical extent is statistically independent of their horizontal extents. This includes both the usual 2-D isotropic turbulence (with essentially no variation in the z direction so that $\Delta\rho(0, 0, \Delta z) = \rho(x, y, z + \Delta z) - \rho(x, y, z) \approx 0$, independent of Δz), as well as the case of flat statistically independent 2-D structures stacked on top of each other in thin layers (so that $\langle |\Delta\rho(0, 0, \Delta z)| \rangle$ is independent of Δz). For $1 > H_z > 0$, the dimensional parameter l_s – the 'sphero-scale' – is the scale at which going a distance $\Delta x = \Delta z = l_s$ yields the same fluctuations: $\langle |\Delta\rho(l_s, 0, 0)| \rangle = \langle |\Delta\rho(0, 0, l_s)| \rangle$. In scaling stratification, the aspect ratio of structures in vertical cross-sections therefore varies as a power law of scale; in addition (assuming horizontal isotropy, i.e. if $\Delta y = \Delta x$ then $\langle |\Delta\rho(\Delta x, 0, 0)| \rangle = \langle |\Delta\rho(0, \Delta y, 0)| \rangle$), the volume of typical structures varies as $\Delta x \Delta x \Delta x^{H_z} = \Delta x^{D_s}$ with $D_s = 2 + H_z$. The intermediate dimension D_s which quantifies how the structures change their volumes with horizontal scale, is called an 'elliptical dimension' because of the typical elliptical shapes of the vertical sections of the average structures. Note that the notion of dimension can also be used in a rather different sense to characterize the intermittency of this stratified turbulence. This different – multifractal – meaning is discussed in section 3 and in parts II (Lilley *et al.*, 2008) and III (Radkevich *et al.*, 2008) of this series of papers. As we shall see below, the proposal by Schertzer and Lovejoy (1983, 1985a, 1985b) that horizontal structures are dominated by energy fluxes, while vertical structures are dominated by buoyancy variance fluxes, implies $H_z = 5/9$ and hence $D_s = 23/9$ (the subscript s indicates 'space'; below we consider the extension to space–time). Here and throughout, the term 'fluxes' refers to Fourier space fluxes, i.e. scale to scale fluxes; 'conserved fluxes' means the statistical constancy with scale (for canonical and micro-canonical conservation, see section 4). These fluxes are only equal to the corresponding dissipations at the small dissipation scales.

Since each atmospheric model implies specific elliptical dimensions D_s (or equivalently, H_z) over different ranges, it ought to be straightforward to empirically test them simply by measuring D_s (or H_z) over the relevant ranges. The difficulty has been that, until recently, tests have primarily been made using either aircraft wind data in the horizontal or balloon wind data in the vertical (the exception is Lilley *et al.* (2004); lidar vertical cross-sections, see below and parts II, III). The results from separate experiments, often from different parts of the world and under different conditions,

could only be compared in an indirect way (with the partial exception of Chigirinskaya *et al.* (1994), Lazarev *et al.* (1994)). An additional problem is that aircraft do not fly in perfectly flat trajectories so that *in situ* wind velocity, temperature or density measurements made with such means can only be made over irregular trajectories. Indeed, it has only recently been discovered (Lovejoy *et al.*, 2004) that – precisely due to non-2-D turbulence – at least stratospheric aircraft can follow fractal trajectories, i.e. can be biased with respect to linear trajectories over large distances. (The fractality is of course cut off by aircraft inertia at scales less than about 1 km below which they become smooth; $D = 1$). Therefore, such *in situ* data can yield spurious statistical exponents, spurious scaling breaks and erroneous interpretations. Finally, huge amounts of data are needed in order to average over the large fluctuations so as to obtain accurate results. These and related empirical issues are the focus of part II, while part III extends this to the time domain.

In spite of several unsatisfactory theoretical assumptions – especially the coexistence of weakly nonlinear waves with strongly nonlinear turbulence driving – the gravity wave theories have the advantage of being close to the wave phenomenology commonly observed in the atmosphere. In contrast, the initial implementation of the main turbulence-based anisotropic alternative – the Fractionally Integrated Flux (FIF) model (Schertzer and Lovejoy, 1987; Schertzer *et al.*, 1997a) – which we will call 'classical FIF' yields realizations clearly missing wave-like structures. Figure 1 shows a vertical section of lidar data (described in part II), and a classical FIF simulation (with the observed scaling and multifractal parameters). Although the classical FIF well reproduces the spectra in the horizontal, vertical and time (and higher-order statistics as well), it is visually imperfect. However FIF is a very general scaling multifractal framework for turbulence dynamics. Below, we show how to use it to produce a continuum of models ranging from an extreme unlocalized ('wave-like') FIF model to the classical localized FIF model.

In outline, the spatial part of the FIF model is based on a physical scale function which defines the notion of scale in the stratified horizontal–vertical space. Whereas classically, scale is imposed *a priori* as the usual (Euclidean) distance, the idea here is that the turbulent dynamics – energy and buoyancy variance (spectral) fluxes – determine the dynamically appropriate notion of scale. This is somewhat analogous to general relativity where the distribution of matter and energy determine the metric. From the physical scale function, the FIF then introduces two propagators (space–time Green's functions), each a (generalized, anisotropic) power law; hence the term 'Fractional Integration' (e.g. Miller and Ross, 1993). The first propagator determines the structure of the turbulent cascade; it must be localized in space–time in order to be compatible with the usual phenomenology of turbulent fluxes. The second propagator links the
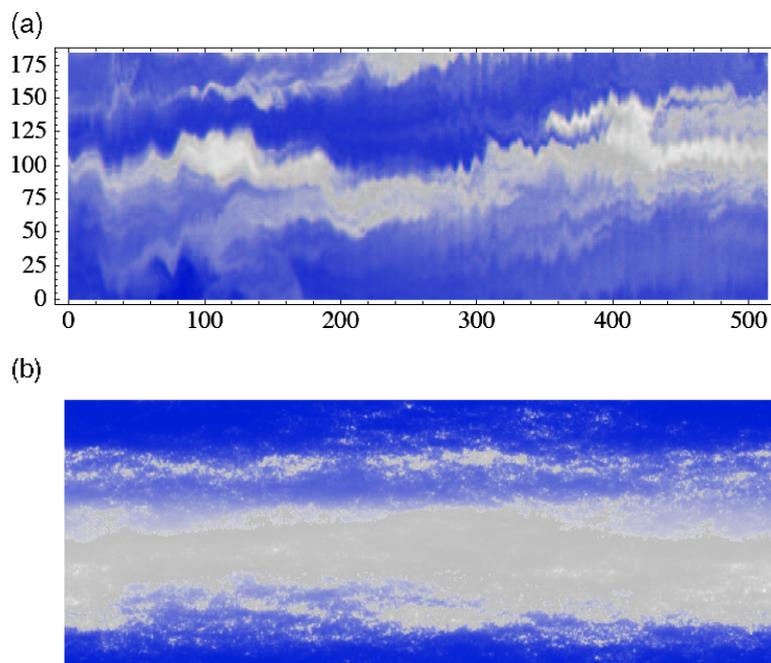

Figure 1. (a) Vertical atmospheric section of lidar backscatter from a passive scalar surrogate: background urban aerosol pollution near Vancouver, British Columbia. This is one of the datasets described and analyzed in part II, with 3 m resolution vertical and 96 m horizontal; $D_s = 2.55 \pm 0.02$. Numbers are pixels; there is a 32 : 1 aspect ratio. Note the wave-like undulations. (b) A simulated vertical cross-section of the 23/9-D Fractionally Integrated Flux model for a passive scalar density (false colours; see details in section 4). This is the FIF model without wave-like fractional integration; it has the same sphero-scale and same differential anisotropy as the data in (a), but lacks the waves. This model has the same multifractal statististical parameters as the data analyzed in part II (universal multifractals Levy index $\alpha = 1.8$, $C_1 = 0.1$, see section 4). For many more simulations (including some in three spatial dimensions and with radiative transfer), see http://www.physics.mcgill.ca/~gang/Lovejoy.htm.

See also the simulations in Lovejoy and Schertzer (2007). This figure is available in colour online at www.interscience.wiley.com/qj

turbulent flux to the observable (e.g. velocity or passive scalar) field. In the classical FIF, the two propagators were essentially the same. Below, we show how this classical FIF can be modified so that the second propagator – although based on the same physical scale function – is wave-like in space-time. Indeed, it allows the classical stationary phase method of asymptotic approximation, so that we can analyse the behaviour in the usual way in terms of wave packets, group velocities, etc. In addition, we show how specific gravito-turbulent dispersion relations can be chosen which are very close to the classical gravity wave dispersion relations. Our model thus leads to many predictions similar to the standard theory yet it is strongly nonlinear, and requires neither scale separation nor linear perturbation analysis. (The waves and turbulence are inseparable; it is a 'tubulence/wave' model.) The overall result is statistically very close to the observations in parts II, III (including the strong intermittency).

The remainder of this paper is organized as follows. In section 2, we discuss some particularly striking recent evidence for the basic wide-range scaling model, and the mainstream gravity wave anisotropic scaling model. In section 3, we review the classical FIF model. In section 4, we outline the turbulence/wave model and discuss the effects of advection. In order to keep the paper focused, several of the technical derivations are given in appendices. Appendix A discusses the effect of intermittency on the scale function and the elliptical dimension, appendix B the effective temporal scaling introduced by the vertical wind, and appendix C gives a few statistical properties of the turbulence/wave model.

2. Wide-range anisotropic scaling in the atmosphere

2.1. Does the atmosphere (even approximately) obey the Kolmogorov law?

The theoretical paradigm of 3D isotropic turbulence was originally developed as a way of simplifying the problem of fully developed dynamically forced turbulence and was initially applied to laboratory isotropic grid-turbulence. The key result is the famous Kolmogorov (1941) $k^{-\beta}$ spectrum for the horizontal velocity with spectral exponent $\beta = 5/3$; in real space, this implies $|\Delta v| \approx |\Delta \mathbf{r}|^H$, where $|\Delta v|$ is a typical (absolute) velocity difference, $\Delta \mathbf{r}$ is a displacement vector over which the difference is measured, and the scaling exponent $H = (\beta - 1)/2 = 1/3$. Since the basic exponent is determined by dimensional analysis on the energy flux (which is conserved by the nonlinear terms of the Navier–Stokes equations), this result is insensitive to the details of the driving mechanism (taking intermittency into account does not change H, but β is changed a little, see section 3.3). However, fundamentally the atmosphere is driven via incoming solar radiation, and Bolgiano (1959) and Obukhov (1959) pointed out that (within the Boussinesq approximation) the buoyancy variance flux, ϕ, is a new quadratic invariant, with dimensions length^2/time^5. They

argued that this ϕ should be dynamically significant in the atmosphere giving rise to an isotropic 'buoyancy subrange' with an isotropic $k^{-11/5}$ spectrum (i.e. $|\Delta v| \approx |\Delta \mathbf{r}|^{H_v}$; $H_v = 3/5$), which would dominate the energy flux (Kolmogorov) subrange for scales larger than the Bolgiano scale l_B (see below), itself estimated to be of the order of metres. In the following years, this theory was largely discarded because of the failure to empirically detect this isotropic range either in time or in the horizontal direction. By the time the vertical $k_z^{-11/5}$ spectrum was finally observed (Endlich, 1969; Adelfang, 1971) it was almost forgotten. It was not until the 1980s that the basic idea was revived in an anisotropic form as part of the 23/9-D model (see section 3 below).

In part II, we examine the literature for the evidence about Bolgiano–Obukhov (BO) scaling of the horizontal wind in the vertical and Kolmogorov scaling in the horizontal. Although as indicated, this review generally favours the BO scaling in the vertical and Kolmogorov scaling in the horizontal, but until recently the data were of too low a quality to give a clear result. However, in a recent paper Lovejoy *et al.* (2007) argued on the basis of 235 state-of-the-art dropsonde vertical profiles that, over vertical scales ranging from 5 m to >10 km and even through the lowest 158 m thick layer, H_v is near the BO value 3/5, increasing slightly with altitude to a value of about 0.75 (apparently influenced by the presence of strong altitude-dependent winds, i.e. jet streams). In other words, the Kolmogorov exponent $H = 1/3$ is (virtually) never observed in the vertical. Dropsondes and stratospheric balloons extend these scaling results to virtually all the other atmospheric fields (S. J. Hovde 2007, personal communication). There thus seem to be few opportunities to apply Kolmogorov isotropic theory.

2.2. Wide-range horizontal scaling

If the vertical scaling exponent H_v were the same as the horizontal exponent H_h – i.e. if the turbulence were isotropic 3-D – it would not be possible for the horizontal scaling to extend much further than the scale height ≈ 7.5 km; this is the problem of the mesoscale gap alluded to earlier. However, since in the vertical we find BO scaling ($H_v \approx 3/5$), then if $H_h < 3/5$, structures will be horizontally stratified at larger and larger scales and it will be possible for the horizontal scaling laws to extend much further than would otherwise be possible. As discussed in part II, there is indeed much reason to believe that the Kolomogorov scaling extends over huge ranges, however due to the difficultes in interpreting *in situ* (aircraft) data, the question still cannot be answered with complete assurance. However, if we turn our attention to remotely sensed data, the situation is quite different. For example, many satellite radiances are strongly coupled to cloud and rain fields which are in turn strongly coupled to the dynamical fields. Since scale invariance is a symmetry principle, we may conclude that a break in horizontal scaling in any one of these fields would be reflected in the others. Even

though we cannot use this indirect approach to estimate the value of the horizontal wind exponent, we can use it to check if basic scaling symmetries are satisfied. This is the basic approach used by Lovejoy *et al.* (2001) and Lovejoy and Schertzer (2006) who used nearly 1000 satellite images in the visible and infrared. While the deviations from scaling were very small, most of the direct results were confined to scales <280 km and the outer scale was estimated by extrapolating to planetary scales. More recently, this limitation was extended to the full planetary scale (20 000 km) using satellite radar reflectivities and visible, infrared and passive microwave radiances from the Tropical Rainfall Monitoring Mission (TRMM) satellite. Lovejoy *et al.* (2008a) and Lovejoy and Schertzer (2008) find that, with the exception of very low-order moments dominated by (nominally) zero-rain reflectivity, Z, the multiscaling holds remarkably well. For example, over the directly observed range 4 to 20 000 km, the moments $<Z_l^q> \approx l^{-K(q)}$ for $0 < q < 2$ predicted by two-parameter universal multifractal were followed to within an average deviation of $\pm 4.6\%$, where l is the resolution of the reflectivities. Figure 2 shows a spectral analysis of the 2.2 km resolution visible and infrared channels, again showing remarkably accurate scaling very close to that theoretically predicted for passive scalars. In Lovejoy *et al.* (2008b), it is shown that these radiances as well as those from from passive microwave channels have gradients whose statistics are within about $\pm 1\%$ of those predicted by mulitfractal cascade models with external cascade scales ranging from about 5 000 to 18 000 km depending on the wavelength.

2.3. The mainstream stratified scaling model with $D_s = 7/3$: quasi-linear gravity waves

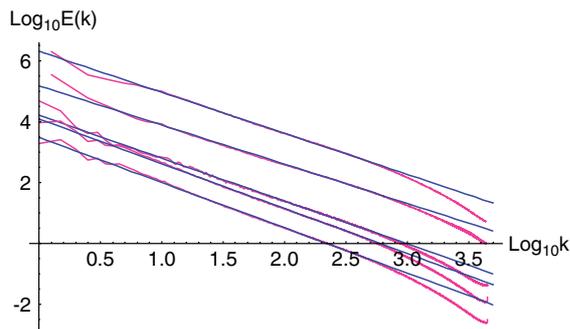The previous two subsections argue that the atmosphere has wide-range scaling in both horizontal and vertical

Figure 2. Spectra from ∼1000 orbits of the TRMM satellite (the Visible and Infrared Scanner) channels 1–5 (at wavelengths 0.630, 1.60, 3.75, 10.8, 12.0 m, from top to bottom, displaced in the vertical for clarity), with nominal resolution 2.2 km. The (straight) regression lines have spectral exponents $\beta = 1.35, 1.29, 1.41, 1.47, 1.49$ respectively, close to the value $\beta = 1.53$ corresponding to the spectrum of passive scalars (= 5/3 minus intermittency corrections). The units are such that $k = 1$ is the wavenumber corresponding to the size of the planet (20 000 km)$^{-1}$. Channels 1, 2 are reflected solar radiation so that only the 15 600 km sections of orbits with maximum solar radiation were used. The high-wavenumber fall-off is due to the finite resolution of the instruments. (We thank S. King for help with the data analysis). This figure is available in colour online at www.interscience.wiley.com/qj

directions implying that the stratification is scaling. Starting with VanZandt (1982), this has indeed become the dominant view in the experimental community. (Part II gives a more detailed review.) VanZandt (1982) innovated in two important respects: (a) he empirically recognized the horizontal/vertical anisotropy of the scaling in the mesoscale and (b) he postulated that classical dispersion relations (based on linear perturbation theory) generally existed between the frequency and horizontal and vertical wavenumbers. The consequence of both assumptions was that fluctuations in the horizontal wind could be described by a 'universal' anisotropic spectrum. This picture differed from previous isotropic models (e.g. Bolgiano, 1959; Lumley, 1964; Weinstock, 1978) because it did not involve a transition from an isotropic inertial subrange to a buoyancy subrange. In addition, the energy source was not explicit; it was not necessarily turbulent 'leakage' as in the Lumley (1964), and Shur (1962) models. Since VanZandt (1982), almost all empirical observations of the horizontal wind and temperature (and for the upper atmosphere, for the density) have assumed different spectral exponents in the horizontal and vertical directions; they all assume $2 < D_s < 3$.

