Atmospheric
Chemistry
and Physics

Atmospheric waves as scaling, turbulent phenomena

J. Pinel and S. Lovejoy

Physics, McGill University 3600 University St., Montreal, Que., Canada

Correspondence to: S. Lovejoy (lovejoy@physics.mcgill.ca)

Abstract. It is paradoxical that, while atmospheric dynamics are highly nonlinear and turbulent, atmospheric waves are commonly modelled by linear or weakly nonlinear theories. We postulate that the laws governing atmospheric waves are in fact high-Reynolds-number (Re), emergent laws so that – in common with the emergent high-Re turbulent laws – they are also constrained by scaling symmetries. We propose an effective turbulence–wave propagator which corresponds to a fractional and anisotropic extension of the classical wave equation propagator, with dispersion relations similar to those of inertial gravity waves (and Kelvin waves) yet with an anomalous (fractional) order $H_{\text{wav}}/2$. Using geostationary IR radiances, we estimate the parameters, finding that $H_{\text{wav}} \approx 0.17 \pm 0.04$ (the classical value $= 2$).

1 Introduction

The atmosphere is a highly turbulent system with the ratio of nonlinear to linear terms – the Reynolds number (Re) – typically of the order $\approx 10^{12}$. At the same time, there is no doubt that atmospheric waves exist and play an important role in transferring energy and momentum. These empirical facts only become problematic when we consider the numerous apparently successful studies comparing data with linear (or weakly nonlinear) theory, commonly (for gravity waves) with the Taylor–Goldstein equations or with the linearized shallow-water equations. For example, in the words of Nappo (2002), "almost all of what we know about the nature of gravity waves is derived from the *linear theory*" (emphasis in the original).

Although one may easily get the impression that linear wave theories have been empirically confirmed, a closer look reveals that what has typically been scrutinized is the linear theory dispersion relations. Considering the example of grav-

ity waves, we find that these have mostly been tested in the horizontal (and occasionally in the vertical) directions. Other predictions of the corresponding linear theory – "polarization relations" – are invoked but are only used in a diagnostic mode so that they cannot be considered to have been convincingly validated (see Placke et al., 2013). Recently, linear gravity wave theory has been directly brought into question by data from dropsonde pairs. For example, such pairs have directly shown that certain terms neglected in the Taylor–Goldstein equations are typically much larger than the corresponding retained terms (Lovejoy and Schertzer, 2013). Also the pairs have clearly shown that the vertical structure of the atmosphere is composed of a fractal hierarchy of unstable layers through which linear gravity waves would not able to propagate (Lovejoy et al., 2008a).

The application of linear wave theories are generally justified in cases where the nonlinear terms are weak, such as in theories of linear advection (e.g. Pielke, 2002) or more generally by the existence of large regions of laminar flow, as well as for the study of terrain-induced or heating-induced mesoscale waves (Smith, 1979; Lin, 1987, 2007, and references therein). Empirical and modelling studies have also attempted to relate linear waves to cyclogenesis in the tropics (Frank and Roundy, 2006; Schreck et al., 2012; Shen et al., 2012, 2013, and references therein). However, since the 1980s – and largely thanks to the development of multifractal cascade models – there has been dramatic progress in understanding atmospheric intermittency (Schertzer and Lovejoy, 1987; Frisch, 1995). It is now clear that a prime characteristic of fully developed turbulence is that most of the important fluxes are concentrated in highly sparse (fractal) sets so that much of the flow appears relatively calm. The modern understanding is that, by its very nature, turbulence is highly intermittent so that on any realization of a turbulent process there will be violent regions in proximity to ones of relative calm.

However, examination of the apparently calm regions shows that they also have embedded regions of high activity, and as we zoom into smaller and smaller regions this strong heterogeneity continues in a scaling manner until we reach the dissipation scale (Tuck, 2010). This explains why aircraft measurements of the wind invariably find roughly $k^{-5/3}$ (i.e. turbulent) spectra even in apparently calm regions. Large-scale regions of true laminar flow have yet to be documented by actual measurements. However, the multifractal, multiplicative cascade picture has been well verified even at large scales (e.g. Lovejoy and Schertzer, 2010). Therefore, it would be a mistake to separate these regions of high and low "turbulent intensities" and associate them with different mechanisms or to apply nonturbulent (linear) wave theories to regions of apparent calm.

In the last few years, (nonlinear) scaling theories of waves have become more compelling. This is because empirical evidence and theoretical arguments have amassed to the effect that atmospheric dynamics give rise to emergent high-Reynolds-number scaling laws with different horizontal and vertical exponents. This allows the horizontal scaling to accurately apply over huge ranges in scale (see Lovejoy and Schertzer (2010) and Lovejoy and Schertzer (2013) for recent reviews). Based on the classical laws of turbulence, they involve extensions to account for (multifractal) intermittency and anisotropy. Their success underlines the fundamental role of scale symmetries in constraining the high-Re dynamics. All this motivates the following question: are atmospheric waves also scaling turbulent phenomena? If this is the case, we may logically expect anomalous wave propagators that could readily have dispersion relations identical to or nearly indistinguishable from their classical counterparts, while simultaneously having nontrivial consequences for the dynamics and for our understanding – for example, we find that energy transport will be modified (Appendix B).

If dispersion relations from linear theory and those from strongly nonlinear theory can be very similar to each other, then how might one empirically distinguish them? The obvious way is to note that linear theory also predicts the entire space–time propagators relating the wave forcings and responses. A key characteristic of linear theories is that they involve integer powers of the (space and time) differential operators, and this strongly constrains the form of the propagators; below, we show how this allows us to test the theory by seeking possible anomalous propagator exponents. We investigate this using geostationary satellite infrared radiances.

This paper attempts to show how scaling propagators with both turbulent and wave-like characteristics could arise while being consistent with both (anisotropic) turbulence theory and observations. However, let the reader be warned that – while the turbulent part of the propagator, which was derived and empirically tested elsewhere (it is summarized here in a Appendix A), is reasonably well grounded – the wave-like part, i.e. the subject of this paper, is in contrast fairly speculative; it is perhaps little more than a proof of concept. On the theoretical side, the reason is that with only scaling symmetries to guide us the possibilities are very broad, while on the empirical side over the scaling range accessible here (60–5000 km in space and 3–100 h in time) the turbulent part of the spectrum is by far the most dominant one, accounting for an empirical range of spectral densities of a factor $\approx 10^5$, leaving the residual wave-like part to account for the remaining factor of 0.9 ± 0.5 in the dynamical spectral scaling range.

2　Fractional propagators and turbulence

In order to motivate our model, consider the classical wave equation for the wave I with forcing f:

$$\left(\nabla^2 - \frac{1}{V^2}\frac{\partial^2}{\partial t^2}\right) I(\boldsymbol{r}, t) = f(\boldsymbol{r}, t). \tag{1}$$

V is the wave velocity, $\boldsymbol{r}$ is the position vector and t the time variable.

As usual, we can solve Eq. (1) by taking Fourier transforms (denoted by tildes):

$$\tilde{I}(\boldsymbol{k}, \omega) = \tilde{g}(\boldsymbol{k}, \omega)\tilde{f}(\boldsymbol{k}, \omega); \tilde{g}(\boldsymbol{k}, \omega) = \left(\omega^2/V^2 - |\boldsymbol{k}|^2\right)^{-1}, \tag{2}$$

where $\boldsymbol{k}$ is the wave vector, ω the frequency and $\tilde{g}(\boldsymbol{k}, \omega)$ is the propagator. This propagator is symmetric with respect to an isotropic space–time scale transformation by factor λ^{-1}:

$$\tilde{g}\left(\lambda^{-1}(\boldsymbol{k}, \omega)\right) = \lambda^H \tilde{g}((\boldsymbol{k}, \omega)); \quad H = 2. \tag{3}$$

However we anticipate that at high Re "effective propagators" may emerge constrained by the same overall scaling symmetry but with some other "anomalous" exponent $H \neq 2$. In this case we obtain fractional propagators corresponding to fractional generalizations of the wave equation

$$\tilde{g}(\boldsymbol{k}, \omega) = \left(\omega^2/V^2 - |\boldsymbol{k}|^2\right)^{-H/2}; \tag{4}$$

$$\left(\nabla^2 - \frac{1}{V^2}\frac{\partial^2}{\partial t^2}\right)^{H/2} I(\boldsymbol{r}, t) = f(\boldsymbol{r}, t).$$

Although we will only require fractional propagators, if needed, we could define the fractional differential operator in Eq. (4) by the inverse Fourier transform of $\tilde{g}(\boldsymbol{k}, \omega)^{-1}$ (or see e.g. Miller and Ross (1993) for fractional differential equations). If we seek the real space solution of Eq. (1) or (4), we can use the fact that Fourier space products (Eq. 2) correspond to real space convolutions ("$*$"); hence the solutions to Eqs. (1) and (4) are $I(\boldsymbol{r}, t) = g(\boldsymbol{r}, t) * f(\boldsymbol{r}, t)$ so that the propagator links the forcing f to the response I.

In order to estimate $g(\boldsymbol{r}, t)$ we can appeal to the method of stationary phase (e.g. Bleistein and Handelsman (1986)) which ensures us that the dominant contribution to $g(\boldsymbol{r}, t)$ is due to the wave-number–frequency region over which

$\tilde{g}(\boldsymbol{k}, \omega)$ is singular; this singularity defines the dispersion relation and accounts for its origin and significance (see Appendix B). For both the classical Eq. (1) and the nonclassical Eq. (4), we find the dispersion relation

$$\omega = \pm V |\boldsymbol{k}|, \tag{5}$$

which is therefore of fundamental importance, a fact which is true for any $H > 0$, not only for positive but even integer values of H.

Before attempting to estimate propagators of real data, we must take into account the fact that atmospheric waves occur in the presence of turbulence. Indeed the spectrum is so strongly dominated by a "turbulent background" that it must first be removed before evidence of any wave-like propagator can be observed. This is paradoxical since the wave-like part implies the existence of a set of points (k_x, k_y, ω) such that $\tilde{g}(k_x, k_y, \omega) \rightarrow \infty$; in the simplest case, such a "singular set" is a surface in (k_x, k_y, ω) space and should be easy to detect, although the topology need not be so simple (see Sect. 5). However, the singularities are apparently of sufficiently low order and the spectral estimates are sufficiently noisy that in practice the singular set is hard to observe. Indeed it is much easier to study 2-D subspaces obtained by integrating out one of the spectral coordinates (which also reduces the "noise"), although if H is small enough (and this is indeed the case here; see below) this can integrate out the singularities. Indeed, one of the main techniques for empirically investigating atmospheric waves (Wheeler and Kiladis, 1999; Hendon and Wheeler, 2008; Dias et al., 2012) integrates over k_y space to yield a (k_x, ω) 2-D spectrum while also using an ad hoc averaging technique for removing the turbulent contribution.

