

Waves and coherent flows in the tropical atmosphere: new opportunities, old challenges

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Stephan, C. C., Zagar, N. and Shepherd, T. G. ORCID: <https://orcid.org/0000-0002-6631-9968> (2021) Waves and coherent flows in the tropical atmosphere: new opportunities, old challenges. Quarterly Journal of the Royal Meteorological Society, 147 (738). pp. 2597-2624. ISSN 1477-870X doi: <https://doi.org/10.1002/qj.4109> Available at <https://centaur.reading.ac.uk/98662/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1002/qj.4109>

Publisher: Royal Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur


CentAUR

Central Archive at the University of Reading

Reading's research outputs online

REVIEW ARTICLE

Waves and coherent flows in the tropical atmosphere: New opportunities, old challenges

Claudia C. Stephan¹  | Nedjeljka Žagar² | Theodore G. Shepherd³¹Max Planck Institute for Meteorology, Hamburg, Germany²Meteorological Institute, Center for Earth System Research and Sustainability, Universität Hamburg, Hamburg, Germany³Department of Meteorology, University of Reading, Reading, UK**Correspondence**C. C. Stephan, Max Planck Institute for Meteorology, Bundesstrasse 53, Hamburg 20146, Germany
Email: claudia.stephan@mpimet.mpg.de**Funding information**

Minerva Fast Track Program of the Max Planck Society

Abstract

Atmospheric science relies on numerical models to simulate the complex multiscale nature of atmospheric variability, but our confidence in weather and climate predictions relies on theory and simplified models that describe scale interactions at a mechanistic level and can provide causal accounts of atmospheric behaviour. Global simulations at kilometre-scale resolution are now feasible and offer new opportunities to the atmospheric science community for testing and expanding our understanding of climate variability and change. Taking full advantage of this new tool requires smart strategies for evaluating and analysing the output, especially as kilometre-scale climate modelling will be limited to relatively short simulations with a rather small number of realizations. We here review some of the available tools for diagnosing and studying the dynamics of waves, coherent flows, and the interactions between them in terms of their ability to provide causal accounts of the behaviour seen in observations and in comprehensive simulation models. We describe their successes but also some of their limitations. The limitations are seen to be especially pronounced in the Tropics, where clouds, convection and atmospheric circulation are inextricably linked. The lack of a natural spatial truncation scale in the Tropics has given rise to many theoretical challenges, but it is for precisely this reason that the Tropics are where we might expect the largest gain from global kilometre-scale models.

KEYWORDS

cause and effect, equatorial waves, gravity waves, mean flows, normal-mode decomposition, scale interactions, tropical dynamics

1 | INTRODUCTION

The atmosphere exhibits a rich spectrum of variability in both space and time. Our knowledge of this variability comes first and foremost from observations. However,

observations of the atmosphere provide an incomplete representation of the relevant fields, and the historical record only samples a subset – potentially a very small subset – of what might have been possible. To make matters worse, the historical record is affected by vari-

ous sources of non-stationarity, some of which may be unknown, meaning that aggregation in time to construct statistics can sometimes be problematical. Comprehensive simulation models based on the fundamental governing equations can overcome these sampling limitations and fill in the phase space of possible behaviour. But models have systematic errors and biases (many of which are poorly known), so the reliability of this infilling is not obvious. Moreover, since the real world is non-stationary, we are actually interested in how the non-stationarity plays out, on a variety of time-scales ranging from probabilistic weather prediction to climate change. As the non-stationarity never exactly repeats itself – and, for climate change, lies in the future – the uncertainty in the effect of non-stationarity is epistemic rather than aleatoric (Beven, 2016), and cannot be represented using standard aleatoric measures (Shepherd, 2019). This requires thinking causally, in terms of theory and mechanistic processes.

Comprehensive global simulations at kilometre-scale resolution offer new opportunities to the atmospheric science community, as their fine resolutions permit an explicit simulation of important small-scale processes, such as convection (Satoh *et al.*, 2019). With fewer processes requiring parametrization, one may hope that model biases reduce. Much of the excitement about global kilometre-scale models derives from the idea that the newly resolved small scales are of immediate relevance for a realistic simulation of larger-scale processes (Schär *et al.*, 2019). These simulations can serve as a tool to test and expand our understanding of climate variability and change. Yet, it is clear that this will require careful strategies for evaluating and analysing the output, especially as kilometre-scale climate modelling will be limited to relatively short simulations with a rather small number of realizations. In order to be able to identify biases, untangle scale interactions and quantify their relevance for weather and climate on different scales from such limited samples of model behaviour, it is necessary to bring theoretical understanding.