Currently, there are two main gravity wave theories used in the literature; the Saturated Cascade Theory (SCT; Dewan and Good, 1986; Dewan, 1997) and the Diffusive Filtering Theory (DFT; Gardner, 1994). They both share the key assumptions of Van Zandt (1982) about anisotropic scaling and linear gravity wave dispersion relations. Both assume the validity of the classical linear perturbation gravity wave dispersion relation:

$$\omega = \frac{Nk_x}{|\mathbf{k}|}; \quad |\mathbf{k}| = (k_x^2 + k_z^2)^{1/2}, \qquad (1)$$

yielding a one-to-one relation between ω and $\mathbf{k}$. (N is the Brunt–Väisäla frequency.) This follows from the Taylor–Goldstein equations (e.g. Nappo, 2002) in the case where there is no overall wind. Both the SCT and DFT assume that the waves 'saturate' i.e. that instabilities limit the horizontal wind u to the horizontal group velocity $u_g = \partial \omega / \partial k_x$ which (for $k_x \ll k_z$) implies the restrictive relation $u(k_z) = Nk_z^{-1}$. (Here and below, we ignore the second horizontal component, y.) In addition, both SCT and DFT use the linear theory 'polarization' relations to uniquely determine the vertical velocity w from u.

We concentrate on the SCT since it gives more precise predictions. In the SCT a (presumably) highly nonlinear wave cascade (*not* turbulent cascade!) is invoked (Dewan, 1997). Since this is controlled by the energy flux ε, dimensional analysis yields the classical $\mathcal{A}(\omega) = \varepsilon^{1/2}\omega^{-1/2}$, where $\mathcal{A}$ is the amplitude of the wave at frequency ω. As a result the model predicts the following (1-D) energy spectra for the horizontal wind:

$$E(\omega) = \varepsilon\omega^{-2}; \quad E(k_x) = \varepsilon^{2/3}k_x^{-5/3}; \quad E(k_z) = N^2 k_z^{-3}. \qquad (2)$$

The first two spectra are in fact the classical Kolmogorov relations (although for anisotropic turbulence), whereas the key third prediction (as noted in Dewan, 1997) actually follows directly from dimensional analysis if N is assumed to be the unique parameter controlling the vertical structure. This fact suggests that many other (less restrictive) derivations are possible. Since $H_z = (\beta_x - 1)/(\beta_z - 1) = 1/3$, we see that the elliptical dimension D_s characterizing the vertical stratification of this model is $2 + 1/3 = 7/3$, (however, since it is a pure unlocalized wave model, the notion of elliptical dimension to characterize the volumes of its structures is not very useful).

2.4. Do stable atmospheric layers exist?

There are two basic difficulties with the quasi-linear gravity wave theories. First, in order for quasi-linear gravity waves to be meaningful, the nonlinearity must be weak enough so that linear dispersion relations can be defined whereas if the spectra have turbulent exponents (and are presumably turbulence driven), then the nonlinearity must on the contrary be strong. Second, quasi-linear gravity waves rely on the existence of homogenous layers with well-defined smoothly varying and real Brunt–Väisälä frequencies, N. Indeed, we saw the crucial role of N in determining the vertical scaling of the horizontal wind spectrum. Whereas in the familiar turbulence laws the dimensional constants are quadratic invariants (and hence positive definite), N has no restrictions; it is even imaginary for statically unstable layers. In addition, even when N is real, it is defined by the derivative of the potential temperature. This is only appropriate if, as we estimate it over thinner and thinner layers,

$$\lim_{\Delta z \to 0} \langle |N^2(\Delta z)| \rangle$$

is well defined. However, in Schertzer and Lovejoy (1985a) and in dropsondes it was found that

$$\langle |N^2(\Delta z)| \rangle \approx \Delta z^{H_N},$$

with $H_N \approx_+ 0.1$ with implying $\lim_{\Delta z \to 0} \langle |N^2(\Delta z)| \rangle \to 0$. In practice, this means that N^2 is dependent on the small dissipation scales. Another implication of the scaling of N^2 is that this buoyancy frequency would not be expected to break the scaling of the dynamics.

But do homogeneous layers really exist? In an attempt to answer this question, Lovejoy *et al.* (2008c) used classical static, convective and dynamic stability criteria to locate unstable layers at various resolutions (i.e. $N^2(\Delta z) > 0$, $N_E^2(\Delta z) > 0$ and $Ri(\Delta z) > 1/4$ respectively, where $Ri(\Delta z)$ is the Richardson number estimated over a layer of thickness Δz and N_E is the analogue of N based on the equivalent potential temperature, which takes the humidity into account). Analysis at low (320 m) resolution simply confirmed the usual stability picture of the daytime atmosphere: the lower atmosphere is unstable, the upper atmosphere is stable. At somewhat higher

resolution (80 m, typical of radiosonde resolutions), the surface layer itself typically has a few stable sublayers, and the stable upper troposphere has a few unstable sublayers. However, when the resolution is increased (to 5 m), each apparently stable sublayer was found to consist of a hierarchy of unstable subsublayers, themselves embedded with stable subsubsublayers, etc., with the same hierarchical structure holding in reverse for the initially unstable layers. Indeed, the unstable layers were found to accurately form a fractal set in the vertical (with correlation dimensions 0.64, 0.78, 0.85 for N, Ri, N_E criteria, respectively), the 'Russian doll'-like hierarchy was indeed scale invariant. At least in several instances, the possibility that the analysis was a spurious consequence of noise could be eliminated since the dropsondes were regularly launched in pairs about 20–30 m apart so they confirmed each other.

These results show that, as the layers get thinner, their probabilities of being stable approach unity. This is in accord with the traditional view that thin enough layers tend to be stable. However at the same time, fixed thick layers are made up of increasing numbers of thinner and thinner sublayers, and the number which are unstable diverges to infinity in the thin-layer limit. Saying the same thing in another way, we could say that in the small-scale limit, the stable layers form a dense set in the vertical with dimension 1. Nevertheless there is a fractal subset with an infinite number of unstable layers but of measure zero.

This picture also brings into doubt the relevance of the 'shear turbulence' discussed by Lumley (1967), Ishihara *et al.* (2002) and Wyngaard and Cote (1972). Here, one considers strongly sheared turbulence which is also strongly stabilized by buoyancy via a large N. If the shear frequency ($\omega_s = \Delta v / \Delta z$) is much smaller than N, then one can obtain corrections to the Kolmogorov law by expanding in the dimensionless ratio ω_s/N. However, for vertical shears, this amounts to expanding in $Ri^{-1/2}$, but we have just seen that Ri does not remain large and positive in a useful way. We could also comment that shear turbulence has $H_h = H_v$ (it is essentially isotropic), so that it would not be compatible with the observed anisotropic scaling.

Finally, it could be mentioned that one can also use the dropsonde pairs to directly investigate the accuracy of the linearizations by directly calculating both mean and perturbed profiles and comparing the mean vertical shear and perturbed vertical shear terms in the usual derivation of the Taylor–Goldstein equations. When this is done, one finds that, even when the mean profile is defined at 80 m resolution, the neglected terms are readily 10 times the size of the kept terms; this is not surprising since the same data show that even at scales as small as 20 m, the fluctuation Reynolds numbers, *Re*, are between 10^6 and 10^7. This large *Re* means that theories of 'wave turbulence' (L'vov *et al.*, 1997) – which share with linear gravity wave theories assumptions of weak turbulence and quasi-linear dispersion relations – would

only be relevant at very small scales (where *Re* is small enough).

3. The anisotropic scaling 23/9-D model

3.1. The basic assumptions

The atmosphere is forced by solar heating which acts via buoyancy forces. In addition to the energy flux, ε, buoyancy introduces a second quadratic invariant (the flux ϕ). The original isotropic model combining these two fluxes predicts that the Kolmogorov law should only be present for scales smaller than the small Bolgiano length, l_B, typically of the order of a metre or so (see below and parts II, III for estimates). This is in contradiction to the horizontal data which are consistent with $k_x{}^{-5/3}$ out to large scales (part II). One is therefore faced with the choice of (a) abandoning the buoyancy variance flux ϕ as being essentially irrelevant, or (b) abandoning the assumption of isotropy. Whereas the mainstream turbulence approaches have opted for (a), the 23/9-D anisotropic scaling model is based on choice (b) i.e. it postulates a regime where both fluxes are important over a wide range of scales.

In the basic 23/9-D model, the dynamics are dominated in the vertical by the buoyancy force variance flux $\phi = \Delta f^2/\tau$ (units of distance^2/time^5; $\Delta f = g\Delta(\log\theta)$ is the buoyancy force gradient across a layer thickness Δz, and τ is the time-scale of the transfer). In the horizontal, the dynamics are dominated by the energy flux $\varepsilon = \Delta v^2/\tau$, where $\Delta v(\Delta x)$ is a horizontal gradient in the horizontal wind, and the time-scale for the transfer is $\tau = \Delta x/\Delta v$. Ignoring for the moment intermittency (see section 3.3 and appendix A), from dimensional analysis on the basic fluxes ϕ and ε, one obtains unique length- and time-scales:

$$l_B = l_s = \phi^{-3/4}\varepsilon^{5/4}; \quad \tau_B = \tau_s = \phi^{-1/2}\varepsilon^{1/2}. \qquad (3)$$

The subscript *B* is used for 'Bolgiano' because analogous scales are introduced in the Bolgiano–Obhukov and other buoyancy subrange theories (where ϕ is defined instead with the help of the Boussinesq approximation). However, the important difference is that, in these theories, they denote transition scales between two isotropic regimes (a small-scale isotropic ε-dominated regime and a large-scale isotropic ϕ-dominated regime). In contrast, in the 23/9-D model, there is no qualitative change in behaviour; $\phi^{-3/4}\varepsilon^{5/4}$ is simply the scale at which structures are roughly isotropic, it is the 'sphero-scale' and hence we prefer the notation l_s with τ_s as the corresponding 'sphero-time' – the lifetime of structures in the flux the size of the sphero-scale. While the l_B, τ_B of the original buoyancy subrange theories were never empirically estimated, the only direct measurements of l_s are aircraft-based ($l_s = 4$ cm for wind in the stratosphere; Lovejoy *et al*., 2004) and for lidar aerosol data $l_s \approx 4$–80 cm (with mean around 10 cm, see part II).

Both results are close to the original (theoretically estimated) values of l_B (of the order of 1 m). Comparable estimates have been obtained indirectly using dropsonde pairs using $\varepsilon \approx \Delta v^3/\Delta x$, $\phi \approx \Delta v^5/\Delta z^3$ and Equation (3) (see also appendix B); over a 12 km thick layer we found $l_s = 1.2$ m, $\tau_s = 13$ s, $v_s = 9.3$ cm s^{-1}. In part III, we use meteorological analyses over North America to estimate comparable values $l_s = 30$ cm, $\tau_s = 7.2$ s, $v_s = 2.2$ cm s^{-1}, although the consistency of these large-scale mean values belies an enormous variability due to intermittency. (See appendix A and part III, where we use meteorological analyses to show that the probability distributions have 'fat' algebraic tails.) Since the turbulent flux ϕ replaces N as the basic dimensional parameter governing the vertical dynamics, we expect the sphero-frequency $\omega_s = 1/\tau_s = \phi^{1/2}\varepsilon^{-1/2}$ to replace N as the basic (flux-dependent) time-scale (however, there is a complication due to the vertical velocity which is discussed in section 3.4). The above estimates of τ_s imply $\omega_s \approx 0.1$ Hz is indeed comparable to gravity-wave-based estimates of the buoyancy frequencies; e.g. Allen and Vincent (1995) give $N \approx 0.03$ Hz as a typical buoyancy frequency.

3.2. Anisotropic scaling and scale functions

The basic hypothesis is that ε dominates the horizontal and ϕ the vertical so that horizontal wind differences follow:

$$\Delta v(\Delta x) = \varepsilon^{1/3}\Delta x^{1/3}, \qquad (4a)$$

$$\Delta v(\Delta y) = \varepsilon^{1/3}\Delta y^{1/3}, \qquad (4b)$$

$$\Delta v(\Delta z) = \phi^{1/5}\Delta z^{3/5}, \qquad (4c)$$

$$\Delta v(\Delta t) = \varepsilon^{1/2}\Delta t^{1/2}, \qquad (4d)$$

where Δx, Δy, Δz, Δt are the increments in horizontal coordinates, vertical coordinate and time respectively. Equations (4a), (4b) describe the real-space horizontal Kolomogorov scaling and (4c) the vertical BO scaling for the velocity; the equality signs should be understood in the sense that each side of the equation has the same scaling properties. The anisotropic Corrsin–Obukov law (see parts II, III) is obtained by the replacements $v \to \rho$; $\varepsilon \to \chi^{3/2}\varepsilon^{-1/2}$ where ρ is the passive scalar density, and χ is the passive scalar variance flux. We have included Equation (4d), which is the result for the pure time evolution in the absence of an overall advection velocity; this is the classical Lagrangian version of the Kolmogorov law. In parts II, III, we indeed show that the anisotropic Corrsin–Obhukov version of Equations (4a)–(4c) hold fairly accurately although we find that, for the temporal scaling (4d), is not directly relevant being replaced by an 'effective' scaling law with very similar exponent, arising from the effects of the vertical velocity. Below we show how to modify the above to take advection into account.

Note that the question about Kolmogorov versus Bolgiano–Obukhov scaling is also debated in laboratory

buoyancy-driven turbulence flows (e.g. Bénard convection; e.g. Ashkenazi and Steinberg, 1999; Shang and Xia, 2001). The 23/9-D theory may apply there as well as in the atmosphere, although at the moment the crucial vertical spectrum of the horizontal velocity has not been adequately investigated (see the discussion in Lilley *et al.*, 2004).

Since there is no characteristic scale, we expect generalized scale changes to define mathematical groups with generators $\mathbf{G}$. Using this idea, the scaling (Equations (4a)–(4d)) can be combined in a single expression valid for any space–time vector displacement $\Delta \mathbf{R} = (\Delta \mathbf{r}, \Delta t) = (\Delta x, \Delta y, \Delta z, \Delta t)$ by introducing a scalar function of space–time vectors called the '(space–time) scale function', denoted $[\![\Delta \mathbf{R}]\!]$, which satisfies the fundamental (functional) scale equation:

$$[\![T_\lambda \Delta \mathbf{R}]\!] = \lambda^{-1}[\![\Delta \mathbf{R}]\!], \tag{5}$$

(Schertzer and Lovejoy, 1985, 1987), where the scale-changing operator T_λ is a continuous one-parameter (Lie) group with space–time generator $\mathbf{G}_{\mathrm{st}}$ such that:

$$T_\lambda = \lambda^{-\mathbf{G}_{\mathrm{st}}}. \tag{6}$$

It will also be convenient to introduce a corresponding spatial scale function $||\Delta \mathbf{r}||$ which is a scalar function of position vector, and which satisfies the scale equation:

$$||\lambda^{-\mathbf{G}_{\mathrm{s}}} \Delta \mathbf{r}|| = \lambda^{-1}||\Delta \mathbf{r}||, \tag{7}$$

with spatial generator $\mathbf{G}_{\mathrm{s}}$. When $\mathbf{G}_{\mathrm{s}}$, $\mathbf{G}_{\mathrm{st}}$ are matrices (corresponding to linear group generators), the notion of scale is position independent. When the generator is nonlinear, the $\mathbf{G}$'s are more general nonlinear operators which will depend on the coordinates; the notion of scale will be position dependent. For the case in which $\mathbf{G}$ is the identity matrix, we have the usual isotropic, self-similar scale changes. In the case of 'linear GSI', where $\mathbf{G}$ is a diagonal matrix, the system is 'self affine' and we obtain stratification along a coordinate axes. Finally, when $\mathbf{G}$ has off-diagonal elements, we have differential rotation and stratification. The idea is that the basic dynamical symmetries determine the $\mathbf{G}$'s and the scale function is then determined by solving the functional scale Equations (5)–(7) for specific boundary conditions, i.e. by specifying all the unit vectors (the 'unit balls'; see below). T_λ then generates all the other vectors; by acting on the unit vectors in this way it determines the scale.

To unify horizontal, vertical and temporal turbulent fluctuations as described in Equation (4), we require:

$$\mathbf{G}_s = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & H_z \end{pmatrix}; \quad H_z = \left(\frac{1}{3}\right) / \left(\frac{3}{5}\right) = \frac{5}{9}, \tag{8}$$

$$\mathbf{G}_{\mathrm{st}} = \begin{pmatrix} \mathbf{G}_s & 0 \\ 0 & H_t \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & H_z & 0 \\ 0 & 0 & 0 & H_t \end{pmatrix};$$

$$H_t = \left(\frac{1}{3}\right) / \left(\frac{1}{2}\right) = \frac{2}{3}, \tag{9}$$

where the rows and columns correspond to the x, y, z and t directions respectively, and $\mathbf{G}_s$ is the matrix corresponding to the spatial part only; we define $D_s = \mathrm{Tr}(\mathbf{G}_s)$ and $D_{\mathrm{st}} = \mathrm{Tr}(\mathbf{G}_{\mathrm{st}}) = D_s + H_t$ as the elliptical dimensions characterizing the spatial and space–time anisotropies respectively. With the above dimensionally determined exponents, we find $D_s = 23/9$, $D_{\mathrm{st}} = 29/9$. The D values are dimensions since changing the scales of the vectors by λ (by operating with $\lambda^{-\mathbf{G}}$) changes their volumes by $\det(\lambda^{-\mathbf{G}}) = \lambda^{-\mathrm{Tr}(\mathbf{G})}$, they therefore quantify the change of volume with scale ('Tr' indicates 'trace' and 'det' indicates 'determinant').

Using the space–time scale function, for an arbitrary space–time displacement $\Delta \mathbf{R} = (\Delta \mathbf{r}, \Delta t)$, we may now write the anisotropic generalization of the Kolmogorov law (Equation (4)) as:

$$\Delta v(\Delta \mathbf{R}) = \varepsilon_{[\![\Delta \mathbf{R}]\!]}^{1/3}[\![\Delta \mathbf{R}]\!]^{1/3}, \tag{10}$$

where the subscripts on the flux indicate the space–time scale over which it is averaged.