Following Wheeler and Kiladis (1999) (and see Kiladis et al. (2008) for a review), we also use infrared data although at hourly not daily resolution: we use instead a theoretically motivated turbulent spectrum to search for evidence of anomalous wave propagators. To understand this, recall the classical Kolmogorov law of three dimensional isotropic turbulence:

$$\Delta I (\Delta \boldsymbol{r}) = \phi |\Delta \boldsymbol{r}|^H; \; \phi = \varepsilon^{1/3}; \; H = 1/3, \tag{6}$$

where I is a component of the wind, ΔI is a fluctuation, $\Delta \boldsymbol{r}$ is a vector displacement, ε is the turbulent energy flux and the equality is understood in a statistical sense. In Fourier space this becomes

$$\tilde{I}(\boldsymbol{k}) = \tilde{g}_{\text{tur}}(\boldsymbol{k})\tilde{\phi}(\boldsymbol{k}) \; ; \; \tilde{g}_{\text{tur}}(\boldsymbol{k}) = |\boldsymbol{k}|^{-H}. \tag{7}$$

Comparing this with Eq. (2), we see that $\tilde{\phi}(\boldsymbol{k})$ is the forcing and $\tilde{g}_{\text{tur}}(\boldsymbol{k})$ is the spatial part of a (fractional order) propagator (a Green's function). Now recall that, for real I, $\tilde{I}(\boldsymbol{k}) = \tilde{I}^*(-\boldsymbol{k})$, if in addition we assume statistical translational invariance ("statistical homogeneity"), then we may define the spectral densities P_I, P_ϕ, by

$$\left\langle \tilde{I}(\boldsymbol{k}) \, \tilde{I}(\boldsymbol{k}') \right\rangle = \delta(\boldsymbol{k} + \boldsymbol{k}') P_I(\boldsymbol{k}), \tag{8}$$

so that $P_I(\boldsymbol{k}) \propto \left\langle \left| \tilde{I} \right|^2 \right\rangle$, $P_\phi(\boldsymbol{k}) \propto \left\langle \left| \tilde{\phi} \right|^2 \right\rangle$, where "$<\cdot>$" denotes ensemble averaging.

To obtain the classical Kolmogorov–Obhukhov $k^{-5/3}$ law we use

$$P_\phi(\boldsymbol{k}) = P_0 |\boldsymbol{k}|^{-s_\phi}; s_\phi = d - K(2), \tag{9}$$

where P_0 is a dimensional constant, d is the dimension of space and $K(2)$ is the second-order intermittency correction. This yields

$$P_I(\boldsymbol{k}) = P_\phi(\boldsymbol{k}) |\tilde{g}_{\text{tur}}|^2 = P_0 |\boldsymbol{k}|^{-s_\phi - 2H} = P_0 k^{-s_\phi - 2H}; \tag{10}$$
$$k = |\boldsymbol{k}|.$$

The angle-integrated ("isotropic") spectral density $E(k)$ is then given by integrating P over shells in Fourier space. Ignoring constant factors (2π in $d = 2$, 4π in $d = 3$), we obtain the (intermittency corrected) isotropic Kolmogorov law

$$E(k) \approx P_I(k)k^{d-1} = k^{-\beta} \; ; \; \beta = 1 + 2H - K(2) \tag{11}$$

(since $H = 1/3$, we see that the nonintermittent $K(2) = 0$ case does indeed have exponent $\beta = 5/3$).

A basic consequence of wide-range spatial scaling of atmospheric fields (in particular the wind) is that the spectrum and spectral density of the turbulent fluctuations in horizontal wave-number–frequency (k_x, k_y, ω) space follow the straightforward space–time extension of Eq. (6):

$$\tilde{I}(\boldsymbol{k}, \omega) = \tilde{g}_{\text{tur}}(\boldsymbol{k}, \omega)\tilde{\phi}(\boldsymbol{k}, \omega); \tag{12}$$
$$P_I(\boldsymbol{k}, \omega) = |\tilde{g}_{\text{tur}}(\boldsymbol{k}, \omega)|^2 P_\phi(\boldsymbol{k}, \omega),$$

where $P_I(\boldsymbol{k}, \omega)$ and $P_\phi(\boldsymbol{k}, \omega)$ are space–time spectral densities, and $\tilde{g}_{\text{tur}}(\boldsymbol{k}, \omega)$ is the turbulent propagator. To obtain the form of $\tilde{g}_{\text{tur}}(\boldsymbol{k}, \omega)$, we follow Lovejoy and Schertzer (2010) and Pinel et al. (2014) outlined in Appendix A (see Eq. (A13)) to obtain the dimensionless propagator

$$\tilde{g}_{\text{tur}}(\boldsymbol{k}, \omega) = \left(-i\omega' + \|\boldsymbol{k}\|\right)^{-H_{\text{tur}}}, \tag{13}$$

where

$$\omega' = (\omega + \boldsymbol{k} \cdot \boldsymbol{\mu})\sigma^{-1}; \quad \sigma = \left(1 - \left(\mu_x^2 + a^2\mu_y^2\right)\right)^{1/2}; \tag{14}$$
$$\|\boldsymbol{k}\| = \left(k_x^2 + a^2 k_y^2\right)^{1/2},$$

where k_x, k_y and ω have been nondimensionalized as discussed in Appendix A, using the size of the Earth and the turbulent velocity V_w, and $\boldsymbol{\mu} = (\mu_x, \mu_y)$ is the mean dimensionless horizontal advection vector. The (horizontal) spatial (Fourier) scale function is $\|\boldsymbol{k}\|$ and a is the north–south/east–west aspect ratio; when $a = 1$, we obtain $\|\boldsymbol{k}\| = |\boldsymbol{k}|$. The transformation $\omega \rightarrow \omega'$ combines the effects of a mean advection by velocity $\boldsymbol{\mu}$ and the statistical variability of the advection wind about its mean (via σ). Note that (a) the factor i in

Eq. (13) is necessary so that the propagator respects causality, and (b) overall $\tilde{g}_{\text{tur}}$ respects the same isotropic scaling symmetry as the wave propagator (Eq. 3) but with exponent H_{tur}.

Note that, if needed, more complex spatial scale functions may be used; they they are only weakly constrained to be of the general form $\|\boldsymbol{k}\| = |\boldsymbol{k}| \, \Phi\left(\widehat{\boldsymbol{k}}\right)$, where $\widehat{\boldsymbol{k}}$ is a unit vector and Φ a fairly general function (see Appendix B for more details). $\|\boldsymbol{k}\|$ replaces the vector norms in the isotropic theories. For another example, below we consider the specific case of Kelvin waves with the help of another wave scale function – $\|\boldsymbol{k}\| = \left(k_{\text{x}}^2 - a^2 k_{\text{y}}^2\right)^{1/2}$ – which has the effect of "channelling" the waves in the zonal direction.

3　Fractional propagators and waves

With the exception of the weak singularities associated with waves, the turbulence dominates the spectral density; the $P_I(\boldsymbol{k}, \omega)$ given in Eq. (12) with the propagator Eq. (13) already gives a good approximation to the empirical spectral density. This may be seen in Fig. 1 using Multifunctional Transport Satellite (MTSAT) data (described below) which shows the 1-D spectral densities $E(k_{\text{x}})$, $E(k_{\text{y}})$, $E(\omega)$ obtained by successively integrating out various pairs of variables from $P_I(k_{\text{x}}, k_{\text{y}}, $ and $\omega)$ (see Pinel et al., 2014). The log–log linearity on this figure directly shows that the spectra are scaling and the near-perfect superposition of the 1-D spectra demonstrates that the scaling exponents are essentially identical so that (in conformity with Eq. (3) the conclusions of Appendix A and the form Eq. (13)) the radiance field structure functions are symmetric with respect to isotropic scale changes $(\Delta x, \Delta y, \Delta t) \rightarrow \lambda^{-1}(\Delta x, \Delta y, \Delta t)$ or, equivalently, $(k_{\text{x}}, k_{\text{y}}, \omega) \rightarrow \lambda(k_{\text{x}}, k_{\text{y}}, \omega)$. This turbulence part corresponds to the "background" spectrum obtained by Wheeler and Kiladis (1999); any wave behaviour is to be found in deviations from this.

A simple model that takes into account waves while respecting both the space–time scaling and the turbulent forcing and background is obtained by including a factor $\tilde{g}_{\text{wav}}$ in the overall propagator. To be "wave-like", $\tilde{g}_{\text{wav}}$ must be causal, unlocalized in space–time and must also be chosen so that the overall scaling symmetry of the system (Eq. 3) is respected by the overall propagator $\tilde{g}_I(\boldsymbol{k}, \omega)$. Following Wheeler and Kiladis (1999), who factored the spectral density into a "red noise" turbulent background and a wave part, and inspired by Eqs. (2) and (13), we can use the form

$$\tilde{I}(\boldsymbol{k}, \omega) = \tilde{g}_I(\boldsymbol{k}, \omega)\tilde{\phi}(\boldsymbol{k}, \omega); \quad \tilde{g}_I(\boldsymbol{k}, \omega) = \tilde{g}_{\text{tur}}(\boldsymbol{k}, \omega)\tilde{g}_{\text{wav}}(\boldsymbol{k}, \omega), \tag{15}$$

with $\tilde{g}_{\text{tur}}$ given by Eq. (13) and $\tilde{g}_{\text{wav}}$ given by

$$\tilde{g}_{\text{wav}}(\boldsymbol{k}, \omega) = \left(\omega'^2 \big/ v_{\text{wav}}^2 - \|\boldsymbol{k}\|^2\right)^{-H_{\text{wav}}/2}. \tag{16}$$

This is a generalization of Eq. (2) to account for spatial anisotropy (with $|\boldsymbol{k}| \rightarrow \|\boldsymbol{k}\|$). The replacement $\omega \rightarrow$

$(\omega + \boldsymbol{k} \cdot \boldsymbol{\mu})\sigma^{-1}$ is the classical advection transformation (see e.g. Nappo, 2002); although just as in the turbulent propagator (where it is more fully justified), we have included the extra factor σ to take into account the statistical variation of the advection velocity (see Eq. 14). Finally, the parameter v_{wav} is the phase speed nondimensionalized by the turbulent velocity V_w (Eq. 13). Note that the overall propagator $\tilde{g}_I$ satisfies the scaling symmetry Eq. (3) with $H = H_{\text{tur}} + H_{\text{wav}}$. Due to $\tilde{g}_{\text{wav}}$, the overall propagator $\tilde{g}_I$ yields the dispersion relation

$$\omega = -\boldsymbol{k} \cdot \boldsymbol{\mu} \pm \sigma v_{\text{wav}} \|\boldsymbol{k}\|. \tag{17}$$

With respect to the background advection ($\boldsymbol{\mu}$), σv_{wav} is the effective wave speed which takes into account the mean wave speed (v_{max}) and the statistical variability via σ. By taking appropriate scale functions $\|\boldsymbol{k}\|$ one can obtain dispersion relations close to gravity and other waves (see Lovejoy et al., 2008b). The factorization of the propagator into a turbulent and wave-like part (Eq. 15) is quite natural since it can be re-written $\tilde{g}_I = \tilde{g}_{\text{wav}}\tilde{\phi}'$, where $\phi' = \tilde{g}_{\text{tur}}\tilde{\phi}$ corresponds to the overall space–time localized (turbulent-like) response to the forcing turbulent flux; the overall result is a turbulent-like forcing $\tilde{\phi}'$ for the waves; in real space, $g_I = g_{\text{wav}} * \phi'$.