The complex, multiscale nature of atmospheric variability, in both space and time, has always presented a challenge for theoretical understanding. Since the fundamental governing equations support both waves and turbulence, and the observations show manifestations of both, Michael McIntyre (2002) evocatively described the situation as a ‘wave-turbulence jigsaw puzzle’. Classical approaches to this challenge have sought to identify specific dynamical regimes based on analysis of the governing equations under particular simplifying assumptions. These provide the basis for conceptual models of atmospheric variability and, ultimately, for causal accounts of its behaviour. A unifying theme that runs through the

theoretical literature is the explanatory power provided by the concepts of waves, coherent flows, and the interactions between them. We use the term ‘coherent flows’ rather than ‘mean flows’, as the former encompasses the latter but is somewhat broader. We here review some of the ways those concepts have been used to understand atmospheric behaviour, based on atmospheric reanalysis data, numerical models, and observations. In this we bring together three strands of knowledge – large-scale dynamics, normal modes, and tropical convection and gravity waves (GWs) – which have historically represented rather distinct communities of researchers, with an eye to their application to the emerging first generation of kilometre-scale global atmospheric models, and with a particular focus on the Tropics. Each strand has been the subject of separate and more comprehensive reviews in the past. However, we see value in bringing the topics together, as a way of encouraging cross-fertilization of ideas across the three communities, especially for early-career researchers who may still be defining the contours of their research ambitions.

In Section 2 we present a general overview of some of the key equation systems that are relevant to different dynamical regimes and of classical approaches to distinguish between waves and coherent flows in these regimes. Starting from the equations for horizontal oscillations on the sphere – the shallow-water equations (SWEs; Section 2.1) – we discuss the concept of ‘balanced’ dynamics which distinguishes between ‘fast’ waves and ‘slow’ coherent flows (Section 2.2), the non-separability of fast and slow large-scale dynamics that arises in the Tropics (even within linear theory) from equatorially trapped modes (Section 2.3), internal gravity waves (Section 2.4), and slow large-scale dynamics in the Tropics (Section 2.5). Whilst we give many examples demonstrating how these concepts are relevant to the real atmosphere, the perspective is very much from the direction of theory.

In Section 3 we invert the perspective, starting from the observed behaviour of the atmosphere, especially as it pertains to interactions across scales (both space and time). We start with the extension of spherical normal-mode theory to the three-dimensional atmosphere (Section 3.1), and discuss wave solutions of the linearized primitive equations on the sphere (Section 3.2), an approach which has considerable explanatory power in the Tropics where the winds are much weaker than in the Extratropics. The focus is on large-scale equatorial Kelvin and mixed Rossby-gravity (MRG) waves, which explain a substantial part of the intraseasonal variance and yet represent a significant source of error for comprehensive simulation models. We then move to tropical mesoscales (Sections 3.3–3.5), where there is essentially no separation of time or space scales between GWs and coherent flows, and where clouds, convection and atmospheric

circulation are inextricably linked. The direct simulation of these scales is also particularly challenging, because the spectrum of vertical motion is essentially white, meaning there is no natural truncation scale and the properties of the solution can be expected to depend sensitively on the wavenumber cut-off. Yet the processes occurring on these scales have first-order effects on larger scales and on mean aspects of climate, and thus represent a major source of uncertainty – not only in models, but in theoretical explanations of atmospheric behaviour. Finally, we bring in the question of cause and effect (Section 3.6), which is always implicit in any physical analysis of observed atmospheric behaviour.

Section 4 provides some thoughts about future directions and opportunities, structured according to three cross-cutting themes that emerge in our review: building causal interpretations (Section 4.1), extending the concept of teleconnection patterns to model biases (Section 4.2), and a better understanding of upscale links (Section 4.3). A brief conclusion is offered in Section 5.

2 | DYNAMICAL REGIMES FROM PLANETARY TO MESOSCALES

2.1 | Horizontal motions on the sphere

Basic insight into the building blocks of atmospheric dynamics can be obtained from the nonlinear rotating SWEs. Their non-dimensional form in spherical coordinates can be written as follows:

$$\frac{\partial u}{\partial t} - v \sin \varphi + \frac{\gamma}{\cos \varphi} \frac{\partial h}{\partial \lambda} = -\gamma (\mathbf{V} \cdot \nabla u - uv \tan \varphi), \quad (1)$$

$$\frac{\partial v}{\partial t} + u \sin \varphi + \gamma \frac{\partial h}{\partial \varphi} = -\gamma (\mathbf{V} \cdot \nabla v + u^2 \tan \varphi), \quad (2)$$