Let us now consider only the spatial scale function $||\Delta \mathbf{r}||$ by finding solutions to Equation (7). Although the method is quite general (applicable to any diagonizable matrix, including those with complex eigenvalues), we consider the case of the (already) diagonal $\mathbf{G}_s$. We start with a particularly simple (but by no means unique) 'canonical' scale function symmetric with repect to $\mathbf{G}_s$ which can be obtained using the following nonlinear transformation of variables:

$$\mathbf{r}' = \left(x, y, l_s \left|\frac{z}{l_s}\right|^{1/H_z} \mathrm{sign}(z)\right), \tag{11}$$

$$||\Delta \mathbf{r}||_{\mathrm{can}} = |\Delta \mathbf{r}'|; \quad |\Delta \mathbf{r}'| = (\Delta \mathbf{r}' \cdot \Delta \mathbf{r}')^{1/2}, \tag{12}$$

so that:

$$||\Delta \mathbf{r}||_{\mathrm{can}} = l_s \left\{ \left(\frac{\Delta x}{l_s}\right)^2 + \left(\frac{\Delta y}{l_s}\right)^2 + \left(\frac{\Delta z}{l_s}\right)^{2/H_z} \right\}^{1/2}. \tag{13}$$

It is easy to check that $||\Delta \mathbf{r}||_{\mathrm{can}}$ satisfies the scale equation (7) and by setting $\Delta r = (\Delta x, 0, 0)$, $\Delta r = (0, \Delta y, 0)$, $\Delta r = (0, 0, \Delta z)$, we can check that the spatial equation $\Delta v(\Delta \mathbf{r}) = \varepsilon_{||\Delta \mathbf{r}||}^{1/3}||\Delta \mathbf{r}||^{1/3}$ reduces to Equations (4a)–(4c). To obtain more general solutions of the scale equation, we note that in the nonlinearly transformed $\Delta \mathbf{r}'$ space, the vector $\Delta \mathbf{r}'$ satisfies the scale equation (7) but with $\mathbf{G}'_s = \mathbf{I} =$ identity:

$$||\lambda^{-\mathbf{I}} \Delta \mathbf{r}'|| = \lambda^{-1}||\Delta \mathbf{r}'||. \tag{14}$$

This isotropic scale equation can be solved by inspection; one family of solutions is:

$$||\Delta \mathbf{r}'|| = \Theta(\widehat{\Omega}')|\Delta \mathbf{r}'|; \quad \widehat{\Omega}' = \frac{\Delta \mathbf{r}'}{|\Delta \mathbf{r}'|}, \tag{15}$$

where $\widehat{\Omega}'$ is a unit vector of spherical polar angles in three-dimensional (x', y', z') space. For a fixed scale $||\Delta \mathbf{r}'||$, this implies $|\Delta \mathbf{r}'| = \left(\Theta\left(\widehat{\Omega}'\right)\right)^{-1} ||\Delta \mathbf{r}'||$ which is simply a spherical polar coordinate expression for the radius $|\Delta \mathbf{r}'|$ in terms of the spherical polar direction $\widehat{\Omega}'$. The equation of the 'unit ball' defining all the unit vectors is obtained with $||\Delta \mathbf{r}'|| = 1$, i.e. in the primed space, scale function is a series of 'blow-ups' of this basic shape. In the original (unprimed) space, using $||\Delta \mathbf{r}|| = ||\Delta \mathbf{r}'||$ yields the following more general solution to Equation (7):

$$||\Delta \mathbf{r}|| = \Theta(\widehat{\Omega}')||\Delta \mathbf{r}||_{\mathrm{can}}. \qquad (16)$$

Note that when $\Theta = 1$, we obtain the canonical scale function which is a sphere in the primed space, but is somewhat elliptical in the unprimed space. As long as Θ is positive, real and continuous, the family of scale functions defined in this way are 'physical scale functions' since in addition to Equation (7), they satisfy a basic 'localization' requirement of physical scale:

$$B_\lambda \subseteq B_{\lambda'} \leftrightarrow \lambda \leq \lambda'; \ \{B_\lambda : ||\Delta \mathbf{r}|| \leq \lambda\}, \qquad (17)$$

where B_λ is the set of all the vectors with scale not exceeding λ. Scale functions satisfying condition (17) define 'balls' (basic sets) B_λ which are strictly decreasing functions of scale and therefore can be used to define anisotropic Hausdorff measures, anisotropic Hausdorff dimensions and hence the notion of spatial integration (Schertzer and Lovejoy, 1985b).

We can now use the physical spatial scale function $||\Delta \mathbf{r}||$ to define the canonical, localized space–time scale function $[\![\Delta \mathbf{R}]\!]_{\mathrm{can}}$ which satisfies $[\![\lambda^{-\mathbf{G}} \Delta \mathbf{R}]\!]_{\mathrm{can}} = \lambda^{-1} [\![\Delta \mathbf{R}]\!]_{\mathrm{can}}$:

$$[\![\Delta \mathbf{R}]\!]_{\mathrm{can}} = l_s \left\{ \left(\frac{||\Delta \mathbf{r}||}{l_s}\right)^2 + \left(\frac{|\Delta t|}{\tau_s}\right)^{2/H_t} \right\}^{1/2}, \quad (18)$$

(Marsan *et al.*, 1996). While this special space–time form will be retained for the turbulent fluxes, for the observables such as the wind, we are not so strongly restricted; we shall see that more general space–time scale functions can be used leading to (unlocalized) wave behaviour. It is useful to note that the space–time scaling of classical (spatially isotropic) turbulence laws can also be expressed with the help of scale functions. For example, the 3-D isotropic Kolmogorov law is obtained using

$$[\![\Delta \mathbf{R}]\!]_{\mathrm{K,can}} = l_{\mathrm{K}} \left\{ \left(\frac{|\Delta \mathbf{r}|}{l_{\mathrm{K}}}\right)^2 + \left(\frac{|\Delta t|}{\tau_{\mathrm{K}}}\right)^{2/H_t} \right\}^{1/2}, \quad (19)$$

in Equation (10) where l_{K}, τ_{K} are the Kolmogorov (viscous, dissipation) length- and time-scales respectively and $H_t = 2/3$ (see 3.4 for a discussion).

3.3. Intermittency

Starting with Novikov and Stewart (1964), Yaglom (1966) and Mandelbrot (1974), stochastic isotropic cascade models have been developed for modelling intermittent ε fields. The energy flux was chosen as the cascade quantity since, in Fourier space, the nonlinear terms in the Navier–Stokes equations conserve ε. (This $\mathrm{d}/\mathrm{d}k = 0$ conservation should not be confused with the more usual $\mathrm{d}/\mathrm{d}t = 0$ conservation). ε is modelled by a 'conservative' cascade, i.e. a cascade whose mean is independent of the scale over which it has been developed. To understand these models, we first give a statistical interpretation to the (anisotropic) Kolmogorov law Equation (10); it is usual to assume that it is actually an equality in the scaling of $\Delta v(\Delta r)$ and $\varepsilon_{[\![\Delta \mathbf{R}]\!]}^{1/3}[\![\Delta \mathbf{R}]\!]^{1/3}$ so that:

$$\langle|\Delta v|^q\rangle = \left\langle \varepsilon_{[\![\Delta \mathbf{R}]\!]}^{q/3} \right\rangle [\![\Delta \mathbf{R}]\!]^{qH}; \ H = \frac{1}{3}. \qquad (20)$$

The cascade processes thus model the highly intermittent ε field by successively breaking up an originally constant, uniform large-scale field ε_L (the 'parent eddy' at initial scale L) into smaller and smaller 'eddies' and 'sub-eddies' each of which multiplicatively modulates the parent with the process repeating down to smaller and smaller scales until it is eventually cut off by viscous dissipation. The generic cascade result is:

$$\left\langle \varepsilon_{[\![\Delta \mathbf{R}]\!]}^{q/3} \right\rangle = \varepsilon_L^{q/3} \left(\frac{L}{[\![\Delta \mathbf{R}]\!]}\right)^{K(q/3)}, \qquad (21)$$

where at this stage, $K(q)$ is relatively arbitrary convex function; the main restriction being $K(3) = 1$ corresponding to the scale by scale conservation condition: $\langle \varepsilon_{[\![\Delta \mathbf{R}]\!]} \rangle = \varepsilon_L$, all $[\![\Delta \mathbf{R}]\!]$. Note that here and below we use ensemble average ('canonical') conservation, although it is frequent in the literature to use the much more restrictive and less intermittent 'microcanonical' conservation (realization by realization, scale by scale). In the 1980s, it was realized that cascades were multifractal i.e. that the increasingly intense turbulent regions are concentrated on sparse fractal sets with dimension decreasing with increasing ε.

With this assumption about the statistics of ε, a component of the horizontal velocity field will satisfy:

$$\langle|\Delta v(\Delta \mathbf{R})|^q\rangle = \left\langle \varepsilon_{[\![\Delta \mathbf{R}]\!]}^{q/3} \right\rangle [\![\Delta \mathbf{R}]\!]^{qH} = \varepsilon_L^{q/3} \left(\frac{L}{[\![\Delta \mathbf{R}]\!]}\right)^{K(q/3)}$$

$$\times [\![\Delta \mathbf{R}]\!]^{qH} = \varepsilon_L^{q/3} L^{K(q/3)} [\![\Delta \mathbf{R}]\!]^{\xi(q)},$$

$$\xi(q) = qH - K\left(\frac{q}{3}\right); \ H = \frac{1}{3}, \qquad (22)$$

where $\xi(q)$ is the (generalized) structure function exponent. In applications, we typically use data from a small

part of the Earth (a region scale $l < L$) so that:

$$\langle (|\Delta v(\Delta \mathbf{R})|^q)_l \rangle = \langle \varepsilon_{L/l}^{q/3} \rangle \left(\frac{\langle \varepsilon_{l/[\![\Delta \mathbf{R}]\!]}^{q/3} \rangle}{\langle \varepsilon_{l/[\![\Delta \mathbf{R}]\!]} \rangle^{q/3}} \right) [\![\Delta \mathbf{R}]\!]^{q/3}$$

$$= \langle \varepsilon_{L/l}^{q/3} \rangle \left(\frac{l}{[\![\Delta \mathbf{R}]\!]} \right)^{K(q/3)} [\![\Delta \mathbf{R}]\!]^{q/3}$$

$$= \langle \varepsilon_{L/l}^{q/3} \rangle l^{K(q/3)} [\![\Delta \mathbf{R}]\!]^{\xi(q)}, \qquad (23)$$

where $\langle (|\Delta v(\Delta \mathbf{R})|^q)_l \rangle$ is the qth-order (ensemble-averaged) structure function average over the regional scale l, $\langle \varepsilon_{L/l}^{q/3} \rangle$ is the (ensemble-averaged) $(q/3)$ order moment of the energy flux from the larger scales down to the data region, i.e. from a cascade developed over a scale ratio L/l. This formula follows from Equation (21) by using the multiplicative property of the cascade which factors into low-frequency and high-frequency parts as indicated. (In the usual presentations, the normalization factor $\langle \varepsilon_{l/[\![\Delta \mathbf{R}]\!]} \rangle^{q/3}$ is absent since the cascade is assumed normalized to unity, i.e. $\langle \varepsilon_{[\![\Delta \mathbf{R}]\!]} \rangle = 1$ for all $[\![\Delta \mathbf{R}]\!]$). Writing the result this way (Equation (23)) is convenient since in practice we often estimate the structure functions by averaging over the data from a region size l, but from a single realization of the overall 'weather' process. In this case we drop the ensemble average on the left-hand side, as well as from the large scale $\varepsilon_{L/l}$ to obtain:

$$(|\Delta v(\Delta \mathbf{R})|^q)_l \approx \varepsilon_{L/l}^{q/3} \left(\frac{\langle \varepsilon_{l/||\Delta \mathbf{R}||}^{q/3} \rangle}{\langle \varepsilon_{l/||\Delta \mathbf{R}||} \rangle^{q/3}} \right) [\![\Delta \mathbf{R}]\!]^{q/3}$$

$$= \varepsilon_{L/l}^{q/3} \left(\frac{l}{[\![\Delta \mathbf{R}]\!]} \right)^{K(q/3)} [\![\Delta \mathbf{R}]\!]^{q/3}$$

$$= \varepsilon_{L/l}^{q/3} l^{K(q/3)} [\![\Delta \mathbf{R}]\!]^{\xi(q)}; \qquad (24a)$$

$$(|\Delta v(\Delta \mathbf{R})|^q)_l \approx (|\Delta v(l)|^q)_l \left(\frac{[\![\Delta \mathbf{R}]\!]}{l} \right)^{\xi(q)}. \qquad (24b)$$

The $\approx$ sign is used since, for the small scales, we have approximated the ensemble average quantities by spatial averages over the (many) small-scale structures. Equation (24b), which will be useful later, uses the notation $(|\Delta v(l)|^q)_l$ to indicate the qth moment of the fluctuation over a scale l of the velocity spatially averaged over a scale l. Equations (24a, b) are valid for $[\![\Delta \mathbf{R}]\!] \leq l$.

Mathematically, practically any convex $K(q)$ is possible so that the theory above effectively has an infinite number of parameters or – from an empirical standpoint – an infinite number of empirical exponents (one for each q). However, we are considering a dynamic process which repeats scale after scale; we may therefore suspect – as is usual in physics – that many of the details of the interactions are unimportant; this is the question of 'multifractal universality'. Indeed, since the cascades are multiplicative, we can appeal to a kind of 'multiplicative central limit theorem'; roughly, the log of the cascade is an additive process and hence in the appropriate limit

one obtains Gaussian and Levy distributions for the logs of the process. In fact, due to the singular small scale cascade limit, things are a little more complicated (hence the debate; Schertzer and Lovejoy, 1997). It turns out that the resulting attractive, stable multifractal processes are not exactly log normal or log Levy, but are nearly so except for high-order moments which diverge (Schertzer and Lovejoy, 1987, 1997). Although this does not constitute a 'proof' (other factors may prevent the system from attaining its universal limit), it is nevertheless natural to suppose that $K(q)$ is of the special two-parameter 'universal' form:

$$K(q) = \frac{C_1}{\alpha - 1} (q^\alpha - q), \qquad (25)$$

where $0 < C_1 < D$ is the co-dimension of the mean field, D is the elliptical dimension of the space, and $0 \leq \alpha \leq 2$ is the index of the Levy noise generator (see below). Except for the case $\alpha = 2$, Equation (25) is only valid for $q \geq 0$, $K(q)$ diverges for $q < 0$. Empirically, we find $C_1 \approx 0.25$, $\alpha \approx 1.5$ for ε (e.g. Schertzer *et al.*, 1995, 1997a; Schmitt *et al.*, 1996); using dropsonde pairs (appendix A), we find $\alpha_\varepsilon \approx 1.6$, $C_{1\varepsilon,v} \approx 0.50$ in the vertical, using $H_z = 5/9$, this yields the comparable (horizontal) estimate $C_{1\varepsilon,h} \approx 0.50 \times 5/9 = 0.28$. (Note: The co-dimension of a (possibly anisotropic) fractal set is simply the difference between the dimension of the space and the corresponding fractal dimension). The co-dimension of the mean field is that which characterizes the fractal set which gives the dominant contribution to the mean. This one-to-one correspondence between the field values (more precisely, orders of singularities) and statistical moments is a consequence of the Legendre transform relation between the two (Parisi and Frisch, 1985). In any case, from an empirical point of view, the turbulent data are very close to the universal form so that the latter is at least a good parametrization of the actual $K(q)$. The only caveat is that due to the singular small-scale limit, we must distinguish 'bare' and 'dressed' cascade quantitites. While the former are those given above – they only take into account the effect of the largest scales down to $[\![\Delta \mathbf{R}]\!]$ – the dressed quantities are the result of a cascade proceeding to the small-scale $[\![\Delta \mathbf{R}]\!]$ limit and then integrated up to the same finite scale. The scaling exponents are identical for all q less than a critical q_D after which the dressed moments diverge; a 'first-order multifractal phase transition'. Thus the universal multifractal formula Equation (25) only holds for the dressed quantities for $q < q_D$. In Schertzer and Lovejoy (1985a) and in part III, we give empirical evidence that $q_D \approx 1.66$ for ε. Our object below is therefore to simulate stochastic velocity fields with statistics satisfying Equations (22) and (25).

3.4. Advection and the effective space–time generator

Before proceeding, we must discuss advection. As pointed out in Schertzer *et al.* (1997b), advection can

be taken into account using the Gallilean transformation matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & w \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad (26)$$

where the mean wind vector has components: $\mathbf{v} = (u, v, w)$. The new generator is $\mathbf{G}_{\text{st,advec}} = \mathbf{A}^{-1}\mathbf{G}_{\text{st}}\mathbf{A}$ and the scale function $[\![\Delta\mathbf{R}]\!]_{\text{advec}}$ which is symmetric with respect to $\mathbf{G}_{\text{st,advec}}$ is: $[\![\Delta\mathbf{R}]\!]_{\text{advec}} = [\![\mathbf{A}^{-1}\Delta\mathbf{R}]\!]$. The canonical advected scale function is therefore:

$$[\![\Delta\mathbf{R}]\!]_{\text{advec,can}} = [\![\mathbf{A}^{-1}\Delta\mathbf{R}]\!]_{\text{can}} = l_s \left\{ \left(\frac{\Delta x - u\Delta t}{l_s}\right)^2 \right.$$
$$\left. + \left(\frac{\Delta y - v\Delta t}{l_s}\right)^2 + \left(\frac{\Delta z - w\Delta t}{l_s}\right)^{2/H_z} + \left(\frac{\Delta t}{\tau_s}\right)^{2/H_t} \right\}^{1/2}.$$
$$(27)$$

Note that, since $D_{\text{st,advec}} = \text{Tr}(\mathbf{G}_{\text{st,advec}}) = \text{Tr}(\mathbf{A}^{-1}\mathbf{G}_{\text{st}}\mathbf{A}) = \text{Tr}(\mathbf{G}_{\text{st}}) = D_{\text{st}}$, constant advection does not affect the elliptical dimension (however, see the next subsection).