To interpret the propagators and dispersion relations in terms of travelling waves, note that in the Fourier expansion of $g(\boldsymbol{r}, t)$, the propagator $\tilde{g}_I(\boldsymbol{k}, \omega)$ is the amplitude of the Fourier coefficient of $\exp(-i(\boldsymbol{k} \cdot \boldsymbol{r} + \omega t))$; i.e. it corresponds to a wave travelling in the direction $-\boldsymbol{k}$. Evanescent waves occur when, for a given real frequency ω, the dispersion relation implies that in some directions $\boldsymbol{k}$ is no longer real but complex with the imaginary part corresponding to exponential damping. In the propagator framework, the equivalent phenomenon is obtained by considering, for a given wave vector $\boldsymbol{k}$, the corresponding frequency (e.g. Eq. 17); when the latter is complex, then waves in the direction $-\boldsymbol{k}$ will decay rapidly in time and will thus not propagate.

For the most general propagator respecting the constraints of scaling, causality and reality, see Appendix B. Although the dispersion relation is independent of the propagator exponent H_{wav}, the exponent does determine the (power law) rate of decay of the forcing so that the value of H_{wav} will affect the transport of momentum and energy (for the general result, see Bleistein and Handelsman (1986); for an application to gravity waves, see Lovejoy et al (2008)).

Of more relevance here are Kelvin waves which are the low Coriolis parameter/high "effective thickness" limit of the inertial gravity (Poincaré) wave dispersion relations often invoked at these space–time scales. First, for only one spatial (zonal) dimension, we may note that Kelvin waves are a special case of Eq. (17) with $\|\boldsymbol{k}\| = k_{\text{x}}$. Considering the full horizontal plane, Kelvin waves are channelled; they only propagate in the zonal direction. To obtain some channelling while maintaining the same overall scaling symmetry, in $\tilde{g}_{\text{wav}}$ we could replace the spatial (Fourier) scale

function $\|\boldsymbol{k}\| = \left(k_{\mathrm{x}}^2 + a^2 k_{\mathrm{y}}^2\right)^{1/2}$ by $\|\boldsymbol{k}\| = \left(k_{\mathrm{x}}^2 - a^2 k_{\mathrm{y}}^2\right)^{1/2}$, which only allows meridional propagation for small-scale (high-wave-number) structures. For example, when $\boldsymbol{\mu} = 0$, large structures with $k_{\mathrm{x}} < \omega/(\sigma v_{\mathrm{wav}})$ cannot propagate in the meridional direction, they are channelled.

Finally, combining Eq. (13), (15) and (16), we obtain the turbulent-wave spectral density:

$$P_I(\boldsymbol{k}, \omega) = P_\phi(\boldsymbol{k}, \omega) \, |\tilde{g}_{\mathrm{tur}}|^2 \, |\tilde{g}_{\mathrm{wav}}|^2 ; \qquad (18)$$

$$P_I(\boldsymbol{k}, \omega) = P_\phi(\boldsymbol{k}, \omega) \left(\omega'^2 + \|\boldsymbol{k}\|^2\right)^{-H_{\mathrm{tur}}} \left(\omega'^2/v_{\mathrm{wav}}^2 - \|\boldsymbol{k}\|^2\right)^{-H_{\mathrm{wav}}} ;$$

$$P_\phi(\boldsymbol{k}, \omega) = P_0 \left(\omega'^2 + \|\boldsymbol{k}\|^2\right)^{-s_\phi/2} .$$

In Eq. (18), we have followed the assumption in the isotropic case (Eq. 10) that the forcing of the flux has the same scale symmetries as $|\tilde{g}_{\mathrm{tur}}|^2$; from Eq. (A13) we see that $s_\phi = d - K(2)$ is the spectral exponent of the flux and P_0 a dimensional constant determined by the climatological (low-frequency) average forcing.

4 Data analysis

We follow Wheeler and Kiladis (1999) and Hendon and Wheeler (2008) but estimate the turbulent background using regressions to estimate the parameters of $\tilde{g}_I$ (i.e. of $\tilde{g}_{\mathrm{tur}}$ and $\tilde{g}_{\mathrm{wav}}$). The data set was comprised of 1386 images ($\sim$two months of data, September and October 2007) of radiances measured by the first "thermal" infrared channel (10.3–11.3 µm, particularly sensitive to temperature near the top of clouds) of the geostationary satellite MTSAT over the western South Pacific at resolutions 30 km and 1 hr over latitudes 40° S–30° N and longitudes 80° E–200° E. We separated the sample into five 277 h ($\sim$12-day) blocks, calculating for each block the spectral density of fluctuations of the field with respect to the mean image (we used a standard Hann window to reduce spectral leakage). Note that as opposed to Wheeler and Kiladis (1999), who averaged their data in order to estimate the turbulent contribution to the signal, we rather averaged our data to obtain a better statistical estimate of the ensemble spectrum; the regression to the theoretical form provides the smooth background. In contrast, Wheeler and Kiladis (1999) performed an additional ad hoc smoothing to better bring out wave-like signals. Our choice of 12-day blocks was made since the temporal scaling has a break at about 5–10 days and we were only interested in analysing the high-frequency "weather" regime. Choosing a longer block period would allow us to examine lower frequencies, but would take us outside the unique (weather) scaling regime considered in this paper and would decrease the number of blocks and the corresponding amount of averaging.

To see how a purely turbulent spectrum already provides a good approximation, we performed a multivariate regression on the empirical MTSAT spectral density and theoretical
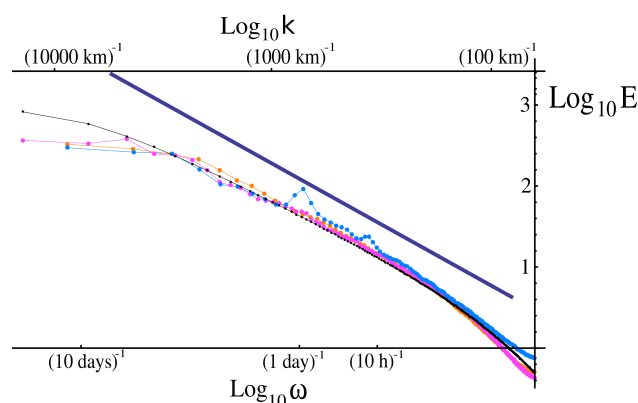
Fig. 1. 1-D spectra from MTSAT data; blue: temporal; orange: meridional; purple: zonal and a multivariate regression curved due to the finite empirical domain; black, using $H_{\mathrm{wav}} = 0$, $V_w = 41 \pm 3\,\mathrm{km\,h^{-1}}$; $\tau_w = L_e/V_w \approx 20 \pm 1$ days; $a \approx 1.2 \pm 0.1$; $s_I \approx 3.4 \pm 0.1$; $P_0 = 2.8 \pm 0.2°\,\mathrm{C^2\,km^2 h}$; $\mu_{\mathrm{x}} \approx -0.3 \pm 0.1$; ($\overline{v_{\mathrm{x}}} \approx -12 \pm 4\,\mathrm{km\,h^{-1}}$); $\mu_{\mathrm{y}} \approx 0.10 \pm 0.08$; ($\overline{v_{\mathrm{y}}} \approx 4 \pm 3\,\mathrm{km\,h^{-1}}$); $\sigma = 0.95 \pm 0.03$.

form (Eq. 18) with $H_{\mathrm{wav}} = 0$. Figure 1 shows the 1D spectra obtained by integrating the 3-D density over the complementary coordinates using $s_\phi = 2.88 \pm 0.01$ and $H_{\mathrm{tur}} = H = 0.26 \pm 0.05$. The log-log linearfit is good over the range of scales 120–5000 km in space and 3–100 h in time (except for small diurnal contributions at 12 and 24 h); it is especially good if we numerically take into account finite sample size effects at the large and small scales (the curvature in the black line in Fig. 1 i.e. good agreement to the highest wavenumber at 60 km scales). The excellent superposition confirms the scale symmetry of the type Eq. (3): $P_I\left(\lambda^{-1}(\boldsymbol{k}, \omega)\right) = \lambda^{s_I} P_I(\boldsymbol{k}, \omega)$, with $s_I = s_\phi + 2H = 3.4 \pm 0.1$ (Eq. 18).