$$\frac{\partial h}{\partial t} + \gamma \nabla \cdot \mathbf{V} = -\gamma (\mathbf{V} \cdot \nabla h + h \nabla \cdot \mathbf{V}), \quad (3)$$

where $\mathbf{V} = (u, v)$ is the non-dimensional horizontal velocity, t is non-dimensional time and h is the non-dimensional perturbed fluid height about its mean state D . The non-dimensional variables t , h , u and v are obtained by replacing dimensional time by $t/2\Omega$, where Ω is the rotation rate, dimensional perturbed height by Dh , and dimensional velocities by $\sqrt{gD}(u, v)$, where \sqrt{gD} is the phase speed of surface GWs in the absence of rotation and g is Earth's gravity. The non-dimensional 'del', $\nabla = \mathbf{i}_\lambda (\cos \varphi)^{-1} \partial / \partial \lambda + \mathbf{i}_\varphi \partial / \partial \varphi$, defines the horizontal spatial derivatives in spherical coordinates, with latitude and longitude denoted φ and λ , respectively. The unit vectors pointing towards the east and north are denoted \mathbf{i}_λ and \mathbf{i}_φ ,

respectively. The nonlinear terms on the right-hand sides of (1)–(3) will be referred to as N_u , N_v and N_h for the zonal velocity, meridional velocity and continuity equations, respectively.

The parameter γ is the inverse of the square root of the Lamb's parameter,

$$\gamma = \frac{\sqrt{gD}}{2\Omega a},$$

with a being the Earth's radius. The Lamb's parameter (γ^{-2}) is a single non-dimensional constant which partly characterizes the nature of shallow-water flows on the sphere, and is the eigenvalue of (1)–(3) when the right-hand side is zero (i.e., the Laplace tidal equation without forcing) (Longuet-Higgins, 1968). On Earth, γ is a small parameter with the Lamb's parameter γ^{-2} ranging from about 8 for D around 10 km to several 100 for D below 1 km.

A wave solution of the system (1)–(3) (with $N_u = N_v = N_h = 0$) is defined in terms of the eastward and westward propagating waves with latitudinal structure given by the Hough functions $\Theta_{n,r}^k$,

$$(u_{n,r}^k, v_{n,r}^k, h_{n,r}^k)^T = W_{n,r}^k \Theta_{n,r}^k(\varphi) e^{i(k\lambda - \omega_{n,r}^k t)}, \quad (4)$$

where $W_{n,r}^k$ is the expansion coefficient common to all three variables, and $r = 1, 2, 3$ indicates the three types of wave solutions: eastward-propagating inertia-gravity (EIG), westward-propagating inertia-gravity (WIG), and Rossby waves. For every wave type, the associated frequency ω depends on D , the zonal wavenumber k and the meridional index n . The horizontal structures of geopotential height, zonal wind and meridional wind, that is, the Hough harmonics $\mathbf{H}_{n,r}^k(\lambda, \varphi) = \Theta_{n,r}^k(\varphi) e^{ik\lambda}$, are a well-known result of a classical theory that describes small-amplitude perturbations about a motionless mean state in the stratified atmosphere (Longuet-Higgins, 1968; Andrews *et al.*, 1987; Kasahara, 2020).

Often the treatment of the SWEs is simplified by adopting a planar geometry for the study of dynamics restricted to a limited range of latitudes. Here there is an essential difference depending on whether the application is extratropical or tropical. For extratropical dynamics, where the Coriolis parameter is bounded away from zero, the frequencies of the eastward- and westward-propagating inertia-gravity wave (IGW) solutions are well separated from the frequencies of the Rossby wave solutions, as discussed in the next section. On the sphere or on the equatorial β -plane, however, the notion of separate fast and slow dynamics becomes problematic. Two special wave solutions, which owe their existence to the change of sign of φ at the Equator, fill the gap (Figure 1). These are the equatorial Kelvin wave and the MRG wave. The