3.5.   The special role of vertical velocity, effective temporal scaling, effective space-time generators

Taking into account the advection velocities u, v, w, at a single point $\Delta x = \Delta y = \Delta z = 0$, the scale function has the following temporal scaling:

$$[\![(0, 0, 0, \Delta t)]\!] = l_s \left\{ \left(\frac{u\Delta t}{l_s}\right)^2 + \left(\frac{v\Delta t}{l_s}\right)^2 \right.$$
$$\left. + \left(\frac{w\Delta t}{l_s}\right)^{2/H_z} + \left(\frac{\Delta t}{\tau_s}\right)^{2/H_t} \right\}^{1/2}, \qquad (28)$$

(cf. Equation (27)). This formula is valid due to the Gallilean invariance of the equations and boundary conditions: it assumes that the advection velocity is essentially constant over the region. Let us first consider the horizontal velocity term. We note that even if we consider a flow with zero imposed mean horizontal velocity, that the typical largest eddy (size L) will have a mean velocity $u \approx v \approx v_L \approx \varepsilon_L^{1/3} L^{1/3}$ and will survive for the corresponding eddy turnover time $\tau_{\text{e},L} = L/v_L = \varepsilon_L^{-1/3} L^{2/3}$. If we use the typical values $\varepsilon \approx 10^{-4}$–$10^{-3}$ m^2 s^{-3}, (cf. the probability distribution in part III) then in our picture with global horizontal scaling with outer scale $L = 2 \times 10^4$ km, this implies $v_L \approx 20$ m s^{-1}, $\tau_{\text{e},L} \approx 10^6$ s (2 weeks, see part III). These are indeed typical values of the large-scale mean wind and of the 'synoptic maximum' respectively; the latter is the lifetime of typical structures of atmospheric extent; Kolesnikov and Monin (1965). The synoptic maximum is the natural time period separating weather from climate variability so that our model naturally predicts the transition to climate (and possibly some of the different scaling properties of the

climate regime; e.g. Tessier *et al.* (1996) for empirical results for rain). (Interestingly, just as our model predicts the correct time-scale for planetary-scale eddies, using the observed $l_s \approx 1$ m, it predicts that the vertical scale corresponding to the outer horizontal scale ($L = 20\,000$ km) is $l_s(L/l_s)^{H_z} \approx 11$ km which implies that the thickness of the troposphere roughly equals the natural vertical outer scale.)

Tennekes (1975) argued that, due to this large mean horizontal wind, the smaller structures are 'swept' along, an effect which dominates that of 'pure' temporal scaling. This can be seen by using space–time scale functions to express the standard Kolmogorov law (Equation (19)); but with advection. Considering a single point, we obtain the scale function Equation (28), but with $l_s \rightarrow l_{\text{K}}$, $\tau_s \rightarrow \tau_{\text{K}}$, $H_z \rightarrow 1$ where l_{K}, τ_{K} are the Kolmogorov (dissipation) scales. The condition that the horizontal wind dominates is $u\Delta t/l_{\text{K}} > (\Delta t/\tau_{\text{K}})^{3/2}$ which implies $\Delta t < u^2/\varepsilon$ (using $\varepsilon = (l_{\text{K}}^2/\tau_{\text{K}}^3)$). Since in the 23/9-D model only the vertical is different, and since $\varepsilon = (l_s^2/\tau_s^3)$, we see that that the same conclusion continues to apply in our anisotropic 23/9-D model. Using $\tau_{\text{e},L} = u^2/\varepsilon$ for our estimate of the eddy turnover time at the largest scale, we see that we expect horizontal wind dominance for periods $\Delta t < \tau_{\text{e},L}$, i.e. less than about 2 weeks. Although this result needs to be nuanced due to intermittency effects (part III), the basic conclusion – that the pure temporal development term $(\Delta t/\tau_s)^{1/H_t}$, is generally subdominant – holds when tested against the meteorological data discussed in more detail in part III.

However, unlike the isotropic 3-D Kolmogorov turbulence considered by Tennekes (1975), in the 23/9-D model, the scale function has a new term $(w\Delta t/l_s)^{9/5}$ with a large exponent which – under the right conditions, (including strong enough vertical velocity w) – could be dominant. In part III, we argue that - depending on scale - that this vertical wind term is indeed frequently dominant, occasionally at scales as small as 10 s.

Let us therefore consider the effect of a mean vertical wind. Unlike the horizontal wind which will typically have a large mean value even when averaged over large space-time regions, the vertical wind is typically small and has a mean that generally decreases approaching zero for large enough averaging scales. Indeed, although the w statistics are not well established, a symptom of atmospheric anisotropy is that w behaves very differently from u, v. For example, according to limited analysis of dropsonde data (Lovejoy *et al.*, 2007; these sondes have trouble estimating w), and meteorological analyses (also inaccurate for w), it appears that the (horizontal, vertical, temporal) spectrum of w is a power law with exponent $\beta_w < 1$. Indeed, it seems likely that the w field can be modelled by a fractional integration of order $H_w < 0$ (i.e. a positive order fractional differentiation $-H_w > 0$) with respect to a multifractal flux. (Recall that the corresponding exponent for the horizontal Kolmogorov law is $H = 1/3$; see the FIF model description in section 4.) Ignoring intermittency corrections, this would yield $\beta_w = 1 + 2H_w < 1$. In the case $H_w < 0$, following

Equation (24b), the statistics are:

$$|(w(\Delta t))_l|^q = |(w(l))_l|^q \left(\frac{[\![\Delta t]\!]}{l}\right)^{\xi_w(q)}; \quad (29a)$$

$$\xi_w(q) = H_w q - K_w(q).$$

(When $H_w > 0$, we need to use Δw in place of w in the above.) We have used the shorthand notation $[\![\Delta t]\!] = [\![(0, 0, 0, \Delta t)]\!]$. $|(w(\Delta t))_l|$ represents the absolute value of the fluctuation over time Δt in the vertical velocity spatially averaged over a region size l, and $|(w(l))_l|$ is the corresponding absolute value of the fluctuation over distance l.

To see what this implies, ignore intermittency (which is discussed in appendix B) i.e. put $K_w = 0$, so that:

$$|(w(\Delta t))_l| = |(w(l))_l| \left(\frac{[\![\Delta t]\!]}{l}\right)^{H_w}. \quad (29b)$$

However, considering the scale function at a single point ($\Delta x = \Delta y = \Delta z = 0$), and neglecting the horizontal wind (i.e. $u = v = 0$) and the pure temporal development terms (i.e. τ_s is large), we see that

$$[\![\Delta t]\!] = l_s \left(|(w(\Delta t))_l| \, \Delta t / l_s\right)^{1/H_z}.$$

Comparing this with Equation (29b), we see that $[\![\Delta t]\!] \approx \Delta t^{1/H_t'}$ with $H_t' = H_z - H_w$, so that if $H_w \approx -1/6$ then $H_t' \approx 2/3$, close to the empirical result.

In appendix B, we take into account intermittency and average over the vertical velocity fluctuations, and obtain the same type of result:

$$[\![\Delta t]\!] = l_s \left(\frac{\Delta t}{\tau_s'}\right)^{1/H_t'}; \quad \tau_s' = \tau_s \left(\frac{v_s}{|(w(l))_l|}\right) \left(\frac{l}{l_s}\right)^{H_z - H_t'}, \quad (30)$$

where $H_t' > H_z$ depends on the statistics of the vertical wind via $\xi_w(q)$ (see appendix B for an estimate); with $H_w \approx -0.1$, we can readily have $H_t' \approx H_t$ which is compatible with the analyses in part III. For large enough vertical velocities, τ_s' can be appreciably smaller than τ_s so that this velocity term with effective exponent H_t' can dominate the vertical advection terms, i.e.

$$\left(\frac{\Delta t}{\tau_s'}\right)^{1/H_t'} > \left(\frac{\Delta t}{\tau_s}\right)^{1/H_t}.$$

In part III using meteorological analyses, this is found to be generally the case. Therefore, replacing $(\Delta t/\tau_s)^{1/H_t}$ with $(\Delta t/\tau_s')^{1/H_t'}$ (and putting the 'effective vertical velocity' to zero) we may replace $G_{\rm st}$ with an 'effective generator' and effective advection matrix:

$$\mathbf{G}_{\rm st,eff} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & H_z & 0 \\ 0 & 0 & 0 & H_t' \end{pmatrix}; \quad \mathbf{A}_{\rm eff} = \begin{pmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

$$D_{\rm st,eff,advec} = \mathrm{Tr}(\mathbf{A}_{\rm eff}^{-1} \mathbf{G}_{\rm st,eff} \mathbf{A}_{\rm eff})$$

$$= \mathrm{Tr}(\mathbf{G}_{\rm st,eff}) = 2 + H_z + H_t', \quad (31)$$

with corresponding effective scale function:

$$[\![\Delta \mathbf{R}]\!]_{\rm advec,eff,can} = [\![A^{-1}\Delta \mathbf{R}]\!]_{\rm eff,can}$$

$$= l_s \left\{ \left(\frac{\Delta x - u\Delta t}{l_s}\right)^2 + \left(\frac{\Delta y - v\Delta t}{l_s}\right)^2 \right.$$

$$\left. + \left(\frac{\Delta z}{l_s}\right)^{2/H_z} + \left(\frac{\Delta t}{\tau_s'}\right)^{2/H_t'} \right\}^{1/2}. \quad (32)$$

4. The FIF model

4.1. The classical FIF model

We now describe the FIF model (Schertzer and Lovejoy, 1987) which satisfies the anisotropic generalization of the Kolmogorov law (10) and the multiscaling statistics (22), with universal $K(q)$, Equation (25). One starts with a subgenerator $\gamma_\alpha(\mathbf{r}, t)$ which is a noise composed of independent identically distributed (i.i.d.) Levy random variables (since the variance is generally infinite, this i.i.d. noise is not generally 'white', although it is a generalization of white noise; appendix C of Schertzer and Lovejoy, 1987). The normalization of $\gamma_\alpha(\mathbf{r}, t)$ depends on C_1 and on the form of the probability distribution which depends on α; the special case $\alpha = 2$ is the Gaussian leading to the 'log-normal' multifractals. In order for the moments of the resulting ε field to be finite, we must use 'extremal' Levy variables, i.e. those which are maximally asymmetric, in this case with only a single (heavy) algebraic tail on the negative side. (For $0 < \alpha < 1$, γ_α will be strictly negative, for $1 < \alpha \le 2$ it will have a rapid 'stretched exponential' fall-off for $\gamma_\alpha > 0$). One next produces the generator $\Gamma(\mathbf{r}, t)$ by convolving ($*$) this with the scaling propagator (the space–time Green's function, to which we return soon) $g_\varepsilon(\mathbf{r}, t)$:

$$\Gamma(\mathbf{r}, t) = \gamma_\alpha(\mathbf{r}, t) * g_\varepsilon(\mathbf{r}, t);$$

$$\widetilde{\Gamma}(\mathbf{k}, \omega) = \widetilde{\gamma}_\alpha(\mathbf{k}, \omega) \widetilde{g}_\varepsilon(\mathbf{k}, \omega), \quad (33)$$

where we have indicated Fourier transforms by tildas. The conserved flux ε is then obtained by exponentiation:

$$\varepsilon(\mathbf{r}, t) = e^{\Gamma(\mathbf{r}, t)}. \quad (34)$$

We note in passing that the small-scale limit of ε only exists in the sense of weak measures (Kahane, 1985), so that only averages over finite sets converge. This is the origin of the interesting 'dressed' properties including divergence of high-order statistical moments, mentioned earlier.

The horizontal velocity field is obtained by a final convolution with the (generally different) propagator g_v:

$$v(\mathbf{r}, t) = \varepsilon^{1/3}(\mathbf{r}, t) * g_v(\mathbf{r}, t);$$

$$\widetilde{v}(\mathbf{k}, \omega) = \widetilde{\varepsilon^{1/3}}(\mathbf{k}, \omega) \widetilde{g}_v(\mathbf{k}, \omega). \quad (35)$$

In order to satisfy the scaling symmetries, it suffices for the propagators to satisfy the generalized scale equation:

$$g_{v,\varepsilon}(T_\lambda(\Delta\mathbf{r}, \Delta t)) = \lambda^{(D_{st}-H)}g_{v,\varepsilon}(\Delta\mathbf{r}, \Delta t), \qquad (36)$$

where H must be chosen $= D_{st}(1 - 1/\alpha)$ for g_ε, and $H = 1/3$ for g_v. (Recall that $D_{st} = \text{Tr}(\mathbf{G}_{st})$ is the 'elliptical dimension' characterizing the overall stratification of space–time). The relevant solutions of Equation (36) are powers of scale functions with a Heaviside function $h(t)$ needed to ensure that causality is respected:

$$g_{v,\varepsilon}(\Delta\mathbf{r}, \Delta t) = h(\Delta t)[\![(\Delta\mathbf{r}, \Delta t)]\!]^{-(D-H)};$$
$$h(t) = \begin{cases} 1 & \text{for } t > 0 \\ 0 & \text{for } t \le 0 \end{cases}. \qquad (37)$$

We note that for ε and for modelling positive fields (such as the passive scalar fields below and parts II, III), this is adequate. However, for the velocity field, it may be of interest to have symmetric positive, negative fluctuations; this can be achieved by multiplying g_v by the factor $\text{sign}(\Delta x)\,\text{sign}(\Delta y)\,\text{sign}(\Delta z)$. (Obviously other factors can be used for less symmetric fluctuations.)

4.2. An extreme unlocalized (wave) model

Although the FIF is quite general, its classical implementation (e.g. Marsan *et al.*, 1996; Schertzer *et al.*, 1997a) is obtained by using the same localized space-time scale function for g_ε and g_v. Such propagators are 'power law' localized in space–time (recall that the term 'localization' is often taken to mean the much stronger exponential localization). For g_ε this is justified by the usual turbulence phenomenology – that turbulent energy fluxes are indeed localized (as required for example by the Navier–Stokes equations). However, we shall now consider the possibility that the propagator g_v is non-local in space–time (although still with some localization in space, i.e. wave packets). The key is to recognize that the physical spatial scale function $||\mathbf{r}||$ defines a different Fourier scale function $||\mathbf{k}||$ satisfying Equation (7) but

with generator $\mathbf{G}_s^{\text{T}}$ (T indicates the transpose). Indeed, more precisely:

$$||\mathbf{k}||^{-H} \xrightarrow{\text{F.T.}} ||\mathbf{r}||^{-(D_s-H)}; \quad D_s = \text{Tr}(\mathbf{G}_s), \qquad (38)$$

where F.T. denotes a Fourier transform. Note that in terms of space–time scale functions we also have:

$$[\![\mathbf{k}, \omega]\!]^{-H} \xrightarrow{\text{F.T.}} h(t)[\![\mathbf{r}, t]\!]^{-(D-H)};$$
$$||\mathbf{k}|| = [\![\mathbf{k}, 0]\!]; \quad D = D_s + H_t. \qquad (39)$$

Equations (35), (39) indicate that the Fourier transforms of generalized power laws with respect to $\mathbf{G}_{st}$ are in turn generalized power laws with respect to $\mathbf{G}_{st}^{\text{T}}$, i.e.

$$[\![\lambda^{-\mathbf{G}_{st}}(\mathbf{r}, t)]\!] = \lambda^{-1}[\![\mathbf{r}, t]\!] \text{ and}$$
$$[\![\lambda^{-\mathbf{G}_{st}^{\text{T}}}(\mathbf{k}, \omega)]\!] = \lambda^{-1}[\![\mathbf{k}, \omega]\!].$$

This is an anisotropic extension of classical 'Tauberian theorems' (e.g. Feller, 1971). However, in spite of the notation, the Fourier and real-space scale functions are generally different.

Due to Equations (37), (39), the velocity propagator must be chosen in Fourier space to respect the appropriate scaling symmetries:

$$\widetilde{g}_v(\mathbf{k}, \omega) = [\![\mathbf{k}, \omega]\!]^{-H}. \qquad (40)$$

However, the key to the extreme unlocalized (wave) model is to choose

$$[\![\mathbf{k}, \omega]\!] = \left\{ \mathrm{i}\left(\omega - ||\mathbf{k}||^{H_t}\right) \right\}^{1/H_t}, \qquad (41)$$

(see Table I for comparisons). In order to understand the implications of this scale function, it is instructive to take the inverse Fourier transform of $\tilde{g}_v(\mathbf{k}, \omega)$ with repect to ω:

$$\widetilde{g}_v(\mathbf{k}, t) = h(t)t^{-1+H/H_t}\mathrm{e}^{\mathrm{i}||\mathbf{k}||^{H_t}t}, \qquad (42)$$

(we ignore constant factors). This shows that the propagator defined by Equations (40), (41) is a causal – due

Table I. Intercomparison of flux- and wave-like velocity propagators. In both cases, the physical (spatial) scale function $||\Delta\mathbf{r}||$ and Fourier counterpart $||\mathbf{k}||$ are the same and are linked by Equation (38).