We now consider the three 2-D spectra, obtained by successively integrating the 3-D spectral density over k_{x}, k_{y} and ω. The fit is sufficiently good that we can use the above regression with $H_{\mathrm{wav}} = 0$ to estimate all the turbulent parameters. However for the 1-D spectra to have fixed exponents, when fitting the wave part we must use the constraint $H = H_{\mathrm{tur}} + H_{\mathrm{wav}}$ so that the 1-D spectral slopes are not affected. In this way we find an optimum relative weighting for the turbulence and wave contributions. Figure 2a–d show these for three different values of H_{wav}. As before, the purely turbulent ($H_{\mathrm{wav}} = 0$) case gives a good fit with mean deviations $\pm 11\%$ in $\mathrm{Log}_{10} P(k, \omega)$ in the three 2-D spaces (excluding the diurnal spikes and the origin), which is small considering that the 2-D space signal $P(k, \omega)$ varies over about four orders of magnitude. The orientations of the contours of $P(k_{\mathrm{x}}, \omega)$ and $P(k_{\mathrm{y}}, \omega)$ is a consequence of the nonzero mean zonal velocity $\overline{v_{\mathrm{x}}} \sim -12 \pm 4\,\mathrm{km\,h^{-1}}$ and smaller mean meridional velocity $\overline{v_{\mathrm{y}}} \sim 4 \pm 3\,\mathrm{km\,h^{-1}}$, the wave part is the residual: $|\tilde{g}_{\mathrm{wav}}|^2 \propto P_I \left(\omega'^2 + \|\boldsymbol{k}\|^2\right)^{s_\phi/2 + H_{\mathrm{tur}}}$; see Eq. (18)
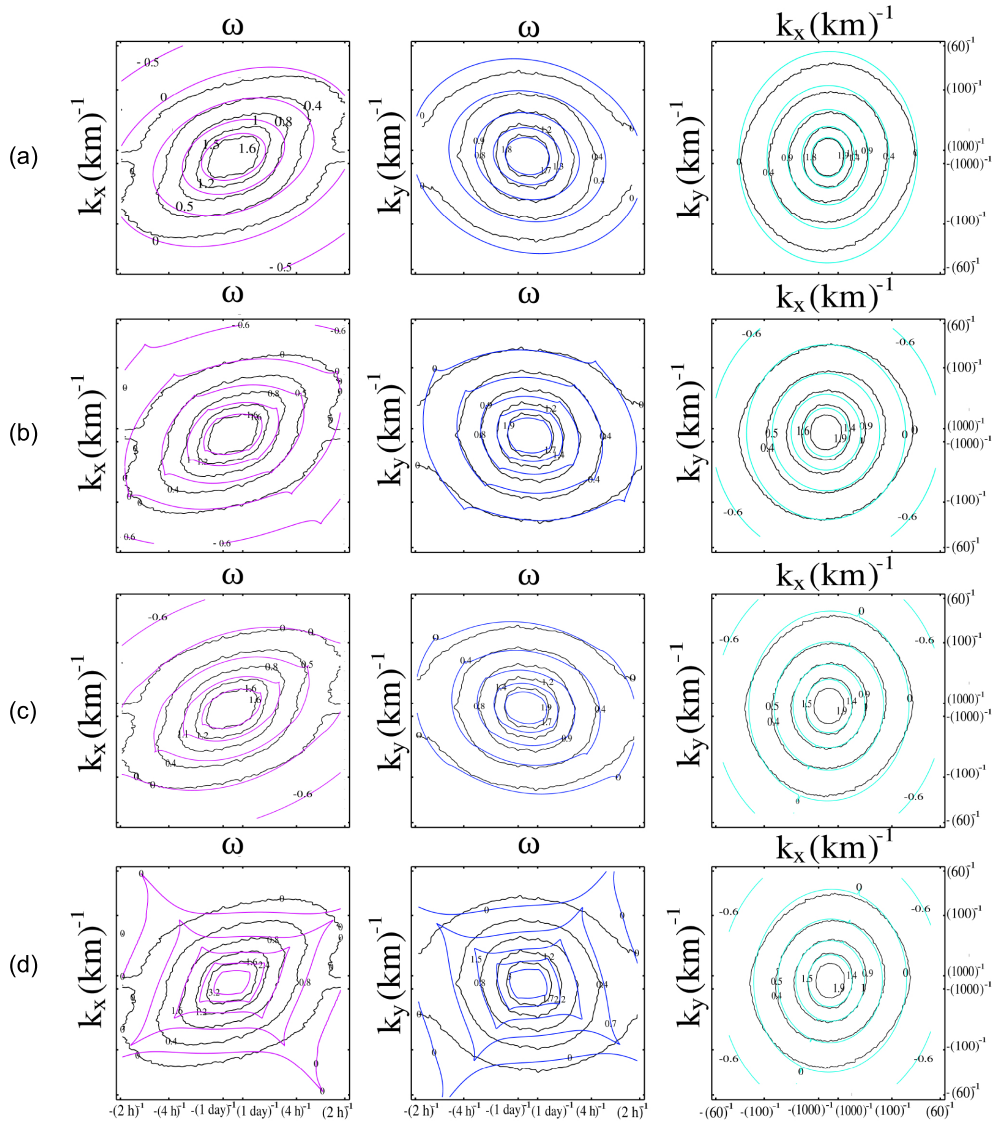
Fig. 2. Comparison of 2-D spectral densities from MTSAT data (in black) and a multivariate regression from theoretical $P_I(\boldsymbol{k}, \omega)$ given by Eq. (18) (in colour). The 2-D subspaces (left to right) are (ω, k_x), (ω, k_y) and (k_x, k_y). The ranges are ω from $(2\,\text{h})^{-1}$ to $(277\,\text{h})^{-1}$; k_x from $(60\,\text{km})^{-1}$ to $(\approx 13\,000\,\text{km})^{-1}$ and k_y from $(60\,\text{km})^{-1}$ to $(\approx 8000\,\text{km})^{-1}$. **(a)** With imposed $H_{\text{wav}} = 0$, which corresponds to the purely turbulent case (with $H_{\text{tur}} = H - H_{\text{wav}}$, $H = 0.26 \pm 0.05$, $s_\phi = 2.88 \pm 0.01$). The other parameters are $V_w = 41 \pm 3\,\text{km h}^{-1}$; $\tau_w = L_e / V_w \approx 20 \pm 1$ days; $a \approx 1.2 \pm 0.1$; $s_I \approx 3.4 \pm 0.1$; $P_0 = 2.8 \pm 0.2°$, $\text{C}^2\,\text{km}^2\text{h}$; $\mu_x \approx -0.3 \pm 0.1$; $(\overline{v_x} \approx -12 \pm 4\,\text{km h}^{-1})$; $\mu_y \approx 0.10 \pm 0.08$; and $(\overline{v_y} \approx 4 \pm 3\,\text{km h}^{-1})$. Hence, $\sigma = 0.95 \pm 0.03$. **(b)** $H_{\text{wav}} = 0.17 \pm 0.04$ (best-fit value) and nondimensional wave speed $v_{\text{wav}} = 1.0 \pm 0.8$. **(c)** Fit from Eq. (18) with $\tilde{g}_{\text{wav}}$ from Eq. (19). $H_{\text{wav}} = 0.08 \pm 0.04$ and $v_{\text{wav}} = 1.4 \pm 0.8$. **(d)** Same parameters as **(a)**, but with imposed $H_{\text{wav}} = 1$.

with $\|\boldsymbol{k}\| = \left(k_x^2 + a^2 k_y^2\right)^{1/2}$. Although this is noisy, the value $H_{\text{wav}} \approx 0.17 \pm 0.04$ (so that $H_{\text{tur}} = H - H_{\text{wav}} = 0.09 \pm 0.06$; $H = 0.26 \pm 0.05$ is fixed) gives the best overall fit and nondimensional wave speed $v_{\text{wav}} = 1.0 \pm 0.8$. Recall that the case $v_{\text{wav}} = 1$ means the wave speed is equal to that of the turbulent wind. Note that, even though $H_{\text{wav}} > H_{\text{tur}}$, the turbulence still dominates the overall spectrum: due to the factor P_ϕ (Eq. (18)) one should compare $s_\phi + 2H_{\text{tur}} \sim 3.06$ with $2H_{\text{wav}} \sim 0.34$.

In order to isolate the wave contribution to the spectrum, Wheeler and Kiladis (1999) removed a turbulent background (estimated with an ad hoc averaging technique) from their (k_x, ω) 2-D spectrum and tried to identify maxima in the residual with linear theory dispersion relations. Following them, we removed (by dividing) from the empirical 3-D spectral density the turbulent background estimated from the fit of Eq. (18) with $H_{\text{wav}} = 0$ (i.e. the purely turbulent spectral density from which we obtained Fig. 2a). The residual from
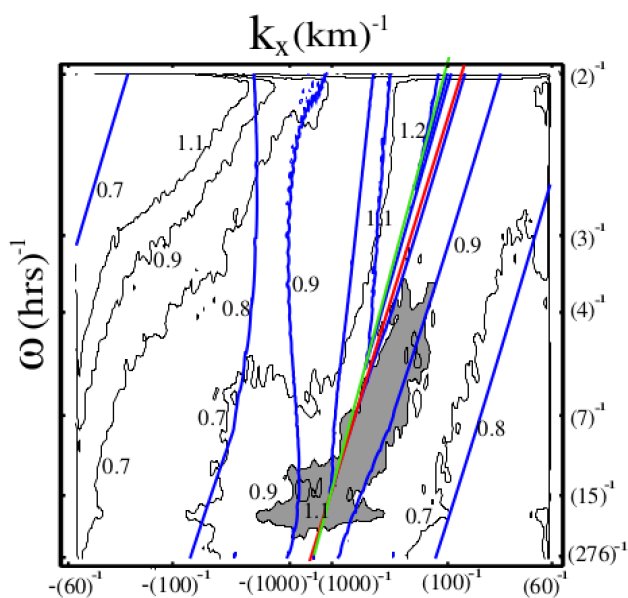
Fig. 3. Contour plot in the 2-D (k_x, ω) space (i.e. integrated over k_y). Black lines: the empirical spectral density divided by the "turbulent background" (i.e. the fit of Eq. (18) with imposed $H_{wav} = 0$; region of maxima relevant to Kelvin waves indicated in grey). Blue lines: $|\tilde{g}_{wav}|^2 = |\omega'/v_{wav} + \| \boldsymbol{k} \| \operatorname{sign}(\boldsymbol{k} \cdot \boldsymbol{\mu})|^{-2H_{wav}}$ with parameters $H_{wav} = 0.08 \pm 0.04$ and $v_{wav} = 1.4 \pm 0.8$ integrated over k_y (the maximal line is indicated in green). In red: the dispersion relation for Kelvin waves (corresponding to $h = 12$ m in Fig. 3 of Wheeler and Kiladis, 1999).

which wave behaviour is to be identified (and which is to be described by the wave part of Eq. 18) is presented in Fig. 3 for the (k_x, ω) 2-D space with $\omega > 0$ (i.e. after integrating over k_y). We observe a region of maxima (Fig. 3 in grey) for $k_x > 0$ which is similar to the residual obtained by Wheeler and Kiladis (1999) although for larger wave numbers and frequencies. Also shown in Fig. 3 is the theoretical dispersion relation for Kelvin waves which was obtained by Wheeler and Kiladis (1999) (compare with their Fig. 3 for the equivalent depth $h = 12$ m).

A key point about Kelvin waves and Fig. 3 is that they are asymmetrical in the zonal direction (k_x). In contrast, the simple form of $\tilde{g}_{wav}$ used in Eq. (18) involves a Fourier space scale function $\| \boldsymbol{k} \| = \left(k_x^2 + a^2 k_y^2 \right)^{1/2}$ which is symmetrical in $\boldsymbol{k}$ and which involves, in the (k_x, ω) space, maxima lines (coming from the singularities) for $k_x > 0$ as well as for $k_x < 0$, which is incompatible with Fig. 3 (grey region). However, the only constraints on the form of $\tilde{g}_{wav}$ are that it must respect causality, that g is real (hence $\tilde{g}(\boldsymbol{k}, \omega) = \tilde{g}^*(-\boldsymbol{k}, -\omega)$), and that the overall scaling symmetry (Eq. 3) is respected (see Appendix B for the general form). We can therefore modify the form of $\tilde{g}_{wav}$ so that it is no longer invariant under $k_x \rightarrow$

$-k_x$. For example, the following form is adequate:

$$\tilde{g}_{wav}(\boldsymbol{k}, \omega) = \left\{ i \left(\omega'/v_{wav} + \| \boldsymbol{k} \| \operatorname{sign}(\boldsymbol{k} \cdot \boldsymbol{\mu}) \right) \right\}^{-H_{wav}}; \quad (19)$$

$$\| \boldsymbol{k} \|_{wav} = \left(k_x^2 + a^2 k_y^2 \right)^{1/2}.$$

Replacing $\tilde{g}_{wav}$ in Eq. (18) with Eq. (19) preserves the quality of the fit of the total (turbulent-wave) spectral density (see the 2-D subspaces in Fig. 2c) and gives a spectra $P_{wave}(k_x, \omega)$ close to the data, including a maxima line which is close to the maxima in the residual presented in Fig. 3. With this asymmetrical propagator, we find that the only parameters that change significantly are $v_{wav} = 1.4 \pm 0.8$, $H_{wav} = 0.08 \pm 0.04$ (so that $H_{tur} = H - H_{wav} = 0.18 \pm 0.06$).

5 Refined singularity analysis

The above analysis is paradoxical since our hypothesis is that there is a singular surface in (k_x, k_y, ω) space yet analysis of the 1-D and 2-D sections showed no direct evidence of singular behaviour. This is consistent with the finding that $0 < H_{wav} < 1$, implying that the singularities are integrated out in the lower-dimensional sections. In order to display potential singularities, we are therefore forced to study the full 3-D density $P(k_x, k_y, \omega)$ recognizing that most of the variation is due to the turbulent part and that the wave part – being only weakly singular – is expected to manifest itself in maxima, perhaps with surface-like topology (line-like in 2-D sections). If one considers 2-D sections of $P(k_x, k_y, \omega)$ – for instance for ω fixed – and using a landscape analogy, these maxima would be either isolated peaks or crests of mountain ranges (including saddle points in such crests). To detect these peaks or crests, we implemented an ad hoc singularity detection algorithm that "scans" parallel to the axes to estimate maxima successively in the k_x, and k_y directions (for ω fixed). In principle, considering the maxima in a single direction is adequate, but in practice the singular surface has parts that are roughly parallel to a given axis; the resulting ambiguity can be resolved by determining maxima in two orthogonal directions.