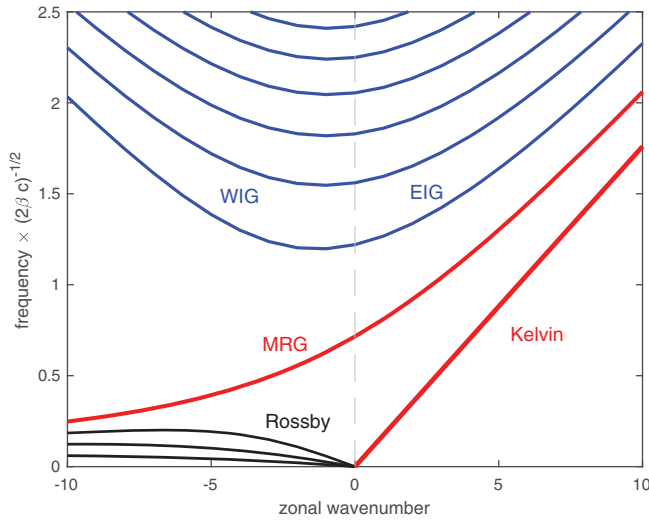


FIGURE 1 Frequencies of linear wave solutions on the sphere for $D = 300$ m. Frequencies are computed solving the eigenvalue problem following Swartrauber and Kasahara (1985) and normalized by $\sqrt{2\beta c}$, where $c = \sqrt{gD}$. Discrete values of frequencies are connected by continuous lines

former is the slowest among the eastward-propagating modes whereas the latter is the fastest low-frequency, westward-propagating mode. Since the 1960s, when the two modes were derived analytically (Matsuno, 1966) and observed both in the oceans and in the atmosphere, they have been the subject of intense research, especially the Kelvin wave. Its phase speed on the equatorial β -plane is equal to that of the surface GWs in the absence of rotation, \sqrt{gD} , a convenient scale for other waves, used in the non-dimensionalization of (1)–(3).

2.2 | Fast and slow dynamics

In geophysical fluid dynamics, the distinction between waves and coherent flows arises naturally in the small-amplitude limit about a state of rest where the characteristic velocity magnitude U tends to zero with all else held fixed. In that limit, waves retain a fixed frequency but the Doppler-shift frequency induced by advection, $U\kappa$, where κ is a wavenumber, tends to zero. There is then an asymptotic time-scale separation between (fast) waves and (slow) advective motion, represented by a small parameter. We identify the advective motion with the coherent flow.

We illustrate how this works for the SWEs on an f -plane. As is well known, the dispersion relation about a state of rest has three roots for the frequency ω as a function of total horizontal wavenumber κ :

$$\omega_0 = 0, \quad \omega_{IG} = \pm \sqrt{f_0^2 + gD\kappa^2}, \quad (5)$$

where f_0 is the Coriolis parameter. The first root represents the normal mode associated with the time evolution of disturbance potential vorticity, which on the f -plane is purely advective (i.e., there are no Rossby waves) and hence has zero frequency in the small-amplitude limit, whilst the other two roots represent (fast) IGWs. We can heuristically represent the effect of coherent flows on these frequencies through a Doppler shift, whence their ratio (assuming all quantities positive for convenience, and also assuming isotropy) takes the approximate magnitude

$$\frac{\omega_0}{\omega_{IG}} \approx \frac{U\kappa}{U\kappa + \sqrt{f_0^2 + gD\kappa^2}}. \quad (6)$$

It is easy to see that this ratio is asymptotically small, that is, there is a time-scale separation, if and only if

$$\epsilon \equiv \frac{U\kappa}{\sqrt{f_0^2 + gD\kappa^2}} \ll 1. \quad (7)$$

Defining the Rossby number Ro and the Froude number Fr in the usual way as $Ro = U\kappa/f_0$, $Fr = U/\sqrt{gD}$, our small parameter can be written instead (Saujani and Shepherd 2006) as

$$\epsilon = \frac{Ro \cdot Fr}{\sqrt{Ro^2 + Fr^2}}. \quad (8)$$

We see that $\epsilon \ll 1$ if either $Ro \ll 1$ or $Fr \ll 1$, even if the other parameter diverges. Thus, for example, there is even a time-scale separation at the Equator if $Fr \ll 1$, even though $Ro \rightarrow \infty$. A similar analysis applies to stratified flow (Saujani and Shepherd, 2006). The existence of a time-scale separation provides a mathematical basis for separating waves from coherent flows in model simulations, either through the time-stepping method itself (e.g., the use of implicit methods) or by applying a time filter to the output.

Following Warn *et al.* (1995), we can represent systems with such a time-scale separation in the general form

$$\frac{\partial f}{\partial \tau} + \Gamma f = \epsilon F(s, f; \epsilon), \quad \frac{\partial s}{\partial \tau} = \epsilon S(s, f; \epsilon), \quad (9)$$

where f represents the fast variables (not to be confused with the Coriolis parameter), s represents the slow variable(s), Γ is an invertible operator, the right-hand sides of the equations contain the nonlinear terms (such as advection), and τ is the fast time-scale (as in the non-dimensionalization employed to derive (1)–(3)). Note that the parameter γ in the previous section can be written as $\gamma = Ro/Fr$, if the Rossby number is defined as $Ro = U/(2\Omega a)$. In this case, γ becomes the Burger number.