Flux-like (Power law!) Localization in space–time	Wave-like Unlocalized
$\tilde{g}_v(\mathbf{k}, \omega) = [\![(\mathbf{k}, \omega)]\!]^{-H}$	$\tilde{g}_v(\mathbf{k}, \omega) = \left\{ \mathrm{i}\left(\omega - ||\mathbf{k}||^{H_t}\right) \right\}^{-H/H_t}$
e.g. $[\![(\mathbf{k}, \omega)]\!] \approx \widetilde{h} * \left(|\omega| + ||\mathbf{k}||^{H_t}\right)^{1/H_t}$	
$\tilde{g}_v(\mathbf{k}, t) = h(t)t^{-1+H/H_t}f(z)$	$\tilde{g}_v(\mathbf{k}, t) = h(t)t^{-1+H/H_t}f(z)$
$z = ||\mathbf{k}||^{H_t}\,t; \; f(z) \underset{z \to 0}{\to} z^{H/H_t-1}; \; f(z) \underset{z \to \infty}{\to} 1$	$f(z) = \mathrm{e}^{-\mathrm{i}z}$
$g_v(\mathbf{r}, t) = h(t)[\![(\mathbf{r}, t)]\!]^{-(D+H_t-H)}$	$g_v(\mathbf{r}, t) \approx h(t)\dfrac{\mathrm{e}^{\mathrm{i}\{\mathbf{k}\cdot\mathbf{r}-\omega_d(\mathbf{k})t+\phi_0(\mathbf{k})\}}}{t^{5/2-H/H_t}\det\left(\dfrac{\partial^2\omega_d(\mathbf{k})}{\partial k_i \partial k_j}\right)},$
e.g. $[\![(\mathbf{r}, t)]\!] = ||\mathbf{r}|| + t^{1/H_t}$	$\mathbf{r} = \mathbf{v}_g(\mathbf{k})t; \; \mathbf{v}_g(\mathbf{k}) = \nabla\omega_d(\mathbf{k}); \; \omega_d(\mathbf{k}) = ||\mathbf{k}||^{H_t}$

to $h(t)$ – temporal fractional integration of order H/H_t of waves. Indeed, we can use the standard method of stationary phase (e.g. Bleistein and Handelsman, 1986) to obtain an asymptotic approximation to the space–time convolution for v (Equation (35)):

$$v(\mathbf{x}, t) \approx \tilde{\varepsilon}^{1/3}(\mathbf{k}(\mathbf{x}), t) *_t \tilde{g}_v(\mathbf{k}(\mathbf{x}), t), \qquad (43)$$

where $*_t$ indicates convolution with respect to time only and the propagator is:

$$
\begin{aligned}
g_v(\mathbf{x}, t) &\approx \tilde{g}_v(\mathbf{k}(\mathbf{x}), t) \\
&= h(t) \frac{e^{i\{\mathbf{k}\cdot\mathbf{x} - \omega_d(\mathbf{k})t + \phi_0(\mathbf{k})\}}}{t^{5/2 - H/H_t} \det\left(\dfrac{\partial^2 \omega_d(\mathbf{k})}{\partial k_i \partial k_j}\right)}, \qquad (44)
\end{aligned}
$$

where ϕ_0 is a phase and

$$\mathbf{x} = \mathbf{v}_g(\mathbf{k})t; \quad \mathbf{v}_g(\mathbf{k}) = \nabla \omega_d(\mathbf{k}); \quad \omega_d(\mathbf{k}) = ||\mathbf{k}||^{H_t}. \quad (45)$$

Equations (44), (45) should be understood as parametric equations; $\mathbf{k}$ is the wave vector which satisfies the 'ray' equation $\mathbf{x} = \mathbf{v}_g(\mathbf{k})t$, where $\mathbf{v}_g$ is the group velocity and $\omega_d(\mathbf{k}) = ||\mathbf{k}||^{H_t}$ is the dispersion relation. (Note that we would have obtained $\omega_d(\mathbf{k}) = -||\mathbf{k}||^{H_t}$ if we had chosen the equally valid scale function Equation (41) with $-\omega$ in the place of ω.) If we consider inversion-symmetric real-space scale functions $||\Delta \mathbf{r}|| = || - \Delta \mathbf{r}||$ then $||\mathbf{k}||^{-H}$ is real so that, whenever it is >0, we obtain a corresponding positive real $\omega_d(\mathbf{k})$, i.e. waves at the corresponding wave vector $\mathbf{k}$ are non-attenuating. Finally, we may take the real part of $g_v(\mathbf{x}, t)$ to obtain a real v field after the convolution.

Equations (44), (45) show that the velocity field is the fractional time integral of wave packets propagating along rays at the group velocity, dispersing and decreasing in amplitude as they travel as t^{-2} (the exponent is $5/2 - H/H_t = 2 > 3/2$). The classical time-dependence of the attenuation of the packet is $t^{-3/2}$ so that the waves attenuate a little faster. Also, as usual, the above breaks down when the determinant in the denominator vanishes;

these singular curves are the 'caustics'. Table I shows the comparison of the turbulent and velocity propagators.

Although there are two different Green's functions used to obtain v, the overall field is still symmetric with respect to the same generator $\mathbf{G}_{st}$, and the structure function exponent $\xi(q)$ is also unchanged. In addition, the spatial $||\Delta \mathbf{r}||$ – which is the basic physical scale function – can (if necessary) be the same for both g_ε and g_v; it is only the space–time scale function which need be different. Appendix C gives some more information on the statistical properties of the resulting turbulence/wave model.

4.3. Gravito-turbulence dispersion relations

The standard gravity wave model assumes for a layer thickness Δz, a uniform stratification characterized by $N^2 = g(\Delta \log \theta / \Delta z)$, and weak enough nonlinearity so as to allow a linear perturbation analysis. The resulting Taylor–Goldstein equations then lead to the dispersion relation Equation (1). (We discuss only its simplest form which is quite adequate for our purpose.) In contrast, our turbulence flux-based approach assumes a highly heterogeneous vertical structure whose statistics are controlled by the (large-scale averaged) buoyancy variance flux

$$\phi = g^2 \left\{ \frac{(\Delta \log \theta)^2}{\tau_b} \right\}_l,$$

via its effect on l_s (the subscript indicates that the flux is measured at space–time resolution l; see appendix A). The combined ε, ϕ fluxes lead to a physical scale function $||\Delta \mathbf{r}||$, and thus to the dispersion relation:

$$\omega_d(\mathbf{k}) = ||\mathbf{k}||^{H_t}; \quad ||\mathbf{k}||^{-H} \xleftrightarrow{\text{F.T.}} ||\mathbf{r}||^{-(D_s - H)}. \quad (46)$$

However, the scale function is fairly general. For example, considering only the vertical (x, z) plane, it is of the form:

$$
\begin{aligned}
||\mathbf{k}|| &= \tilde{\Theta}(\theta_p) ||\mathbf{k}||_{\text{can}}; \\
||\mathbf{k}||_{\text{can}} &= l_s^{-1} \{(k_x l_s)^2 + |k_z l_s|^{2/H_z}\}^{1/2}, \qquad (47)
\end{aligned}
$$

Table II. Comparison of the standard gravity wave dispersion relations, with a turbulent/wave model with a gravity wave-like choice of $\tilde{\Theta}(\theta_p)$: a 'gravito-turbulence' dispersion relation. To make the comparison clearer, we have expressed the flux ϕ in terms of the potential temperature θ and g. l_s is the sphero-scale.

	Linear theory gravity wave dispersion	'Gravito-turbulent' dispersion
General form	$\omega(\mathbf{k}) \approx g^{1/2} \left(\dfrac{\Delta \log \theta}{\Delta z}\right)^{1/2} \dfrac{|k_x|}{|k|};$ $|k| = \left(k_x^2 + k_z^2\right)^{1/2}$	$\omega(\mathbf{k}) = \varepsilon^{1/3} \dfrac{|k_x|}{||\mathbf{k}||^{1/3}};$ $||\mathbf{k}|| = l_s^{-1}\{(l_s k_x)^2 + (l_s k_z)^{18/5}\}^{1/2};$ $l_s = \phi^{-3/4}\varepsilon^{5/4}$
Near-horizontal propagation	$\omega(\mathbf{k}) \approx g^{1/2} \left(\dfrac{\Delta \log \theta}{\Delta z}\right)^{1/2} \dfrac{|k_x|}{|k_z|};$ $|k_x| \ll |k_z|$	$\omega(\mathbf{k}) = g^{2/5} \left[\overline{\dfrac{(\Delta \log \theta)^2}{\tau_b}}\right]^{1/5}_{||\Delta x, \Delta z||} \dfrac{|k_x|}{|k_z|^{3/5}};$ $|k_x| \ll |k_z|^{9/5} l_s^{4/5}$
Near-vertical propagation	$\omega(\mathbf{k}) \approx g^{1/2} \left(\dfrac{\Delta \log \theta}{\Delta z}\right)^{1/2};$ $|k_x| \gg |k_z|$	$\omega(\mathbf{k}) = \varepsilon^{1/3}_{||\Delta x, \Delta z||} |k_x|^{2/3};$ $|k_x| \gg |k_z|^{9/5} l_s^{4/5}$

where $\tilde{\Theta}(\theta_p)$ is a relatively arbitrary function of direction (polar angle, θ_p) in the vertical plane (see Equation (16)).

Several of the predictions of gravity wave theory have been at least roughly empirically verified; it is therefore of interest to choose $\tilde{\Theta}(\theta_p)$ so that the turbulence/wave theory gives a similar dispersion relation and hence gives similar predictions. Since the classical dispersion relation is symmetric with respect to isotropic scale changes, i.e. with x, z generator

$$\mathbf{G} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ rather than the anisotropic}$$

$$\mathbf{G} = \begin{pmatrix} 1 & 0 \\ 0 & H_z \end{pmatrix},$$

the two dispersion relations cannot be identical. However, they can be chosen to be sufficiently similar so that the new relation can plausibly be compatible with the results of previous atmospheric gravity wave studies.

Using the values $H_t = 2/3$, $H_z = 5/9$ from the dimensional analysis, we find that the choice $\tilde{\Theta}(\theta_p) = (\cos \theta_p)^{3/2}$ leads to the following gravity wave-like 'gravito-turbulent' dispersion relation:

$$\omega(\mathbf{k}) = \varepsilon^{1/3} \frac{|k_x|}{||\mathbf{k}||^{1/3}}. \tag{48}$$

To display the similarity with the classical dispersion relation more clearly, Table II shows the two special cases which are most commonly tested empirically: near-horizontal and near-vertical propagation.

It can be seen that in both cases, for horizontal propagation, the dispersion relation becomes linear (or near linear) in k_x so that the horizontal group velocity is independent of k_x, i.e. it 'saturates'. In addition, the dependencies on $(\Delta \log \theta)$ are very similar (a 2/5 power instead of a power) 1/2 although it should be recalled that in the turbulence case, the potential temperature

profile is considered highly variable, not linear. Also, for near-vertical propagation, in both cases ω is independent of k_z. A final physically significant similitude is the fact that, in both cases, the group velocity has a 'restoring' vertical component i.e. w_g is opposite in sign to ω/k_z so that for example if the wave front is propagating upwards, then the wave energy propagates downwards (in the absence of a mean advection; Nappo, 2002). The comparison of the group velocities is shown in Figure 3. In Figure 4 we show (x, z) and (t, z) sections of (x, z, t) numerical multifractal simulations showing the stratified wave-like structures that the model produces, including in the presence of overall advection. (These models were actually for a passive scalar, produced by replacing $\varepsilon^{1/3}$ by $\varepsilon^{-1/6}\chi^{1/2}$, where χ is the passive scalar flux, see parts II, III). In Figure 5 we show a time sequence, and in Figure 6 we show the effect of changing the sphero-scale and the vertical wind; all these use the grativo-turbulence dispersion/scale function). Finally in Figure 7 we show simulations of horizontal sections with varying scale function/dispersion relations (by changing the shape of the unit ball B_1 via the function Θ defining the unit vectors), showing that quite realistic morphologies are readily produced. In appendix C, we derive various properties of the turbulence/wave model including its space–time energy spectrum.

4.4. Intermediate 'leakage' models, energy transport and the self-consistency of the turbulence/wave FIF models

We have presented a model in which the space–time propagator corresponds to a fractional integral over waves with a nonlinear turbulent dispersion defined via the physical scale function. The implications for energy transport are very strong: the turbulent energy flux input will generally be transported far from the source via the
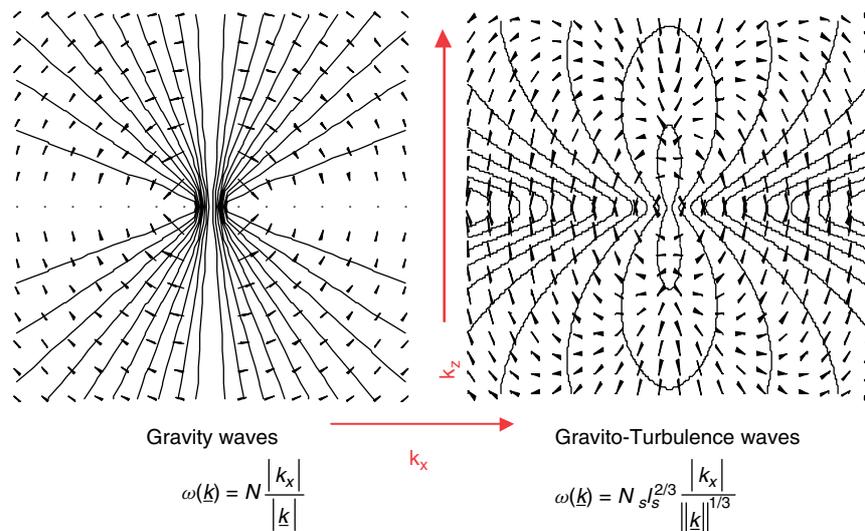

Gravity waves

$$\omega(\underline{k}) = N \frac{|k_x|}{|\underline{k}|}$$

Gravito-Turbulence waves

$$\omega(\underline{k}) = N_s l_s^{2/3} \frac{|k_x|}{\|\underline{k}\|^{1/3}}$$

Figure 3. Contour lines of frequency, ω, and the corresponding gradients (group velocities, arrows). The formula for the gravito-tubulence dispersion waves is the same as in Table II, with $N_s = \omega_s = (\phi/\varepsilon)^{1/2}$. Note that $N_s l_s^{2/3} = \phi^{1/5}$. This figure is available in colour online at www.interscience.wiley.com/qj
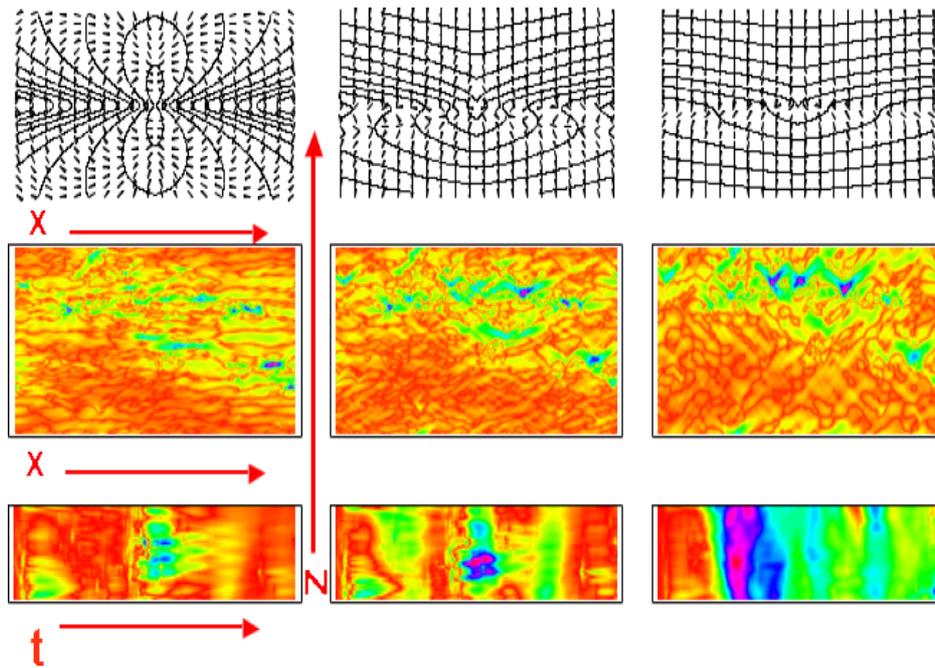
Figure 4. A multifractal simulation of a passive scalar in (x,z,t) space with the observed multifractal parameters ($\alpha = 1.8$, $C_1 = 0.05$, see parts II, III) and theoretical values $H_t = 2/3$, $H_z = 5/9$. The simulations show the vertical wind increasing from 0 (left) to 0.25 to 0.5 pixels per time step. (Only a single time step is shown.) The top row shows the dispersion relation and group velocity, the second is an (x,z) cross-section while the third row is a (t,z) cross-section. The numerical simulation techniques are based on those described in Schertzer and Lovejoy (1987), Wilson *et al.* (1991), Pecknold *et al.* (1993), Marsan *et al.* (1996). This figure is available in colour online at www.interscience.wiley.com/qj
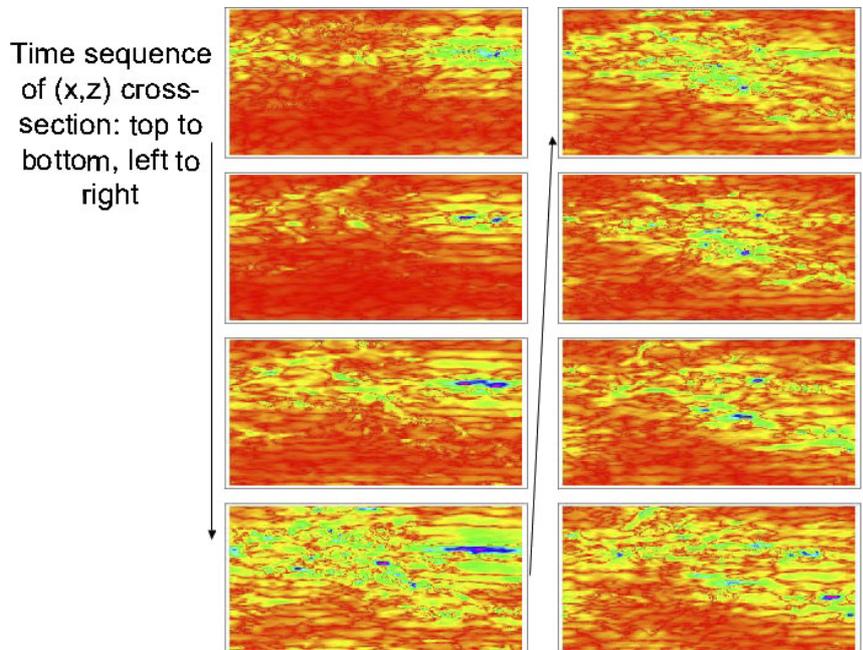
Figure 5. Eight time steps in the evolution of the vertical (x, z) cross-section of a passive scalar component from the zero-wind case of Figure 4. The structures are increasingly stratified at larger and larger scales and displays wave phenomenology. This figure is available in colour online at www.interscience.wiley.com/qj