The results are shown in Fig. 4, where we compare such an analysis with the theoretical behaviour (with $\tilde{g}_{wav}$ from Eq. (19)) for constant ω sections. A drawback of the method is that it does not distinguish maxima due to the turbulent contribution and from the (presumed) wave contribution, and in the empirical case the separation is not always evident. In the figure, the two have been distinguished by the colour of the lines. We see that, although far from perfect, the semi-ellipses indicating the theoretical singularity (dispersion relation) are close to the empirical ones. Given that we used a straightforward generalization of the classical wave equation with only one new parameter, v_{wav} (two if we include $H_{wav} = 0.08$, but this does not affect the singular surface), and given that the wave part of the overall propagator Eq. (16) as the postulated multiplicative decomposition (Eq. 15) are not
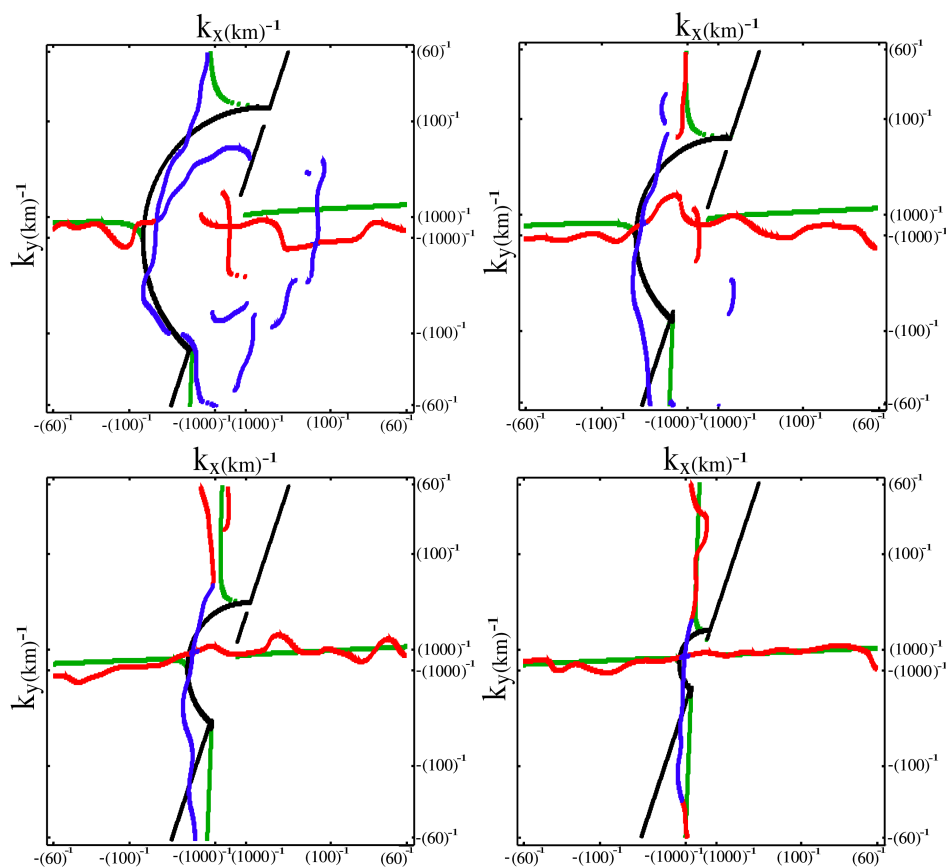
Fig. 4. Left to right, top to bottom, four (k_x, k_y) sections of $P(k_x, k_y, \omega)$ for $\omega = 2, 3, 5$ and $10\ \mathrm{hr}^{-1}$, origin in the centre. The black line is the theoretical singularity (dispersion) curve ($|\tilde{g}_{wav}|^2$ with $\tilde{g}_{wav}$ from Eq. 19), the blue the empirically estimated curve using the ad hoc algorithm, and the green and red show maxima but are presumed to originate in the turbulence "background" (they are very close to the axes).

more than the simplest analytical hypotheses, the results are quite encouraging, yet they indicate some of the difficulties.

6   Conclusion

The atmosphere is highly nonlinear yet displays both turbulent and wave-like behaviour over huge ranges of space–time scales. Theories explaining the turbulent aspects assume that the dynamics are strongly nonlinear and scaling; in contrast, the corresponding wave theories are generally linear or weakly nonlinear. We proposed that the paradox can be resolved by noting that, although linear theory predicts propagators, only the relations implied by the singular part of the latter – the dispersion relations – have been tested to any extent. However, linear theories invariably involve integer-ordered differential operators and corresponding integer-ordered propagators, so we may test these theories by examining the propagators – or at least their squared moduli which are amenable to empirical spectral determination. Although linear theories are commonly extended using the WKB (Wentzel–Kramers–Brillouin) approximation, this

amounts to a transformation of variables, which itself is only valid if it is weakly nonlinear: it cannot compensate for the neglect of the nonlinear terms. Yet even aside from the issue of linearization, at least in the classical case of gravity waves (Appendix C), we showed that in any case – due to turbulence – the necessary WKB smoothness conditions were rarely satisfied so that the method cannot be justified.

The mathematical structure of the turbulent laws that link the observables to driving turbulent fluxes (such as energy fluxes) use scaling (turbulent) propagators which are very similar to those of wave equations except that the latter are singular. To account for both waves and turbulence, the actual propagators need only respect scale symmetries and can be modelled as products of turbulent-like and wave-like (space–time localized and unlocalized) propagators, with both involving anomalous exponents. The wave propagator involves the mean horizontal turbulent wind and energy flux as well as a mean background wave advection velocity. Although in Appendix B we give the more general results, in the case studied here it was obtained as an (anisotropic and fractional) generalization of the classical wave equation (which is approximately satisfied by inertial gravity waves and Kelvin

waves). Using 2 months of MTSAT hourly geostationary IR radiances, using a simple form for the wave propagator, we found that the best fit involved an anomalous wave scaling exponent $H_{wav} \approx 0.17 \pm 0.04$; for comparison, the classical wave equation has the integer value $H_{wav} = 2$. Since the propagator determines the rate at which the forcing decays in time (Appendix B), such an anomalous exponent will affect the transport of energy and momentum and thus has implications for the dynamics.

Investigating the wave structure, we followed Wheeler and Kiladis (1999), dividing our empirical spectral density by a theoretically estimated turbulent background; the maxima in the residual are to be identified with wave dispersion relations. Since there are only weak theoretical constraints on the form of the wave propagator, we chose a simple ansatz that is compatible with the observations.

This paper is simply an early attempt to understand waves in highly turbulent media using scaling symmetries as constraints. On the one hand, these symmetries are so broad that they provide only limited guidance; on the other hand, the turbulent part – with no wave contribution at all – explains almost all of the observed dynamical spectral scaling range of factor $\approx 10^5$, leaving only a small residual (a factor which has a mean of 0.9 ± 0.5) for the wave-like part. The empirical situation would certainly be improved if data from other fields covering wide scale ranges in the full (x, y, z, t) space could be found, but at present the (x, y, t) geostationary radiances are apparently the best available. Therefore, this paper should be seen more as a proof of concept than as providing definitive results. The main conclusion is thus that a priori strongly turbulent atmospheric dynamics are compatible with the observed waves. If this is true, to understand them requires neither the existence of large laminar regimes nor linear theories.

Linear theories of atmospheric waves were developed starting in the 1930s, well before the ongoing revolutions in both computing and in atmospheric data. At the time, they were attractive because they provided a rare tool for theoretically understanding atmospheric dynamics. As atmospheric science becomes increasingly dominated by brute force numerical modelling, they remain attractive as being among the last analytically tractable approaches to the subject. In addition – at least so far – they have survived the "golden age" of meteorological data because even today few data sets exist with high-enough space–time resolutions to fully test them and they have had some apparent successes. Nevertheless, we find their theoretical bases implausible, and detailed examination of both gravity and Kelvin waves suggests they are empirically untenable. It is likely that their success lies in their – at least rough – capture of the appropriate scaling symmetries which constrain the system at both low and high Reynolds number.

References

Bleistein, N. and Handelsman, R. A.: Asymptotic Expansions of Integrals, Dover, Mineola, 425 pp., 1986.

Dias, J., Tulich, S. N., and Kiladis, G. N.: An object based approach to assessing tropical convection organization. J. Atmos. Sci., 69, 2488–2504, 2012.

Frank, W. M. and Roundy, P. E.: The Role of Tropical Waves in Tropical Cyclogenesis. Mon. Weather Rev. 134, 2397–2417, 2006.

Frisch, U.: Turbulence: the legacy of A. N. Kolmogorov, Cambridge University Press, Cambridge, 296 pp., 1995.

Gill, A. E.: Atmosphere-Ocean Dynamics, Academic Press, New York, , 452 pp., 1982.

Hendon, H. H. and Wheeler, M.: Some space-time spectral analyses of tropical convection and planetary waves, *J. Atmos. Sciences*, 65, 2936–2948, 2008.

Kiladis, G. N., Wheeler, M. C., Haertel, P. T., Straub, K. H., and Roundy, P. E.: Convectively coupled equatorial waves, Rev. Geophys., 47, RG2003, doi:10.1029/2008RG000266, 2009.

Lin, Y.-L.: Two-dimensional response of a stably stratified shear flow to diabatic heating. J. Atmos. Sci., 44, 1375–1390, 1987.

Lin, Y.-L.: Mesoscale Dynamics. Cambridge University Press, 630 pp., 2007.

Lovejoy, S. and Schertzer, D.: Towards a new synthesis for atmospheric dynamics: space-time cascades, Atmos. Res., 96, 1–52, doi:10.1016/j.atmosres.2010.1001.1004, 2010.

Lovejoy, S. and Schertzer, D.: The Weather and Climate: Emergent Laws and Multifractal Cascades, Cambridge University Press, Cambridge, 496pp., 2013.

Lovejoy, S., Tuck, A. F., Hovde, S. J., and Schertzer, D.: Is isotropic turbulence relevant in the atmosphere? Geophys. Res. Lett., L14802, doi:10.1029/2007GL029359, 2007.

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

Lovejoy, S., Schertzer, D., Lilley, M., Strawbridge, K. B., and Radkevitch, A.: Scaling turbulent atmospheric stratification, Part I: turbulence and waves, Q. J. Roy. Meteor. Soc., 134, 277–300, doi:10.1002/qj.1201, 2008b.

Matsuno, T.: Quasi-geostrophic motions in the equatorial area, J. Meteorol. Soc. Jpn., 44, 25–43, 1966.

Miller, K. S. and Ross, B.: An introduction to the fractional calculus and fractional differential equations, Wiley-Interscience, New York, 365 pp., 1993.

Nappo, C. J.: An introduction to Gravity Waves, 276 pp., Academic Press, Amsterdam, 2002.

Pielke, R. A.: *Mesoscale meteorological modeling*, 2nd Edition, Academic Press, San Diego, CA, 676 pp., 2002.

Pinel, J., Lovejoy, S., and Schertzer, D.: The horizontal space-time scaling and cascade structure of the atmosphere and satellite radiances, Atmos. Res., 140–141, 95–114, doi:10.1016/j.atmosres.2013.11.022, 2014.

Placke, M., Hoffmann, P., Gerding, M., Becker, E., and Rapp, M.: Testing linear gravity wave theory with simultaneous wind and temperature data from the mesosphere, J. Atmos. Sol.-Terr. Phys., 93, 57–69, 2013.

Schertzer D. and Lovejoy, S.: The dimension and intermittency of atmospheric dynamics, Turbulent Shear Flow 4, edited by: Launder B., Springer-Verlag: Berlin, 7–33, 1985a.

Schertzer, D. and Lovejoy, S.: Generalised scale invariance in turbulent phenomena, Phys.-Chem. Hydrodynam. J., 6, 623–635, 1985b.

Schertzer, D. and Lovejoy, S.: Physical Modeling and Analysis of Rain and Clouds by Anisotropic Scaling Multiplicative Porcesses. J. Geophys. Res., 92, 9692–9714, 1987.

Schreck, C. J., Molinari, J., and Aiyyer, A.: A global view of equatorial waves and tropical cyclogenesis. Mon. Weather Rev, 140, 774–787, 2012.

Shen, B.-W., Tao, W.-K., Lin, Y.-L., and Laing, A.: Genesis of Twin Tropical Cyclones as Revealed by a Global Mesoscale Model: The Role of Mixed Rossby Gravity Waves, J. Geophys. Res., 117, D13114, doi:10.1029/2012JD017450, 2012.

Shen, B. W., DeMaria, M., Li, J.-L. F., and Cheung, S.: Genesis of Hurricane Sandy (2012) Simulated with a Global Mesoscale Model, Geophys. Res. Lett., 40, 4944–4950,, doi:10.1002/grl.50934, 2013.

Smith, R. B.: The influence of mountains on the atmosphere, Advances in Geophysics, Academic Press, 21, 87–230, 1979.

Tuck, A. F.: From molecules to meteorology via turbulent scale invariance, Q. J. Roy. Meteor. Soc., 136, 1125–1144, 2010.

Wheeler, M. and Kiladis, G. N.: Convectively coupled equatorial waves: analysis of clouds and temperature in the wavenumber-frequency domain, J. Atmos. Sciences, 56, 374–399, 1999.

Appendix A

The space–time turbulent spectrum

The 23/9D model of spatial turbulence (Schertzer and Lovejoy, 1985a, b) involves wide-range scaling in the horizontal and vertical directions but with different scaling exponents – the horizontal being dominated by energy fluxes, and the vertical by buoyancy variance fluxes. Since the infrared radiances are essentially (x, y, t) (horizontal-time) fields, we need not explicitly consider the vertical; however we do need to extend the model to space-time. In this appendix, we summarize the arguments developed in Lovejoy et al. (2008b), Lovejoy and Schertzer (2010, 2013) and in Pinel et al. (2014).

The first step is to rewrite the isotropic Eq. (6) in a more general anisotropic scaling manner by replacing the vector norm by a space–time scale function $[[\Delta R]]$:

$$\Delta I (\Delta R) = \phi_{[[\Delta R]]} [[\Delta R]]^H, \tag{A1}$$

where we have used the subscript $[[\Delta R]]$ to emphasize that the flux is at resolution $[[\Delta R]]$. The space–time scale function is symmetric with respect to generalized scale changes T_λ:

$$[[T_\lambda \Delta R]] = \lambda^{-1} [[\Delta R]]; \quad T_\lambda = \lambda^G, \tag{A2}$$

where G is the generator of the scale-changing operator T_λ. Ignoring for the moment advection, and taking I as a horizontal velocity component, the canonical (simplest) nondimensional scale function compatible with the Kolmogorov law is

$$[[\Delta R]] = L_w \left\{ \left(\frac{\Delta x}{L_w}\right)^2 + \left(\frac{\Delta y}{L_w/a}\right)^2 + \left(\frac{\Delta t}{\tau_w}\right)^{2/H_t} \right\}^{1/2}, \tag{A3}$$

where $H_t = (1/3)/(1/2) = 2/3$, L_w and τ_w are the outer scales of the scaling in space and in time, and a is a north–south/east–west aspect ratio. The outer scales are linked by the overall average energy flux ε: $\tau_w = \varepsilon^{-1/3} L_w^{2/3}$ (τ_w is the lifetime/"eddy-turn-over time" of structures size L_w). Successively substituting $\Delta R = (\Delta x, 0, 0)$, $\Delta R = (0, \Delta y, 0)$ and $\Delta R = (0, 0, \Delta t)$ into Eq. (A3) and the latter into Eq. (A1) yields the Kolmogorov law in the horizontal directions (e.g. $\Delta v = \varepsilon^{1/3} \Delta x^{1/3}$) and the Lagrangian law in time ($\Delta v = \varepsilon^{1/2} \Delta t^{1/2}$).

The next step is to consider the scale function corresponding to a constant advection in the horizontal $v = (v_x, v_y)$. Due to Galilean invariance (i.e. under $x \to x - v_x \Delta t$; $y \to y - v_y \Delta t$; $\Delta t \to \Delta t$), we obtain

$$[[\Delta R]] = L_w \left\{ \left(\frac{\Delta x - v_x \Delta t}{L_w}\right)^2 + \left(\frac{\Delta y - v_y \Delta t}{L_w/a}\right)^2 + \left(\frac{\Delta t}{\tau_w}\right)^{2/H_t} \right\}^{1/2} \tag{A4}$$

Since there is no scale separation, v is a turbulent velocity, so that over a given region it will be dominated by the advection due to the largest eddies. For the same reasons, the link between L_w and τ_w is via a turbulent velocity, a consequence of which is that the pure temporal development (Lagrangian) term may be neglected.

Denoting the scale function obtained by averaging over the ensemble of different advection velocities as the "effective" scale function, it is obtained by averaging

$$[[\Delta R]]^2 = \left(\frac{\Delta x - v_x \Delta t}{L_w}\right)^2 + \left(\frac{\Delta y - v_y \Delta t}{L_w/a}\right)^2 \tag{A5}$$

over the turbulence (this argument is not completely rigorous since, due to the intermittency, averaging over other powers of $[[\Delta R]]$ will give somewhat different parameters). Turbulence will have two effects on Eq. (A5): the mean advection by v, and the effect of turbulent variability. The overall result (for more details, see Pinel and Lovejoy, 2014) is

$$\left\langle \Delta I (\Delta R)^2 \right\rangle = \left\langle \phi^2_{[[\Delta R]]_{\text{eff}}} \right\rangle [[\Delta R]]^{2H}_{\text{eff}}, \tag{A6}$$

with the effective scale function given by

$$[[\Delta R]]_{\text{eff}} = \left(\Delta R^T B \Delta R\right)^{1/2}; \underline{B} = \begin{pmatrix} 1 & 0 & -\mu_x \\ 0 & a^2 & -a^2\mu_y \\ -\mu_x & -a^2\mu_y & 1 \end{pmatrix}, \tag{A7}$$

where we have used the nondimensional variables

$$\Delta R \to \Delta R = \left(\frac{\Delta x}{L_w}, \frac{\Delta y}{L_w}, \frac{\Delta t}{\tau_w}\right); \quad \mu = (\mu_x, \mu_y) = (\overline{v_x}, \overline{v_y})/V_w, \tag{A8}$$

with $\overline{v} = (\overline{v_x}, \overline{v_y})$ the overall mean advection over the region considered and $V_w = (\overline{v_x^2} + a^2 \overline{v_y^2})^{1/2}$ the large-scale turbulent velocity at planetary scale.

The intermittency corrections come from the scaling of the flux ϕ:

$$\left\langle \phi^2_{[[\Delta R]]_{\text{eff}}} \right\rangle \approx [[\Delta R]]^{-K(2)}_{\text{eff}}, \tag{A9}$$

so that, overall,

$$\left\langle \Delta I (\Delta R)^2 \right\rangle \approx [[\Delta R]]^{\xi(2)}_{\text{eff}}. \tag{A10}$$

The structure function exponent $\xi(2)=2H-K(2)$ thus takes into account the scaling exponent H as well as the intermittency correction.

To obtain the corresponding spectral density $P_I(k, \omega)$ (and hence $\tilde{g}_{\text{tur}}(k, \omega)$), we follow the development presented by Lovejoy and Schertzer (2010) (see also Pinel et al. (2014)) and use the general relation between structure functions and spectra: $\left\langle |\Delta I (\Delta R)|^2 \right\rangle = 2 \int_{-\infty}^{\infty} dK \left(1 - e^{iK \cdot \Delta R}\right) P_I(K)$, with nondimensionalized wave vector

$$K = \left(L_w k_x, L_w k_y, \tau_w \omega\right),$$

so that the effective real space scale function (that takes into account the averaging over an ensemble of advection velocities) defines an effective Fourier space scale function:

$$[[\boldsymbol{K}]]_{\text{eff}} = \left(\boldsymbol{K}^T \boldsymbol{B}^{-1} \boldsymbol{K}\right)^{-1/2};$$ (A11)

$$\boldsymbol{B}^{-1} = \begin{pmatrix} 1 - a^2\mu_y & \mu_x\mu_y & \mu_x \\ \mu_x\mu_y & \left(1 - \mu_x^2\right)/a^2 & \mu_y \\ \mu_x & \mu_y & 1 \end{pmatrix}.$$

Using

$$\omega' = (\omega + \boldsymbol{k} \cdot \boldsymbol{\mu})\sigma^{-1}; \quad \sigma = \left(1 - \left(\mu_x^2 + a^2\mu_y^2\right)\right)^{1/2},$$ (A12)

this can be simplified to $[[\boldsymbol{K}]]_{\text{eff}} = \left(\omega'^2 + \|k\|^2\right)^{1/2}$, where $\|\boldsymbol{k}\|^2 = k_x^2 + \left(k_y/a\right)^2$ is the (horizontal) spatial (Fourier) scale function. The transformation $\omega \to \omega'$ combines the effects of a mean advection by velocity $\boldsymbol{\mu}$, and the statistical variability of the advection wind about its mean is accounted for by σ.

With this scale function, we have the nondimensional spectra

$$P_I(\boldsymbol{K}) = |\tilde{g}_{\text{tur}}|^2 P_\phi(\boldsymbol{K}); \quad |\tilde{g}_{\text{tur}}|^2 = [[\boldsymbol{K}]]_{\text{eff}}^{-2H};$$ (A13)

$$P_\phi = [[\boldsymbol{K}]]_{\text{eff}}^{-d+K(2)}$$

(with $d = 3$ for horizontal space–time).

Appendix B

The general form for space–time propagator, dispersion relations, group velocities and wave amplitudes

In sect. 3, we discussed some specific examples of both space–time localized (turbulent) and space–time unlocalized (wave-like) propagators; here we outline their general properties.

B1 Scaling

The basic scaling symmetry of the Fourier transform of the propagator is

$$\tilde{g}\left(\lambda^{\tilde{\mathbf{G}}} \mathbf{K}\right) = \lambda^{-H} g\left(\mathbf{K}\right) \; ; \; K = (\mathbf{k}, \omega); \quad \mathbf{k} = \left(k_x, k_y\right), \quad \text{(B1)}$$

where $\tilde{\mathbf{G}}$ is a matrix; it is the transpose of the real space generator G of the scale-changing operator $T_\lambda = \lambda^{-G}$ discussed in Appendix A. This generality is needed at minimum to take into account the vertical stratification. The solution is

$$\tilde{g}\left(\mathbf{K}\right) = [[\mathbf{K}]]^{-H}, \quad \text{(B2)}$$

where the space–time (Fourier) scale function satisfies

$$\left[\left[\lambda^{\tilde{\mathbf{G}}} \mathbf{K}\right]\right] = \lambda \, [[\mathbf{K}]]. \quad \text{(B3)}$$

However, for the horizontal radiance field that we empirically found (Fig. 2), $\tilde{\mathbf{G}} = $ identity matrix, in this case $\lambda^{\tilde{\mathbf{G}}} = \lambda$, and the solution in terms of unit vectors $\hat{\mathbf{K}}$ is

$$[[\mathbf{K}]] = |\mathbf{K}| \Theta\left(\hat{\mathbf{K}}\right); \quad \hat{\mathbf{K}} = \frac{\mathbf{K}}{|\mathbf{K}|}, \quad \text{(B4)}$$

where Θ is a (complex) function of the unit vector $\hat{\mathbf{K}}$; this can be easily checked. When $\tilde{G}$ is not the identity matrix, we must first diagonalize the system and then perform a nonlinear coordinate transformation after which the same basic solution method may be applied; see Lovejoy and Schertzer (2013) for details.