waves. Since the energy flux is related to the velocity differences via $\varepsilon \approx \Delta v^3 / \Delta x$, most of the energy must remain localized for the model to be self-consistent. One way of achieving this energy localization would be if the dispersion relation had a negative imaginary part. However this would imply a dissipation mechanism which – if too strong – would contradict the picture of a cascade of conservative fluxes upon which the FIF is built. A more satisfying method is to combine the wave with the turbulent scale functions so that the final model
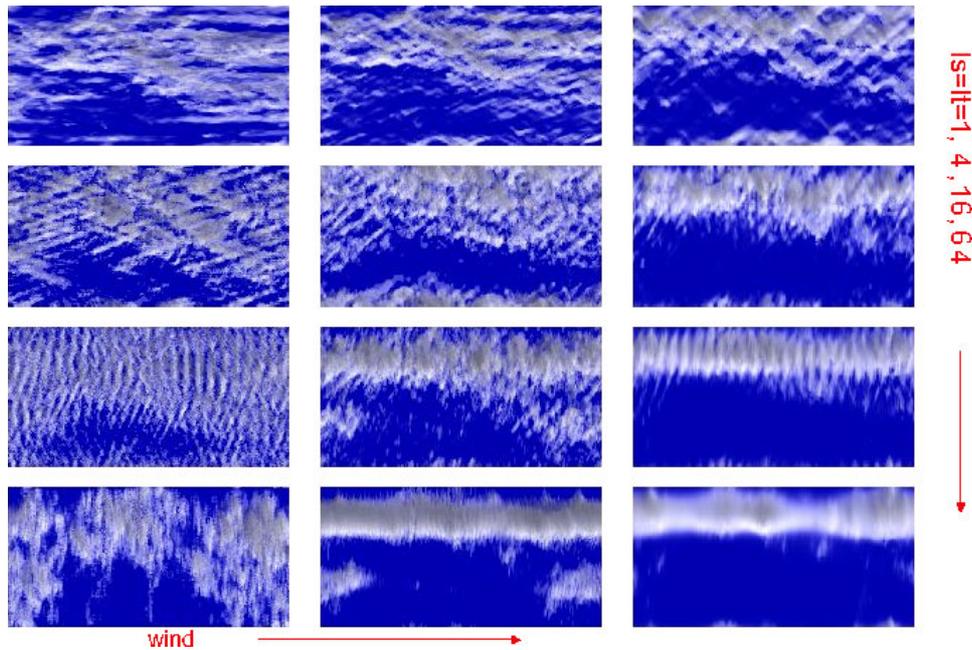
Figure 6. The effect of a constant vertical wind (left to right 0, 0.25, 0.5 pixels per time step), and sphero-scale increasing top to bottom from 1 pixel to 4, 16, 64 pixels for vertical cross-sections of simulated passive scalar, for the gravito-turbulence dispersion relation described in the text. In order to visualize the time evolution, the successive vertical sections are 'stacked' on top of each other; the rendition uses simulated single-scattering visible radiation through the stack (as though the time dimension is the 'depth' of the cloud). This figure is available in colour online at www.interscience.wiley.com/qj

has aspects of both; in this case, the wave energy could considered as 'leakage' in analogy with the Lumley–Shur model. A simple way to achieve this is to use:

$$\widetilde{g}_v(\mathrm{k}, \omega) = \left[\overbrace{h(t) [\![\mathbf{r}, t]\!]^{-(D-H_{\mathrm{tur}})}} \right]$$
$$\times \left[\mathrm{i} \left(\omega - ||\mathbf{k}||^{H_t} \right)^{-1/H_t} \right]^{H_{\mathrm{wav}}};$$
$$H_{\mathrm{tur}} + H_{\mathrm{wav}} = H. \qquad (49)$$

We see that the extreme localized and extreme unlocalized models correspond to $H_{\mathrm{tur}} = H$, $H_{\mathrm{wav}} = 0$ and $H_{\mathrm{wav}} = H$, $H_{\mathrm{tur}} = 0$ respectively (recall $H = 1/3$). In Figure 8 we show the effect of increasing H_{wav}; one can see how structures become progressively more and more wave-like while retaining the same scaling symmetries, close to observations.

5. Conclusions

One of the most fundamental unsolved problems in atmospheric science is to understand the nature of atmospheric stratification over wide ranges of space–time scales, i.e. its strong vertical anisotropy. Since the 1980s, in the face of mounting evidence that the vertical and horizontal scaling laws are different, the classical isotropic 3-D/isotropic 2-D model for atmospheric dynamics has been increasingly abandoned by the experimentalists. Indeed, state-of-the-art dropsonde analyses find no evidence of isotropic turbulence even in the lowest 158 m layer, down to 5 m in scales, while other analyses find a fractal hierarchy of alternating unstable layers within stable layers bringing into question the existence of 'homogeneous' layers (Lovejoy *et al.*, 2007, 2008c). This result, supported by the empirical results described in part II in this series, shows that this anisotropic scaling applies to scales as small as several metres. This suggests that, when we discretize the equations of the atmosphere at these scales (or larger), the isotropic gradient, divergence and Laplacian operators should be replaced by anisotropic 'effective' or 'renormalized' operators which will involve non-integer derivatives in either the vertical or horizontal directions (or both).

The need to develop anisotropic scaling models leads to several theories based on linear gravity waves; here we present a strongly nonlinear alternative – a turbulence flux-based model (the Fractionally Integrated Flux model) in which the horizontal dynamics are controlled by energy fluxes, and the vertical by buoyancy force variance fluxes. This model is based on the notion of physical scale – that the nonlinear turbulent dynamics determine the physically relevant scale; it is not imposed *a priori* (i.e. the classical (Euclidean) scale). The turbulence-based FIF model is more satisfying (and in parts II, III we show closer to the observations) – if only because the Reynold's numbers are far too large for linearizations to be plausible. Even at 20 m resolutions, dropsonde pairs show that the fluctuation Reynolds numbers are typically in the range 10^6–10^7. In part II we review some of this evidence and some of these arguments.

Although the FIF model has few restrictions, initial implementations focused on the special case in which structures were localized in both space and in space–time. This implied that they lacked wave
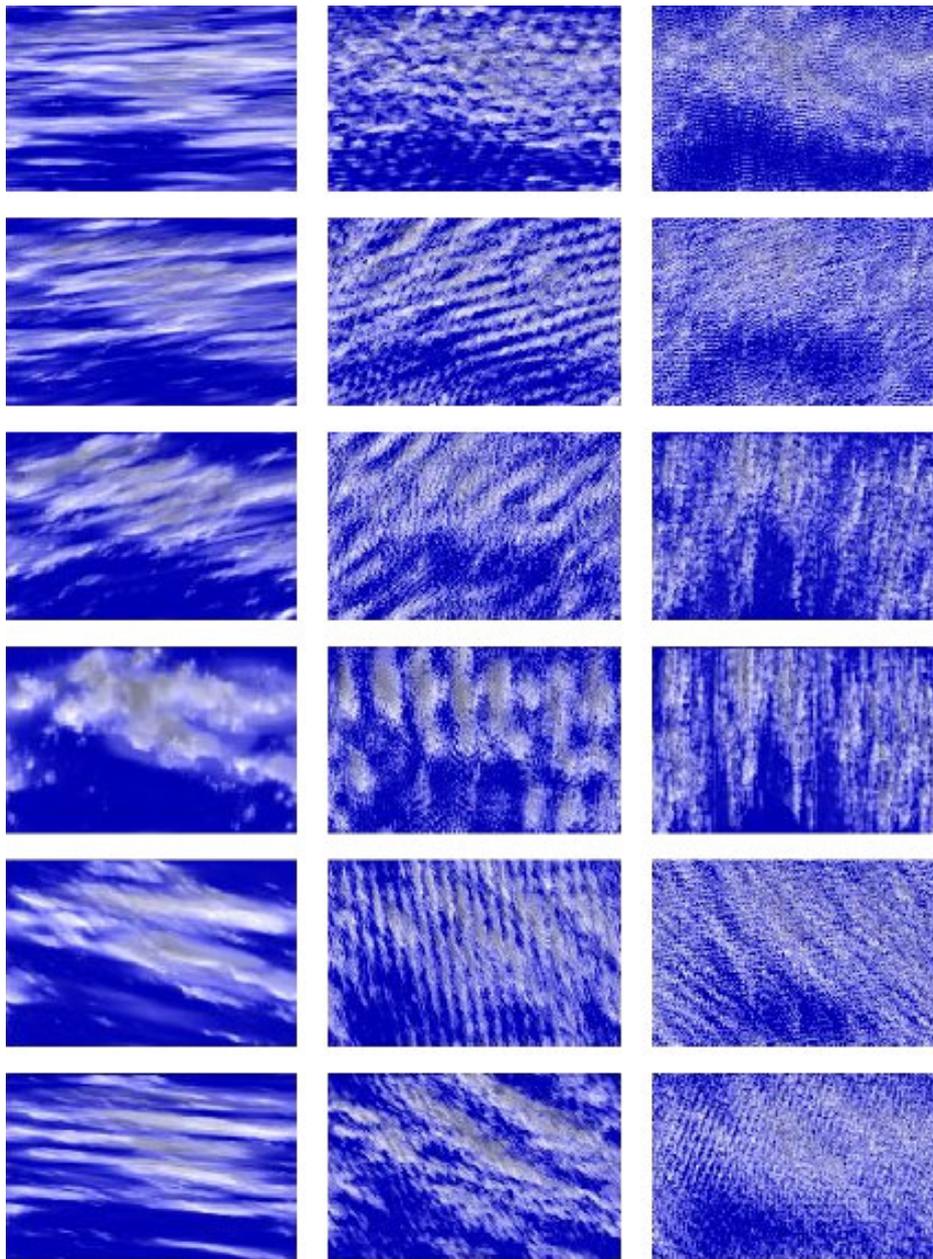
Figure 7. A series of horizontal sections of (x,y,t) passive scalar cloud simulations with horizontal generator $\mathbf{G} = \begin{pmatrix} 1.2 & 0.05 \\ -0.05 & 0.8 \end{pmatrix}$, with $H_t = 2/3$ as usual. From left to right, the horizontal sphero-scale $= 1, 8, 64$ pixels. The horizontal unit ball is characterized by $\Theta(\theta_p) = 1 + a\cos(2\theta_p - 2\theta_0)$, with $a = 0.65$, and from top to bottom the orientation θ_0 is varied from 0 to $5\pi/6$. (This is the real (x,y) space function; cf. Equation (16).). These simulations show how sensitive the morphologies are to the unit balls (i.e the spatial scale function/dispersion relation). This figure is available in colour online at www.interscience.wiley.com/qj

phenomenologies; numerical multifractal modelling confirmed this feature.

The key result in this paper is to show that the FIF framework is wide enough not only to include the wave effects, but also to accommodate dispersion relations close to the standard gravity wave dispersion relations. The model can thus potentially explain the empirical results in much of the atmospheric gravity wave literature. The key point is that the FIF requires two propagators (space–time Green's functions). The first determines the space–time structure of the cascade of fluxes; this must be localized in space–time in order to satisfy the

usual turbulence phenomenology. In contrast, the second propagator relates the turbulent fluxes to the observables; this propagator can still be localized in space but can be unlocalized in space–time. (The spatial part is the same as before; it is spatially localized in wave packets.)

This model is still very general – its main constraint is to produce stochastic realizations which respect the anistotropic, multifractal extensions of Kolmogorov's law (or in the case of passive scalars, Corrsin–Obukov laws). The turbulence-determined 'physical' scale function defines an anisotropic dispersion relation; by changing the scale function we can change the dispersion
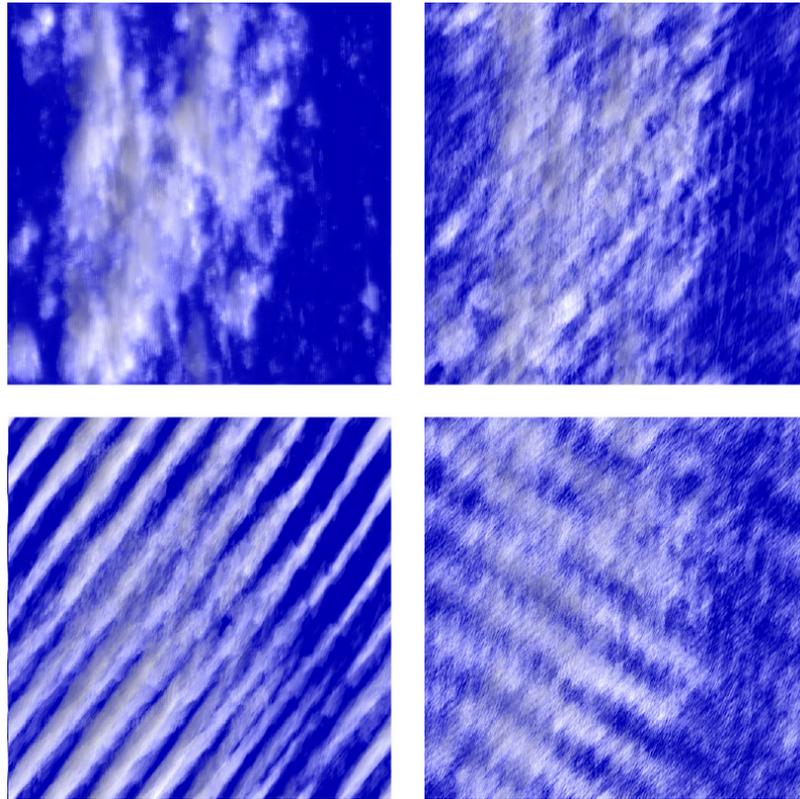
Figure 8. The effect of increasing H_{wav} with $H_{wav} + H_{tur} = H = 0.33$, $H_t = 0.66$; clockwise from the upper left we have $H_{wav} = 0$, 0.33, 0.52, 0.38 (i.e. $H_{tur} = 1/3 - H_{wav} = 0.33$, 0, -0.19, -0.05), $C_1 = 0.1$, $\alpha = 1.8$. There is a small amount of differential anisotropy characterized by $\mathbf{G} = \begin{pmatrix} 0.95 & -0.02 \\ 0.02 & 1.05 \end{pmatrix}$. The horizontal unit ball is characterized by $\Theta(\theta_p) = 1 + a\cos(2\theta_p - 2\theta_0)$, with $a = 0.65$ and $\theta_0 = 0$. The random seed is the same in all cases, so that one can see how structures become progressively more and more wave-like while retaining the same scaling symmetries, close to observations. This figure is available in colour online at www.interscience.wiley.com/qj

relation. In order to allow the model to account for the numerous observations of gravity waves, we show how a specific 'gravito-turbulence' dispersion relation can be chosen which has many of the key qualitative features of the classical linear dispersion relation. We show by numerical simulations that the new model does indeed generate wave-like phenomenologies. Finally, we show how to interpolate between the extreme wave and turbulence propagators so that any degree of localization/delocalization can be accommodated. In parts II and III (Lilley *et al.*, 2008; Radkevich *et al.*, 2008) we use high-powered lidar data of passive scalars to accurately confirm the overall FIF framework; however the observations are still not accurate enough to distinguish the various specific localized/unlocalized FIF models discussed here. This means that theory and numerical multifractal modelling will be particularly important.

Appendix A: Sphero-Scales, Sphero-Times and Intermittency

A.1. Discussion

The turbulent fluxes ε, ϕ are highly intermittent, presumably the result of (multifractal) cascade processes concentrating them into smaller and smaller regions of space; it is the statistics of these fluxes that determine the scale functions, sphero-scales and times l_s, τ_s; they 'emerge' from the statistics of the fluxes. If we consider applying Equation (10) to a given (space–time) region, then we must use appropriate space–time averages. In this appendix, we clarify this point and consider the effects of intermittency on the anisotropy exponents and on l_s, τ_s. In section 3.2 we considered constant advection; here we consider the advection-free formulae, and in appendix B we consider advection in connection with a statistically varying vertical wind.

Consider a cascade process starting at a horizontal scale L_x, vertical scale L_z with corresponding external time-scale T. (It is probably best to think of T as the typical time-scale of planetary-scale structures, i.e. as a consequence of the cascade, not as something given *a priori*, but this is not relevant for the argument below.) Let us consider taking data over a space–time region with horizontal scale l_x, vertical scale l_z taken over a time τ. (We ignore the y coordinate since we assume horizontal isotropy.) For simplicity, consider that the data region is the result of the cascade over a scale ratio λ:

$$\lambda = \frac{L_x}{l_x}, \; \lambda_z = \lambda^{H_z} = \frac{L_z}{l_z}, \; \lambda_t = \lambda^{H_t} = \frac{T}{\tau}. \tag{50}$$

(Due to the anisotropy, the ratio is different along the z, t axes as indicated.) Due to the multiplicative nature of the cascade, the effect of the cascade from the

largest (L_x, L_z, T) scales down to (l_x, l_z, τ) scales is to multiplicatively modulate the (l_x, l_z, τ) region by constant large-scale fluxes which we will denote $\varepsilon_\lambda, \phi_\lambda$; λ is the dimensionless reduction factor L/l in the notation of section 3.3. We can thus estimate the qth-order velocity structure functions averaged over the (l_x, l_z, τ) region:

$$
\left.
\begin{aligned}
&(|\Delta v(\Delta x, 0, 0, 0)|^q)_{(l_x,l_z,\tau)} \\
&= \varepsilon_\lambda^{q/3}\left(\frac{l_x}{\Delta x}\right)^{K_\varepsilon(q/3)}\Delta x^{q/3} = \varepsilon_\lambda^{q/3}l_x^{K_\varepsilon(q/3)}\Delta x^{\xi_h(q)}; \\
&(|\Delta v(0, 0, \Delta z, 0)|^q)_{(l_x,l_z,\tau)} \\
&= \phi_\lambda^{q/5}\left(\frac{l_z}{\Delta z}\right)^{K_\phi(q/5)}\Delta z^{3q/5} = \phi_\lambda^{q/5}l_z^{K_\phi(q/5)}\Delta z^{\xi_v(q)}; \\
&(|\Delta v(0, 0, 0, \Delta t)|^q)_{(l_x,l_z,\tau)} \\
&= \varepsilon_\lambda^{q/2}\left(\frac{\tau}{\Delta t}\right)^{K_\varepsilon(q/2)}\Delta t^{q/2} = \varepsilon_\lambda^{q/2}\tau^{K_\varepsilon(q/2)}\Delta t^{\xi_t(q)}.
\end{aligned}
\right\}.
$$
(51)

Note that, in this appendix, when needed we use the subscript ε on $K(q)$ in order to avoid confusion. The structure function exponents are

$$
\begin{aligned}
\xi_h(q) &= \frac{q}{3} - K_\varepsilon\left(\frac{q}{3}\right); \\
\xi_v(q) &= \frac{3q}{5} - K_\phi\left(\frac{q}{5}\right); \\
\xi_t(q) &= \frac{q}{2} - K_\varepsilon\left(\frac{q}{2}\right),
\end{aligned}
$$
(52)

where, we have added the ε index for the energy flux exponents. Note that in the above we assume a single cascade realization (the structure functions in Equation (51) only involve space-time averages, not ensemble averages); however, we have used the multiplicative property to factor out the low-frequency part of the cascade $(\varepsilon_\lambda, \phi_\lambda)$, and approximate the space–time averaging of the high frequencies using ensemble averages. An exact result would be obtained if we ensemble averaged both sides of Equations (51).

The hypothesis of anisotropic scaling means that when the equations (51) are expressed in terms of scale functions, they should be equivalent. This means that if we take

$$
\Delta x = l, \quad \Delta z = l_s\left(\frac{l}{l_s}\right)^{H_z}, \quad \Delta t = \tau_s\left(\frac{l}{l_s}\right)^{H_t}, \quad (53)
$$

so that: $l = [\![(\Delta x, 0, 0, 0)]\!] = [\![(0, 0, \Delta z, 0)]\!] = [\![(0, 0, 0, \Delta t)]\!] = [\![\Delta \mathbf{R}]\!]$, then all the left-hand sides of Equation (51) become $(|\Delta v(\Delta \mathbf{R})|^q)_{(l_x,l_z,\tau)} = \varepsilon_\lambda^{q/3}l_x^{K_\varepsilon(q/3)}l^{\xi_h(q)}$. Expressing the right-hand sides of Equation (51) in terms of l (using Equation (53)) and equating the exponents of l we obtain

$$
H_z = \frac{\xi_h(q)}{\xi_v(q)}; \quad H_t = \frac{\xi_h(q)}{\xi_t(q)}. \quad (54)
$$

This is equivalent to:

$$
H_z = \frac{5}{9} + \Delta H_z;
$$

$$
\Delta H_z = \frac{5}{9}\left\{\frac{\frac{5}{3q}K_\phi\left(\frac{q}{5}\right) - \frac{3}{q}K_\varepsilon\left(\frac{q}{3}\right)}{1 - \frac{5}{3q}K_\phi\left(\frac{q}{3}\right)}\right\}. \quad (55)
$$

Similarly, for horizontal/time cross-sections we obtain

$$
H_t = \frac{2}{3} + \Delta H_t; \quad \Delta H_t = \frac{2}{3}\left\{\frac{\frac{2}{q}K_\varepsilon\left(\frac{q}{2}\right) - \frac{3}{q}K_\varepsilon\left(\frac{q}{3}\right)}{1 - \frac{2}{q}K_\varepsilon\left(\frac{q}{2}\right)}\right\}. \quad (56)
$$

Using data from dropsonde pairs to estimate the universal multifractal parameters, we find $C_{1\varepsilon} \approx 0.50, \alpha_\varepsilon \approx 1.6,$ $C_{1\phi} \approx 0.65, \alpha_\phi \approx 1.8,$ (Lovejoy *et al.*, 2007) gives more information about this experiment). Figure A.1 shows the variations in the exponents $\Delta H_z(q), \Delta H_t(q)$ over the range $0 < q < 3$. (Since the dropsondes estimate the exponents in the vertical, the above exponents had to be used to estimate their horizontal counterparts; to first order this may be done by multiplying the C_1 values by $H_z = 5/9$). Note that, for $\alpha \geq 1$, the limit as $q \to 0$ of $K(q)/q$ is non-zero, finite. We see that the variation is fairly small, so that the non-intermittent estimates $H_z = 5/9$, $H_t = 2/3$ should not be so bad. In addition, we see that within the data region, since the qth-order statistics are dominated by structures with well-defined orders of singularities (intensities), we have both a small variation in H_z with intensity, but also a small variation in l_s, τ_s with the intensity of the structures. In part II, we determine ΔH_z directly from vertical cross-sections of passive scalars and find agreement to within ± 0.02.

Since in multifractals there is a precise one-to-one correspondence between structures (singularities) and qth-order statistical moments, these equations combined with the corrections for the exponents ΔH_z, ΔH_t show how the more and more intense structures (larger and larger q) change their anisotropy with scale. Since a smaller H_z means that stratification increases more rapidly as a function of scale, we see that there is a small tendency for the vertical stratification to increase for intense structures whereas the space–time stratification tends to decrease them. We see in part II that direct estimates of ΔH_z for passive scalars are fairly close to the ΔH_z curve in Figure A.1, although with smaller variation.

In order to get the full picture, we must consider also the change in sphero-scale and sphero-time. Using the above equations for the exponents, we can solve the x, z equation for l_s, and then the x,t equation for τ_s. Solving for l_s, we obtain

$$
l_s = l_{s,b}\left(\frac{l^*}{l_{s,b}}\right)^{\frac{\Delta K}{1+\Delta K}}; \quad l_{s,b} = \frac{\varepsilon_\lambda^{5/4}}{\phi_\lambda^{3/4}}, \quad (57)
$$

where

$$
l^* = \left(\frac{l_x^{K(q/3)}}{l_z^{K_\phi(q/5)}}\right)^{\frac{15}{4q\Delta K}};
$$

$$
\Delta K = \frac{15}{4q}\left\{K_\varepsilon\left(\frac{q}{3}\right) - K_\phi\left(\frac{q}{5}\right)\right\}. \quad (58)
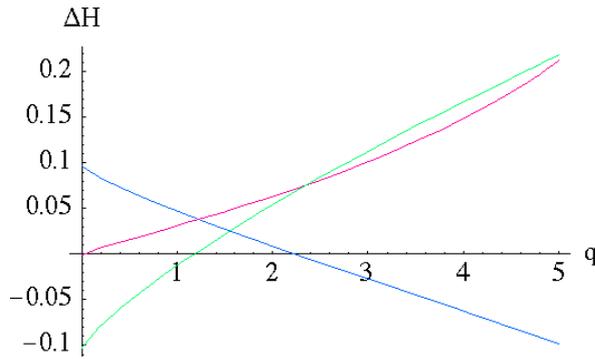$$

Figure A.1. The intermittency corrections as estimated from dropsonde pairs at 20 m resolution in the vertical and (roughly) the same in the horizontal. (Vertical parameters are $C_{1\varepsilon} \approx 0.50$, $\alpha_\varepsilon \approx 1.6$, $C_{1\phi} \approx 0.65$, $\alpha_\phi \approx 1.8$) ΔH_z (from Equation (55), top at left), ΔH_t (from Equation (56), middle at left), and ΔK (from Equation (58), bottom at left).

We can see see that, since ΔK appears to change sign around $q = 1$ (Figure A.1), the more intense structures (moments larger than this critical value) will have smaller sphero-scales than weaker structures.

The explicit equation for τ_s is complicated; it is more convenient to use implicit relations obtained by introducing the nondimensional quantities:

$$l'_x = \frac{l_x}{l_s}; \quad l'_z = \frac{l_z}{l_s}; \quad \tau' = \frac{\tau}{\tau_s}. \quad (59)$$

We obtain the following (implicit) equations:

$$l_s = l_{s,b} \left(\frac{l_x'^{K_\varepsilon(q/3)}}{l_z'^{K_\phi(q/5)}} \right)^{\frac{15}{4q}}, \quad (60)$$

$$\tau_s = \tau_{s,b} \left(\frac{l_x'^{\frac{9}{2q}K_\varepsilon(q/3)}}{l_z'^{\frac{5}{2q}K_{phi}(q/5)} \tau'^{\frac{2}{q}K_\varepsilon(q/2)}} \right);$$

$$\tau_{s,b} = \varepsilon_\lambda^{-1/3} l_{s,b}^{2/3} = \frac{\varepsilon_\lambda^{1/2}}{\phi_\lambda^{1/2}}. \quad (61)$$

When the intermittency is vanishingly small ($K \to 0$), we see that the above reduce to the values l_{sb}, τ_{sb}, The subscript b is for 'bare' quantities; they only take into account the large scales down to the scale of the data region.

A.2. Ensemble averages

If we consider averages over an ensemble of realizations, then we must use

$$\langle \varepsilon_\lambda^q \rangle = \varepsilon_1^q \lambda^{K_\varepsilon(q)},$$

$$\langle \phi_\lambda^q \rangle = \phi_1^q \lambda^{K_\phi(q)}, \quad (62)$$

where ε_1, ϕ_1 are the mean outer scale (scale ratio $\lambda = 1$) fluxes. This means that the formulae for l_s, τ_s can

be obtained from Equations (57)–(61) if we make the replacements

$$l_x \to \lambda l_x; \quad l_z \to \lambda_z l_z; \quad \tau \to \lambda_t \tau$$

$$\varepsilon_\lambda^q \to \varepsilon_1^q \lambda^{K_\varepsilon(q)}; \quad \phi_\lambda^q \to \phi_1^q \lambda^{K_\phi(q)}. \quad (63)$$

A.3. Passive scalars

Since we analyse passive scalar surrogates (lidar backscatter), we require the passive scalar equations corresponding to the above. Introducing the space–time scale function, we obtain

$$\langle |\Delta \rho(\Delta \mathbf{R})|^q \rangle = \left\langle \varphi_{[\![\Delta \mathbf{R}]\!]}^{q/3} \right\rangle [\![\Delta \mathbf{R}]\!]^{q/3};$$

$$\varphi = \chi^{3/2} \varepsilon^{-1/2}, \quad (64)$$

where χ is the passive scalar variance flux which is (scale by scale) conserved by the equations of passive scalar advection. Since χ, ε are presumably statistically coupled (in a non-trivial way), it is convenient to introduce the new 'flux' φ, although strictly speaking it is not expected to be conserved. (i.e. whereas $K_\chi(1) = K_\varepsilon(1) = 0$, we expect $K_\varphi(1) \neq 0$, although it is small). We now specialize the equation to purely horizontal, vertical and temporal lags, obtaining in analogy to Equation (51),

$$\left(|\Delta \rho(\Delta x, 0, 0, 0)|^q \right)_{(l_x, l_z, \tau)} = \varphi_\lambda^{q/3} \left(\frac{l_x}{\Delta x} \right)^{K_\varphi(q/3)} \Delta x^{q/3}$$

$$= \varphi_\lambda^{q/3} l_x^{K_\varphi(q/3)} \Delta x^{\xi_h(q)},$$

$$\left(|\Delta \rho(0, 0, \Delta z, 0)|^q \right)_{(l_x, l_z, \tau)} = \kappa_\lambda^{q/5} \left(\frac{l_z}{\Delta z} \right)^{K_\kappa(q/5)} \Delta z^{3q/5}$$

$$= \kappa_\lambda^{q/5} l_z^{K_\kappa(q/5)} \Delta z^{\xi_v(q)},$$

$$\left(|\Delta \rho(0, 0, 0, \Delta t)|^q \right)_{(l_x, l_z, \tau)} = \chi_\lambda^{q/2} \left(\frac{\tau}{\Delta t} \right)^{K_\chi(q/2)} \Delta t^{q/2}$$

$$= \chi_\lambda^{q/2} \tau^{K_\chi(q/2)} \Delta t^{\xi_\tau(q)}, \quad (65)$$

$$\xi_{h,\rho}(q) = q/3 - K_\varphi(q/3);$$

$$\xi_{v,\rho}(q) = 3q/5 - K_\kappa(q/5);$$

$$\xi_{\tau,\rho}(q) = q/2 - K_\chi(q/2);$$

where we have introduced the product of conserved fluxes $\kappa = \chi^{5/2} \varepsilon^{-5/2} \phi$. With these definitions, we see that passive scalars yield identical equations for the intermittency effects, on condition that in the x equation we make the replacement $\varepsilon \to \varphi$, in the z equation $\phi \to \kappa$ and in the time equation $\varepsilon \to \chi$ so that we obtain

$$H_{z,\rho} = \xi_{h,\rho}(q)/\xi_{v,\rho}(q),$$

$$H_{t,\rho} = \xi_{h,\rho}(q)/\xi_{\tau,\rho}(q). \quad (66)$$

Due to basic dimensional analysis, the bare sphero-scale and time-scales (which are good approximations) are

Appendix B: The effect of vertical velocity: effective generators effective temporal scaling exponents

The application of the advection formula (Equation (27)) is straightforward for the generally large, quasi-constant horizontal velocity, but is problematic for the vertical velocity which is always close to zero and which is likely to be scale dependent. In section 3.5 we discussed the implications in the case where the intermittency could be neglected. Here, we rederive the result including intermittency.

Let us assume no horizontal advection and neglect the 'pure temporal development' term (i.e. assume that τ_s is sufficiently large) so that

$$[\![\Delta t]\!] = l_s \left(\frac{|(w(\Delta t))_l| \Delta t}{l_s} \right)^{1/H_z},$$

$$[\![\Delta t]\!] = [\![(0, 0, 0, \Delta t)]\!], \tag{67}$$

$$|(w(\Delta t))_l| = w_l \left(\frac{[\![\Delta t]\!]}{l} \right)^{H_w - \gamma_w}, \tag{68}$$

where we have used the lighter notation $w_l = |(w(l))_l|$ for the absolute fluctuation in the vertical velocity over distances l spatially averaged over scales l. We have followed appendix A in assuming that the data region size (l_x, l_z, τ), is an (anisotropic) reduction by factor λ of the largest (outer) cascade scale, and have characterized the scale of the region by the effective scale l; this will be roughly the maximum scale of vectors within the region. Equation (67) follows from Equation (27) with $\Delta x = \Delta y = \Delta z = u = v = 0$. Equation (68) follows from an application of Equation (24b) using a vertical velocity structure function, ξ_w in place of ξ, and recognizing that if $H_w < 0$, $|(w(\Delta t))_l|$ should be used in place of. $(|\Delta w|)_l$. The non-random H_w is introduced because the effect of a fractional integration/differentiation of order H_w is to shift the singularities as indicated; e.g. Schertzer *et al.* (1997a). Solving for $[\![\Delta t]\!]$, following the discussion in section 3.5 (without intermittency), we eliminate $|(w(\Delta t))_l|$ between Equations (67), (68) and obtain:

$$[\![\Delta t]\!] = l \left(\frac{\Delta t}{\tau_s^*} \right)^{1/H_t^*}; \quad \tau_s^* = \frac{l_s}{w_l} \left(\frac{l}{l_s} \right)^{H_z};$$

$$H_t^* = H_z - H_w + \gamma_w. \tag{69}$$

Hence for a given (random) vertical velocity (random singularity γ):

$$\left(|\Delta v(0, 0, 0, \Delta t)|^q \right)_l = \varepsilon_\lambda^{q/3} l^{K_\varepsilon(q/3)} [\![\Delta t]\!]^{\xi_h(q)}$$

$$= \varepsilon_\lambda^{q/3} l^{q/3} \left(\frac{\Delta t}{\tau_s^*} \right)^{\xi_h(q)/H_t^*}. \tag{70}$$

We now suppose that $(H_w - \gamma_w)/H_z \ll 1$ (the mean singularities in the velocity field are of the order of 0.1 and we are assuming that $H_w \approx -0.1$ so that this is probably not a bad assumption) and average over the vertical velocity fluctuations:

$$\left\langle \left(\frac{\Delta t}{\tau_s^*} \right)^{\xi_h(q)/H_t^*} \right\rangle_w \approx \left\langle \left(\frac{\Delta t}{\tau_s^*} \right)^{\frac{\xi_h(q)}{H_z} \left(1 + \frac{(H_w - \gamma_w)}{H_z} \right)} \right\rangle_w$$

$$\approx \left(\frac{\Delta t}{\tau_s^*} \right)^{\frac{\xi_h(q)}{H_z} + \xi_w \left(\frac{\xi_h(q)}{H_z^2} \right)} \approx \left(\frac{\Delta t}{\tau_s^*} \right)^{\frac{\xi_h(q)}{H_t'}};$$

$$H_t' = \frac{H_z}{1 + \frac{H_z}{\xi_h(q)} \xi_w \left(\frac{\xi_h(q)}{H_z^2} \right)} \tag{71}$$

where $\xi_w(q) = qH_w = K_w(q)$ is the structure function exponent for the vertical velocity. In Figure B.1 we show the resulting $\Delta H_t' = H_t' - 2/3$ for various plausible vertical velocity parameters. For $H_w \approx -0.08$, the correction for $q = 1$ is small.