B2 Causality

Causality implies that $g(\mathbf{r}, t) = 0$ for $t < 0$: otherwise the future will influence the past. The consequence for the Fourier propagator is that there cannot be any singularities/poles of $\tilde{g}(\mathbf{K})$ in the upper half complex ω plane; therefore we rewrite $[[\mathbf{K}]]$ as a product of zero and nonzero parts:

$$[[K]] = \left[(-i\omega + \|\mathbf{k}\|) F\left(\hat{\mathbf{K}}\right)\right], \quad \text{(B5)}$$

where F is an arbitrary nonzero function of unit vector $\hat{\mathbf{K}}$ and $\|\mathbf{k}\|$ is the spatial (Fourier) scale function. It satisfies

$$\|\lambda \mathbf{k}\| = \lambda \|\mathbf{k}\| \; . \quad \text{(B6)}$$

The solution $\|\mathbf{k}\|$ of this scaling equation is

$$\|\mathbf{k}\| = |\mathbf{k}| \Phi\left(\hat{\mathbf{k}}\right); \; \hat{\mathbf{k}} = \frac{\mathbf{k}}{|\mathbf{k}|} \; . \quad \text{(B7)}$$

The propagator poles are defined by the zeroes of $[[\mathbf{K}]]$, i.e. the frequencies ω satisfying

$$\omega = -i |\mathbf{k}| \Phi\left(\hat{\mathbf{k}}\right). \quad \text{(B8)}$$

Causality now imposes the condition

$$Re\left(\|\mathbf{k}\|\right) > 0 \quad \text{(B9a)}$$

or, equivalently,

$$Re\left(\Phi\left(\hat{\mathbf{k}}\right)\right) > 0. \quad \text{(B9b)}$$

Note that advection by the deterministic vector $\boldsymbol{\mu}$ is equivalent to the transformation $\omega \to \omega + \mathbf{k} \cdot \boldsymbol{\mu}$; this is equivalent to $\Phi\left(\hat{\mathbf{k}}\right) \to \Phi\left(\hat{\mathbf{k}}\right) - i |\boldsymbol{\mu}| \cos\left(\hat{\mathbf{k}} \cdot \boldsymbol{\mu}\right)$. In Appendix A it is argued that a random advection can be modelled by the $\omega \to (\omega + \mathbf{k} \cdot \boldsymbol{\mu})/\sigma$ which is equivalent to $\Phi\left(\hat{\mathbf{k}}\right) \to \sigma\left(\Phi\left(\hat{\mathbf{k}}\right) - i\left(\hat{\mathbf{k}} \cdot \boldsymbol{\mu}\right)\right)$.

B3 Reality

If the forcing and responses are real, then the propagator g is real and its Fourier transform satisfies

$$\tilde{g}\left(\mathbf{K}\right) = \tilde{g}^*\left(\mathbf{K}\right), \quad \text{(B10)}$$

where the asterisk indicates complex conjugate. Hence

$$[[-\mathbf{K}]]^* = [[\mathbf{K}]] \quad \text{(B11)}$$

and

$$\Theta\left(\hat{\mathbf{K}}\right) = \Theta^*\left(-\hat{\mathbf{K}}\right); \Phi\left(\hat{\mathbf{k}}\right) = \Phi^*\left(-\hat{\mathbf{k}}\right); F\left(\hat{\mathbf{k}}\right) = F^*\left(-\hat{\mathbf{k}}\right). \quad \text{(B12)}$$

Overall, we have

$$\tilde{g}\left(\mathbf{k}, \omega\right) = [[\mathbf{K}]]^{-H} = \left[\left(-i\omega + |\mathbf{k}| \Phi\left(\hat{\mathbf{k}}\right)\right) F\left(\hat{\mathbf{K}}\right)\right]^{-H}, \quad \text{(B13)}$$

where Φ, F satisfy Eq. (B12).

Alternatively, in the usual wave tradition, we can use complex waves and simply take real parts when needed. In this case, we can ignore the reality condition (Eq. B10), choosing $\tilde{g}$ simply to satisfy the scaling and causality conditions, and then – if needed – construct a $\tilde{g}$ of a real propagator by taking $\tilde{g}\left(\mathbf{K}\right) \to \frac{1}{2}\left(\tilde{g}\left(\mathbf{K}\right) + \tilde{g}^*\left(-\mathbf{K}\right)\right)$.

B4 Dispersion relations and waves

The propagator form B5 is convenient for discussing the turbulent case since space–time localisation will lead to that form with corresponding $\|\mathbf{k}\| > 0$ – i.e with no poles (see the

discussion in Lovejoy et al. (2008)). However, when there are poles, the latter define dispersion relations and it is convenient to consider the form:

$$\tilde{g}_{\text{wav}}(\boldsymbol{k}, \omega) = [[\boldsymbol{K}]]^{-H_{\text{wav}}};$$ (B14)

$$[[\boldsymbol{K}]] = (-i(\omega - \omega_d(\boldsymbol{k}))) F_{\text{wav}}\left(\hat{\boldsymbol{K}}\right),$$

where $\omega_d(\boldsymbol{k})$ defines the dispersion relation. In this case, the scaling and causality constraints impose

$$\omega_d(\boldsymbol{k}) = \|\boldsymbol{k}\|_{\text{wav}}; \quad \|\lambda\boldsymbol{k}\|_{\text{wav}} = \lambda\|\boldsymbol{k}\|_{\text{wav}}; \quad Im\|\boldsymbol{k}\|_{\text{wav}} < 0 \quad \text{(B15)}$$

(the reality condition can be treated as indicated above). Finally, $F_{\text{wav}}\left(\hat{\boldsymbol{K}}\right)$ is a nonzero function of the unit vector $\hat{\boldsymbol{K}}$ satisfying $F_{\text{wav}}\left(\hat{\boldsymbol{K}}\right) = F_{\text{wav}}^*\left(-\hat{\boldsymbol{K}}\right)$. However, the overall propagator is the product of the wave and turbulence propagators:

$$\tilde{g}_I = \tilde{g}_{\text{tur}}\tilde{g}_{\text{wav}} = \left[(-i\omega + \|\boldsymbol{k}\|) F_{\text{tur}}\left(\hat{\boldsymbol{K}}\right)\right]^{-H_{\text{tur}}} \quad \text{(B16a)}$$

$$\left[(-i(\omega - \omega_d(\boldsymbol{k}))) F_{\text{wav}}\left(\hat{\boldsymbol{K}}\right)\right]^{-H_{\text{wav}}},$$

so that without loss of generality we may take $F_{\text{wav}} = 1$, effectively absorbing it into F_{tur}:

$$\tilde{g}_I = \left[(-i\omega + \|\boldsymbol{k}\|) F_{\text{tur}}\left(\hat{\boldsymbol{K}}\right)\right]^{-H_{\text{tur}}} [(-i(\omega - \omega_d(\boldsymbol{k})))]^{-H_{\text{wav}}}. \quad \text{(B16b)}$$

In order to see what such Fourier propagators imply for the real space behaviour, first take the inverse Fourier transform with respect to ω:

$$g_{\text{wav}}(\boldsymbol{k}, t) = \int_{-\infty}^{\infty} e^{-i(\boldsymbol{k}\cdot\boldsymbol{r} + \omega t)} g_{\text{wav}}(\boldsymbol{k}, \omega) \, d\omega \quad \text{(B17)}$$

$$= \int_{-\infty}^{\infty} e^{-i(\boldsymbol{k}\cdot\boldsymbol{r} + \omega t)} (-i(\omega - \omega_d(\boldsymbol{k})))^{-H_{\text{wav}}} \, d\omega$$

$$= \Theta(t) t^{-1+H_{\text{wav}}} e^{-i(\boldsymbol{k}\cdot\boldsymbol{r} + \omega_d(d)t)},$$

where we have ignored constant numerical factors and $\Theta(t)$ is the Heaviside function ($\Theta(t) = 0$ for $t < 0$, otherwise $= 1$; it assures that causality is respected). We may now use the method of stationary phase (e.g. Bleistein and Handelsman (1986), and see Lovejoy et al. (2008) for the application to gravity waves) to estimate the real space propagator $g_{\text{wav}}(\boldsymbol{r}, t)$:

$$g_{\text{wav}}(\boldsymbol{r}, t) \quad \text{(B18)}$$

$$\approx \Theta(t) t^{-d/2 - 1 + H_{\text{wav}}} \left[\det\left(\frac{\partial^2 \omega_d(\boldsymbol{k})}{\partial k_i \partial k_j}\right)\right]^{-1/2} e^{-i(\boldsymbol{k}\cdot\boldsymbol{r} + \omega_d(k)t + \varphi_0(\boldsymbol{k}))},$$

where d is the dimension of space; for the horizontal space discussed in the paper, $d = 2$ and $\varphi_0(\boldsymbol{k})$ is a phase, and "det" indicates determinant. Equation (B18) is a parametric equation; $\boldsymbol{k}$ is the wave vector which satisfies the "ray" equation:

$$\boldsymbol{r} = -\boldsymbol{v}_g(\boldsymbol{k}) t, \quad \text{(B19)}$$

where $\boldsymbol{v}_g$ is the group velocity:

$$\boldsymbol{v}_g(\boldsymbol{k}) = \nabla\omega_d(\boldsymbol{k}). \quad \text{(B20)}$$

The negative sign in Eq. (B19) is due to the fact that the Fourier propagator $\tilde{g}_{\text{wav}}(\boldsymbol{k}, t)$ is the amplitude of a wave travelling in the direction $-\boldsymbol{k}$ as discussed in Sect. 3. When the determinant vanishes, the approximation breaks down; the corresponding rays are the "caustics".

Of more interest to our present discussion is the amplitude factor $t^{-d/2 - (1 - H_{\text{wav}})}$ which quantifies the rate at which an impulse forcing at $t = 0$ decays as the disturbance spreads. For comparison, the classical wave equation corresponds to $H_{\text{wav}} = 1$ so that in 2-D the waves decay as t^{-1} (to see this, in Eq. (2), use the identity $\left(\omega^2/V^2 - |\boldsymbol{k}|^2\right)^{-1} = \frac{V}{2\omega}\left[(\omega/V - |\boldsymbol{k}|)^{-1} + (\omega/V + |\boldsymbol{k}|)^{-1}\right]$; i.e. there are two terms each with $H = 1$). We see that smaller values of H_{wav} correspond to waves that decay more quickly in time. Although the precise implications of this for energy transport will depend on the physical nature of the field I whose response is described by g, one generally finds that the power p is proportional to the wave amplitude squared multiplied by the group velocity, i.e. $p \propto |g|^2 v_g$; hence the amplitude factor for the wave power is $\propto t^{-d - 2(1 - H_{\text{wav}})}$. If the observed $H_{\text{wav}} < 1$ (as is the case here), then the amplitudes decays anomalously quickly and, when compared with a classical wave propagator, the fractional propagator will deposit the energy more locally.