Finally, we define τ_s' and H_t' via:

$$\left\langle (|\Delta v(0, 0, 0, \Delta t)|^q)_l \right\rangle = \varepsilon_\lambda^{q/3} l^{K_\varepsilon(q/3)} l_s^{\xi_h(q)} \left(\frac{\Delta t}{\tau_s'} \right)^{\xi_h(q)/H_t'} \tag{72}$$

and hence comparing this with expressions (70), (71), we obtain:

$$\tau_s' = \tau_s^* \left(\frac{l_s}{l} \right)^{H_t'} = \tau_s \left(\frac{v_s}{w_l} \right) \left(\frac{l}{l_s} \right)^{H_z - H_t'}, \tag{73}$$

where we have used $l_s = v_s \tau_s$. Since $H_z - H_t'$ is of the order of $5/9 - 2/3 = -1/9$, the term $(l/l_s)^{H_z - H_t'}$ will be of order unity, so that the change in τ_s will mainly reflect the (v_s/w_l) factor. This means that if $H_t' \approx H_t$, the effect of vertical velocity term is to replace τ_s by a factor (v_s/w_l) larger; in addition, if $\tau_s' < \tau_s$ then this term will dominate the pure temporal development term.

Transition times from horizontal wind domination to vertical wind domination occur for $\Delta t > \Delta t_{xz}$ with:

$$\Delta t_{xz} = \frac{u^2}{l_s^2} \tau_s'^3 = \frac{u^2}{\varepsilon} \left(\frac{\tau_s'}{\tau_s} \right)^3, \tag{74}$$
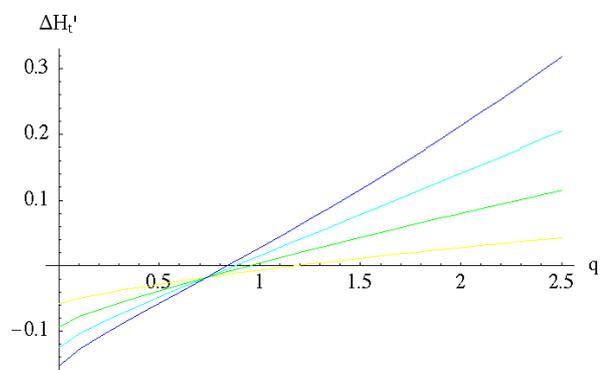

Figure B.1. $\Delta H_t' = H_t' - 2/3$ as a function of q. Assuming vertical velocity parameters $\alpha_w = 1.8$, $H_w = -0.08$, we have $C_{1w} = 0.025$, 0.05, 0.075, 0.1 (flatter to steeper).

(assuming $H_t' = 2/3$); so that the enhancement due to the vertical wind effect is:

$$\Delta t'_{xz} = \Delta t_{xt} \left(\frac{v_s}{w_l}\right)^3 \left(\frac{l_s}{l}\right)^{3(H_t'-H_z)}$$

$$= \frac{u^2}{\varepsilon} \left(\frac{v_s}{w_l}\right)^3 \left(\frac{l_s}{l}\right)^{3(H_t'-H_z)}. \quad (75)$$

In part III we argue that this effect is indeed present in some of the upward-pointing lidar sequences, and using meteorological analyses we show that it ought to dominate the horizontal advection effect with a reasonably high probability.

Appendix C: Some statistical properties of the turbulent/wave model

The basic real-space statistical properties of the turbulence/wave model follow as in the classical FIF (Schertzer *et al.* (1997a)); the fields obey Equation (22). The 1-D spectra then followed by Fourier transforming Equation (20) for $q = 2$ taking $\Delta\mathbf{R} = (\Delta x, 0, 0, 0)$, $\Delta\mathbf{R} = (0, \Delta y, 0, 0)$, $\Delta\mathbf{R} = (0, 0, \Delta z, 0)$, $\Delta\mathbf{R} = (0, 0, 0, \Delta t)$ for $E(k_x)$, $E(k_y)$, $E(k_z)$, $E(\omega)$, respectively. It is however instructive to derive the result by starting with the joint spectral density $P_v(\mathbf{k}, \omega)$:

$$P_v(\mathbf{k}, \omega) = \frac{P_\varepsilon(\mathbf{k}, \omega)}{\left|\omega - ||\mathbf{k}||^{H_t}\right|^{2H/H_t}};$$

$$P_\varepsilon(\mathbf{k}, \omega) = [\![(\mathbf{k}, \omega)]\!]_t^{-s}; \quad s = D_{st} - K\left(\frac{2}{3}\right), \quad (76)$$

where $[\![(\mathbf{k}, \omega)]\!]_t$ is the Fourier counterpart of the turbulent space−time scale function, 'symmetrized' by the modulus operation involved in calculating the spectral energy density, and $D = \mathrm{Tr}(\mathbf{G}) = 2 + H_z + H_t$ is the 'elliptical dimension' of space−time and $K_\varepsilon(2/3)$ is the intermittency correction for the flux, estimated empirically to be ≈ -0.07 (note the sign; e.g. Schmitt *et al.*, 1992); $K_\varepsilon(2/3)$ is the exponent of the second moment of v which is proportional to $\varepsilon^{2/3}$). To determine the 1-D spectra with respect to k_x, k_y, k_z, ω, we successively integrate P_v with respect to the remaining variables. The results are the classical Kolmogorov and Bolgiano−Obukov statistics with multifractal (intermittency) corrections:

$$E(k_x) \propto \varepsilon^{2/3} k_x^{-5/3} \left(\frac{k_x}{k_{x1}}\right)^{-K_\varepsilon(2/3)};$$

$$E(\omega) \propto \varepsilon\omega^{-2} \left(\frac{\omega}{\omega_1}\right)^{-K_\varepsilon(2/3)/H_t};$$

$$E(k_z) \propto \phi^{2/5} k_z^{-11/5} \left(\frac{k_z}{k_{z1}}\right)^{-K_\varepsilon(2/3)/H_z}. \quad (77)$$

We have ignored dimensionless numerical constants. The derivation of the spectra in this way has the advantage that the condition for the convergence is clear: we

must have $2H < H_t$ (due to the singular propagator; cf. Equations (40), (41)). When $2H > H_t$, Parseval's theorem proves that the real-space variance also diverges, hence we expect a divergence of second-order moments for v. This is indeed confirmed numerically. Interestingly, the borderline case $2H = H_t$ is the one apparently relevant (since $H = 1/3$, $H_t = 2/3$, at least for the extreme wave model presented here, see section 3.2.4), so that the question of divergence/convergence depends on more detailed properties of the propagator than those considered here. (There is empirical evidence that the 5th and 7th moments of v diverge in the vertical and horizontal directions respectively (Schertzer and Lovejoy, 1985a; Schmitt *et al.*, 1994; Lovejoy and Schertzer, 2007; and part III). Note that, even if the ensemble-averaged spectrum does indeed diverge, empirical spectra will converge; they will however display large realization to realization fluctuations – perhaps not unlike real radiosonde profiles. Finally, more general velocity propagators can be considered with weaker singularities (wave-like parts).

References

Adelfang SI. 1971. On the relation between wind shears over various intervals. *J. Atmos. Sci.* **10**: 138.

Allen SJ, Vincent RA. 1995. Gravity wave activity in the lower atmopshere: Seasonal and latitudinal variations. *J. Geophys. Res.* **100**: 1327–1350.

Ashkenazi S, Steinberg V. 1999. Spectra and statistics of velocity and temperature fluctuations in turbulent convection. *Phys. Rev. Lett.* **83**: 4760–4763.

Bleistein N, Handelsman RA. 1986. *Asymptotic expansions of integrals.* Dover: Mineola, New York.

Bolgiano R. 1959. Turbulent spectra in a stably stratified atmosphere. *J. Geophys. Res.* **64**: 2226.

Chigirinskaya Y, Schertzer D, Lovejoy S, Lazarev A, Ordanovich A. 1994. Unified multifractal atmospheric dynamics tested in the tropics. Part 1: Horizontal scaling and self organized criticality. *Nonlinear Proc. Geophys.* **1**: 105–114.

Dewan E. 1997. Saturated-cascade similtude theory of gravity wave sepctra. *J. Geophys. Res.* **102**: 29799–29817.

Dewan E, Good R. 1986. Saturation and the 'universal' spectrum vertical profiles of horizontal scalar winds in the stratosphere. *J. Geophys. Res.* **91**: 2742–2748.

Endlich RM, Singleton RC, Kaufman JW. 1969. Spectral analyses of detailed vertical wind profiles. *J. Atmos. Sci.* **26**: 1030–1041.

Feller W. 1971. *An introduction to probability theory and its applications, Vol. 2.* Wiley: New York.

Gardner C. 1994. Diffusive filtering theory of gravity wave spectra in the atmosphere. *J. Geophys. Res.* **99**: 20601–20622.

Ishihara T, Yoshida K, Kaneda Y. 2002. Anisotropic velocity correlation spectrum at small scales in a homogeneous turbulent flow. *Phys. Rev. Lett.* **88**: 154501–154504.

Kahane J.-P. 1985. Sur le chaos multiplicatif. *Annales Sciences Math. Québec* **9**: 435.

Kolesnikov VN, Monin AS. 1965. Spectra of meteorological field fluctuations. *Izvestiya, Atmospheric and Oceanic Physics* **1**: 653–669.

Kolmogorov AN. 1941. Local structure of turbulence in an incompressible liquid for very large Reynolds numbers. *Proc. Acad. Sci. URSS. Geochem. Sect.* **30**: 299–303. (English translation: *Proc. R. Soc. London A* **434**: 9–17, 1991).

Lazarev A, Schertzer D, Lovejoy S, Chigirinskaya Y. 1994. Unified multifractal atmospheric dynamics tested in the tropics: Part II. Vertical scaling and Generalized Scale Invariance. *Nonlinear Proc. Geophys.* **1**: 115–123.

Lilley M, Lovejoy S, Strawbridge K, Schertzer D. 2004. 23/9 dimensional anisotropic scaling of passive admixtures using lidar aerosol data. *Phys. Rev. E* **70**: DOI: 10.1103/70, 036307-1-7.

Lilley M, Lovejoy S, Strawbridge KB, Schertzer D, Radkevich A. 2008. Scaling turbulent atmospheric stratification. II: Spatial

stratification and intermittency from lidar data. *Q. J. R. Meteorol. Soc.* **134**: in press.

Lovejoy S, Schertzer D. 2006. Multifractals, cloud radiances and rain. *J. Hydrol.* **322**: 59–88. DOI: 10.1016/j.hydrol.2005.02.042.

Lovejoy S, Schertzer D. 2007. Scale, scaling and multifractals in geophysics: twenty years on. pp. 311–337 in *Nonlinear dynamics in geosciences.* Elsner J, Tsonis AA. (eds). Springer: New York.

Lovejoy S, Schertzer D, Stanway JD. 2001. Direct evidence of planetary scale atmospheric cascade dynamics. *Phys. Rev. Lett.* **86**: 5200–5203.

Lovejoy S, Schertzer D, Tuck AF. 2004. Fractal aircraft trajectories and nonclassical turbulent exponents. *Phys. Rev. E* **70**: DOI: 10.1103/70, 036306–1-5.

Lovejoy S, Tuck AF, Hovde SJ, Schertzer D. 2007. Is isotropic turbulence relevant in the atmosphere? *Geophys. Res. Lett.* **34**: L14802, DOI: 10.1029/2007GL029359.

Lovejoy S, Schertzer D, Allaire V. 2008a. The remarkable wide range spatial scaling of TRMM precipitation. *Atmos. Res.* in press.

Lovejoy S, Schertzer D, Allaire V, Bourgeois T, King S, Pinel J, Stolle J. 2008b. The remarkable wide range spatial scaling of atmospheric radiances. *Phys. Rev. Lett.* submitted.

Lovejoy S, Tuck AF, Hovde SJ, Schertzer D. 2008c. Do stable atmospheric layers exist? *Geophys. Res. Lett.* **35**: L01802, DOI: 10.1029/2007GL032122.

Lumley JL. 1964. The spectrum of nearly inertial turbulence in a stably stratified fluid. *J. Atmos. Sci.* **21**: 99–102.

Lumley JL. 1967. Rational approach to relations between motions of differing scales in turbulent flows. *Phys. Fluids* **10**: 1405.

L'vov VS, L'vov Y, Newell AC, Zakharov V. 1997. Statistical description of acoustic turbulence. *Phys. Rev. E* **56**: 390–405.

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

Marsan D, Schertzer D, Lovejoy S. 1996. Causal space-time multifractal processes: Predictability and forecasting of rain fields. *J. Geophy. Res.* **31D**: 26333–26346.

Miller KS, Ross B. 1993. *An introduction to the fractional calculus and fractional differential equations.* John Wiley and Sons: New York.

Nappo CJ. 2002. *An introduction to gravity waves.* Internat. Geophys. Series. Academic Press: Amsterdam.

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

Obukhov A. 1959. Effect of Archimedean forces on the structure of the temperature field in a turbulent flow. *Dokl. Akad. Nauk SSSR* **125**: 1246.

Parisi G, Frisch U. 1985. A multifractal model of intermittency. pp. 84–88 in *Turbulence and predictability in geophysical fluid dynamics and climate dynamics.* Ghil M, Benzi R, Parisi G. (eds). North Holland: Amsterdam.

Pecknold S, Lovejoy S, Schertzer D, Hooge C, Malouin JF. 1993. The simulation of universal multifractals. pp. 228–267 in *Cellular Automata: Prospects in astronomy and astrophysics.* Perdang JM, Lejeune A. (eds). World Scientific: Singapore.

Radkevich A, Lovejoy S, Strawbridge KB, Schertzer D, Lilley M. 2008. Scaling turbulent atmospheric stratification. III: Space-time stratification of passive scalars from lidar data. *Q. J. R. Meteorol. Soc.* **134**: xxx–xxx.

Richardson LF. 1922. *Weather prediction by numerical process.* Cambridge University Press: republished (1965) Dover: Mineola, New York.

Schertzer D, Lovejoy S. 1983. On the dimension of atmospheric motions. pp. 141–144 in *IUTAM Symposium on turbulence and chaotic phenomena in fluids,* Kyoto, Japan. Tasumi T. (ed).

Schertzer D, Lovejoy S. 1985a. The dimension and intermittency of atmospheric dynamics. pp. 7–33 in *Turbulent Shear Flow 4* Launder B. (ed). Springer-Verlag: Berlin.

Schertzer D, Lovejoy S. 1985b. Generalised scale invariance in turbulent phenomena. *Physicochem. Hydrodyn.* **6**: 623–635.

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

Schertzer D, Lovejoy S. 1997. Universal multifractals do exist! *J. Appl. Meteorol.* **36**: 1296–1303.

Schertzer D, Lovejoy S, Schmitt F. 1995. Structures in turbulence and multifractal universality. pp. 137–144 in *Small-scale structures in 3D and MHD turbulence,* **462.** Meneguzzi M, Pouquet A, Sulem PL. (eds). Springer-Verlag: New York.

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

Schertzer D, Lovejoy S, Schmitt F, Naud C, Marsan D, Chigirinskaya Y, Marguerit C. 1997b. New developments and old questions in multifractal cloud modeling, satellite retrievals and anomalous absorption. pp. 327–335 in Proceedings of 7th Atmospheric Radiation Measurement (ARM) meeting, San Antonio.

Schmitt F, Lovejoy S, Schertzer D, Lavallée D, Hooge C. 1992. First estimates of multifractal indices for velocity and temperature fields. *Comptes Rendus de l'Académie des Sciences de Paris, serie II* **314**: 749–754.

Schmitt F, Schertzer D, Lovejoy S, Brunet G. 1996. Universal multifractal structure of atmospheric temperature and velocity fields. *Europhysics Lett.* **34**: 195–200.

Schmitt F, Schertzer D, Lovejoy S, Brunet Y. 1994. Empirical study of multifractal phase transitions in atmospheric turbulence. *Nonlinear Processes in Geophysics* **1**: 95–104.

Shang XD, Xia KQ. 2001. Scaling of the velocity power spectra in turbulent thermal convection. *Phys. Rev. E* **64**: 065301.

Shur G. 1962. Eksperimental'nyye issledovaniya energeticheskogo spektra atmosfernoy turbulentnosti. *Tsentral'naya Aerologicheskaya Observatoriya. Trudy* **43**: 79.

Tennekes H. 1975. Eulerian and Lagrangian time microscales in isotropic turbulence. *J. Fluid Mech.* **67**: 561–567.

Tessier Y, Lovejoy S, Hubert P, Schertzer D, Pecknold S. 1996. Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. *J. Geophys. Res.* **31D**: 26427–26440.

Van Zandt TE. 1982. A universal spectrum of buoyancy waves in the atmosphere. *Geophys. Res. Lett.* **9**: 575–578.

Weinstock J. 1978. On the theory of turbulence in the buoyancy subrange of stably stratified flows. *J. Atmos Sci.* **35**: 634–649.

Wilson J, Schertzer D, Lovejoy S. 1991. Physically based modelling by multiplicative cascade processes. pp. 185–208 in *Non-linear variability in geophysics: Scaling and Fractals.* Schertzer D, Lovejoy S. (eds). Kluwer.

Wyngaard JC, Coté OR. 1972. Cospectral similarity in the atmospheric surface layer. *Q. J. R. Meteorol. Soc.* **98**: 590–603.

Yaglom AM. 1966. The influence on the fluctuation in energy dissipation on the shape of turbulent characteristics in the inertial interval. *Sov. Phys. Dokl.* **2**: 26–30.