The above calculation of $g_{\text{wav}}(\boldsymbol{r}, t)$ is fairly straightforward; if we use the full turbulence-wave Fourier propagator $g_I(\boldsymbol{k}, \omega) = g_{\text{tur}}(\boldsymbol{k}, \omega) g_{\text{wav}}(\boldsymbol{k}, \omega)$, this corresponds to $g_I(\boldsymbol{r}, t) = g_{\text{tur}}(\boldsymbol{r}, t)^* g_{\text{wav}}(\boldsymbol{r}, t)$, where "*" indicates "convolution". Alternatively, as indicated in Sect. 3 since both the forcing turbulent flux ϕ and the turbulent propagator g_{tur} are space–time localized, we may consider that the overall forcing is $\phi' = g_{\text{tur}}^*\phi$; the propagator $g_{\text{wav}}(\boldsymbol{r}, t)$ thus transports the disturbance ϕ'.
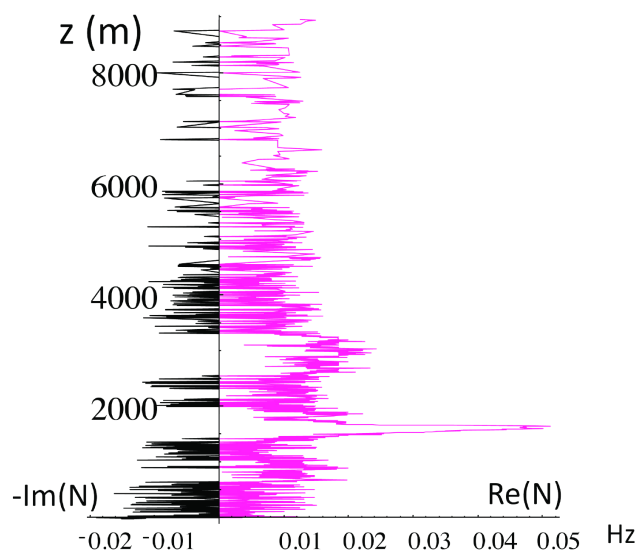
Fig. C1. This shows the Brunt–Vaisälä frequency N (in Hz) as estimated by a single dropsonde over the Pacific; only estimates across layers <20 m thick were used, and it was typical of the 238 sondes analysed in the Pacific 2004 experiment. Analysis of dropsonde pairs (Lovejoy et al., 2008a) indicated that the measurement errors are less than the thickness of the lines. The right-hand side shows the real part, the left hand side the negative of the imaginary part (these unstable layers constituted 15 % of the all the values, and this was typical).

Appendix C

### The WKB approximation, the example of gravity waves

The first step in formulating a linear wave theory is to linearize the equations by eliminating various nonlinear terms, a procedure that is valid only when they are smaller than the retained terms. In the simplest cases, the resulting linear equations have constant coefficients. Since the coefficients are determined by a presumed mean state around which the linearization has been performed, the assumption that the coefficients are constant is very artificial so that linear theory is often extended to cases where the coefficients are slowly varying: the WKB approximation (Wentzel–Kramers–Brillouin, also known as the Liouville–Green method). It is therefore conceivable that, although it does not compensate for the neglect of the nonlinear terms, the WKB theory might provide a more plausible basis for linear theories.

Rather than continue the discussion at a purely theoretical level, let us consider the classical atmospheric application of the WKB method: to gravity waves. Consider the classical wave equation in the z direction only, without forcing and for frequencies N. We follow the standard development from Gill (1982). In this case Eq. (1) reduces to

$$\frac{d^2 w}{dz^2} + m^2 w = 0; \quad m = \frac{N}{V}; \quad I(z, t) = w(z) e^{iNt}. \quad \text{(C1)}$$
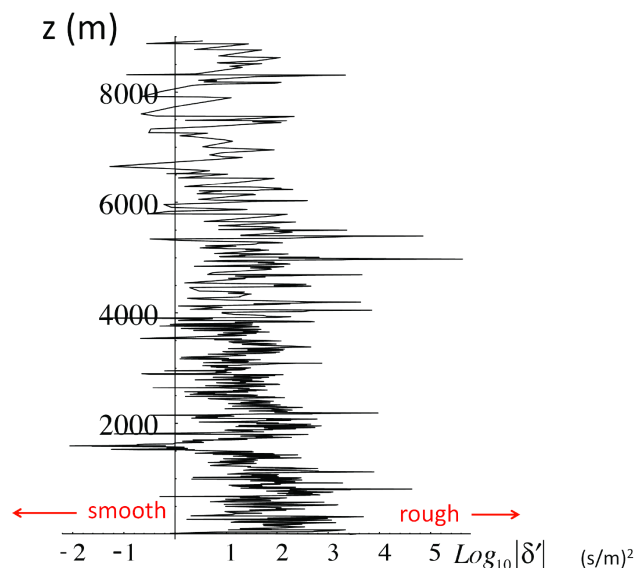
Fig. C2. This shows $\text{Log}_{10}|\delta'|$ where $\delta' = N^{-3/2} \frac{d^2 N^{-1/2}}{dz^2}$ in $(\text{s m}^{-1})^2$. $\delta = V^2 \delta'$ is the nondimensional smoothness parameter determining the applicability of the WKB approximation and V is the wave speed (assumed constant). For the WKB approximation to be valid, we require $\delta \ll 1$.
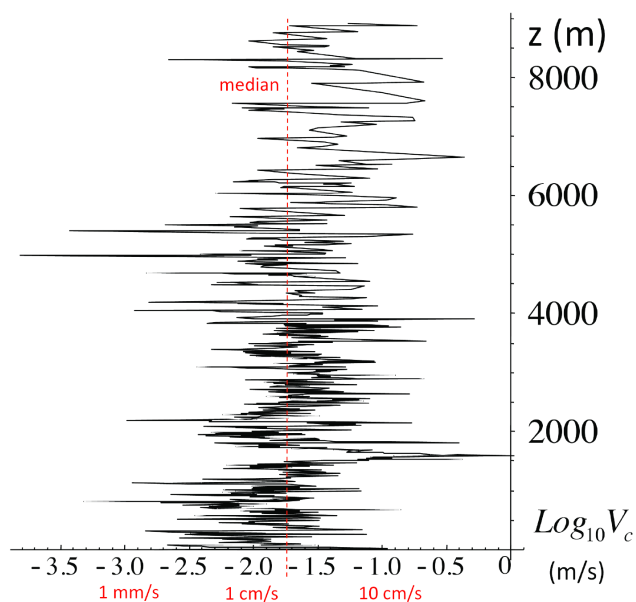
Fig. C3. This shows $\text{Log}_{10} V_c$ where $V_c = (\delta_c / \delta')^{1/2}$ is the critical speed, δ_c is the critical value of the nondimensional smoothness parameter δ, and here we take $\delta_c = 0.01$. For the application of the WKB method, we require $V < V_c$. The median V_c (1.9 cm s^{-1}) is shown by the dashed red line.

Where w is the vertical component of the wind, m is the vertical wave number and $N = (d \log \theta / dz)^{1/2}$ is the Brunt–Vaisälä frequency, (θ is the potential temperature), Eq. (C1) is a special case of the Taylor–Goldstein equations. If N and

V (hence m) are independent of z, then we have the solutions

$$w = Ae^{\pm imz}. \tag{C2}$$

The WKB method extends this to the case where m varies slowly. The approximation starts by defining the phase

$$\phi = \int m\, dz, \tag{C3}$$

from which the approximate solution is obtained:

$$w = m^{-1/2}e^{\pm i\phi}. \tag{C4}$$

However, the condition of validity of the approximation is that ϕ varies slowly; it is quite strict:

$$\delta = m^{-3/2}\frac{d^2 m^{-1/2}}{dz^2} \ll 1. \tag{C5}$$

Rather than taking the theoretical analysis further, we empirically check whether this smoothness criterion is satisfied by using high-resolution dropsondes. Although the wave speed V is not readily empirically accessible, N is accessible and, in any case, V is often assumed to be constant – as for example in the case of hydrostatic waves treated in Gill (1982). If we take $V = $ constant, then the method is applicable for small δ:

$$\delta = V^2 \delta' \ll 1;\ \delta' = N^{-3/2}\frac{d^2 N^{-1/2}}{dz^2}. \tag{C6}$$

In Lovejoy et al. (2008a) data from 238 dropsondes from the Pacific 2004 experiment were analysed with vertical resolutions varying from ≈ 5 to ≈ 12 m (depending on altitude, but with frequent outages; see Lovejoy et al. (2007) for details of the experiment). Three instability criteria were considered: conditional instability corresponding to $N^2 < 0$ (relevant here), as well as dynamical instability and convective instability; Fig. C1 shows a typical example. The occasional use of sonde pairs directly showed that the measurement errors were small (less than the thickness of the lines in the figure). In all cases it was found that the atmosphere is actually a fractal hierarchy of unstable layers. Lovejoy and Schertzer (2013) give a summary as well as an empirical demonstration that the linearization assumptions needed to obtain the Taylor–Goldstein were generally badly violated.

The vertical thus has a large number of layers with imaginary N in which the gravity waves would be evanescent (see the black lines in Fig. C1); this is already quite problematic for the method. However, with the help of the WKB approximation, the smoothness condition can also be examined. Fig. C2 shows the estimate of δ': it is seen that at the sonde resolution it varies wildly; for example over half the layers had $\delta' > 25$ (s m^{-1})^2, while only 6 % of the layers had $\delta' < 1$ (s m^{-1})^2. This places strong limits on the maximum wave speed V that is compatible with the WKB approximation. For example, if we interpret the condition $\delta \ll 1$ as a requirement that $\delta < \delta_{\rm c}$ (for a small $\delta_{\rm c}$), then we require

$$V < V_{\rm c};\ V_{\rm c} = \left(\delta_{\rm c}/\delta'\right)^{1/2}. \tag{C7}$$

Taking $\delta_{\rm c} = 0.01$, we obtain the limits shown in Fig. C3 where we find that 90% of the time, in order for the profile to be smooth enough, the wave speed must be less than 7.5 cm s^{-1}, and that 1% of the time it must be less than 1.4 mm/s. Since gravity waves are usually considered to have wave speeds of the order of several m/s, these very low values are highly problematic for the application of the WKB approximation.

A scaling analysis of the statistics of the fluctuations $\Delta\delta$ shows that $\left\langle \Delta\delta\,(\Delta z)^q \right\rangle \approx \Delta z^{\xi(q)}$. Although $\xi(1) \approx 0.1$ indicates that in the small-scale limit ($\Delta z \to 0$) the mean $\Delta\delta$ does decrease, the intermittency (due to the multifractality) is so extreme that for example the root mean square value $\left\langle \Delta\delta\,(\Delta z)^2 \right\rangle^{1/2}$ has an exponent $\xi(2)/2 \approx -0.35$, indicating that in fact it rapidly diverges in the small-scale limit. In comparison, the corresponding exponents for $\Delta(Re(N))$ are $\xi(1) \approx 0.0$ and $\xi(2)/2 \approx -0.1$, i.e. much less intermittent but with the same qualitative behaviours.