In the limit $\epsilon \rightarrow 0$, the system linearizes with the eigenvalues of Γ providing the frequencies of the fast modes and with a zero frequency for the slow mode. Thus, the normal modes of the linear system provide a natural basis for the representation above, as they cleanly separate the fast and slow degrees of freedom. However, one can also seek solutions on the slow time-scale $t = \epsilon\tau$, which leads to

$$\frac{\partial f}{\partial t} + \frac{\Gamma f}{\epsilon} = F(s, f; \epsilon), \quad \frac{\partial s}{\partial t} = S(s, f; \epsilon). \quad (10)$$

Again considering the limit $\epsilon \rightarrow 0$, the first equation now requires the fast mode to vanish, $f = 0$, whereby the second equation reduces to a closed evolution equation for the slow dynamics. (Note that the slow dynamics are non-linear, even though the amplitude is small in the sense of (7).) Warn *et al.* (1995) illustrate this for the case of the f -plane SWEs under classical quasi-geostrophic scaling with $Ro = Fr$, where the fast variables are divergence and geostrophic imbalance, and the slow variable is the disturbance potential vorticity. Since on the f -plane the disturbance potential vorticity is a materially conserved quantity, its time evolution is purely advective. It thus represents coherent flows, whilst the fast variables represent IGWs.

If the slow dynamics contain low-frequency waves, such as Rossby waves, which occurs if the f -plane is replaced by the β -plane, then there is potentially a regime of small-amplitude, linear slow dynamics. This regime is relevant to the Tropics (see later discussion). More generally, we can consider low-frequency waves to be included within the function S , together with the advective dynamics.

It is important to distinguish between fast/slow *variables*, and fast/slow *dynamics*. If one expands the fast variables in an asymptotic series $f = f^{(0)} + \epsilon f^{(1)} + \epsilon^2 f^{(2)} + \dots$, then $f^{(0)} = 0$ is just the leading-order approximation and the next-order approximation of the system is given by

$$f^{(1)} = \Gamma^{-1} F(s, 0; 0), \quad \frac{\partial s}{\partial t} = S(s, f^{(1)}; \epsilon). \quad (11)$$

We can then say that the fast variables are slaved to the slow variables (Warn *et al.*, 1995), in the sense that the former are uniquely determined by the latter. (The converse is not true, because one cannot determine s from f .) The slow dynamics thus include not only the slow variables, but also the slow, slaved component of the fast variables.

Such slow dynamics derived under a time-scale separation are generally referred to as balanced dynamics, with quasi-geostrophic balance being the prototypical example. The importance of the time-scale separation is somewhat obscured in the traditional derivations (Pedlosky, 1987), which proceed from the momentum and mass equations.

The latter leads to three constraints (non-divergence and geostrophic balance in both horizontal directions), which turn out to be redundant, and potential vorticity conservation is derived as if by magic. Yet, since potential vorticity conservation is a property of all rotating, stratified fluids (Hoskins *et al.*, 1985), the existence of an advective slow mode is guaranteed for any geophysical system in the limit $\epsilon \rightarrow 0$. An example of a slaving relation is the well-known quasi-geostrophic omega equation, which relates the vertical velocity (and thus the horizontal divergence, a fast variable) to the slow dynamics. We can refer to the fast dynamics as unbalanced dynamics, as it may not be entirely wave-like if the amplitudes are large enough. Thus, in general the fast variables will comprise both slow (slaved) and fast (unbalanced) dynamical components.

However, even in this f -plane context, the separation between fast and slow dynamics is only exact in the limit $\epsilon \rightarrow 0$. For small but finite ϵ , the separation is only asymptotic (Bokhove and Shepherd, 1996) and the dynamics become coupled if the slow dynamics are chaotic (Wirosietisno and Shepherd, 2000) – which is generally the case for vortical dynamics. Nevertheless, the degree of separation of the fast and slow dynamics in geophysically realistic regimes can be quite remarkable. Figure 2 shows an example of this from an idealized model derived from a low-resolution version of the SWEs, for which the small parameter in (8) is 0.04 and thus there is a clear separation of time-scales. The frequency power spectrum of the potential vorticity, corresponding to the balanced dynamics, is recognizable as a continuous spectrum falling off rapidly from its maximum at zero frequency. Such a feature is also evident in the power spectrum of divergence, although five orders of magnitude smaller, and corresponds to the slaved component of the fast variables, represented by (11a). The unbalanced dynamics are visible in the divergence spectrum as a discrete set of GWs (since the resolution of the model is very limited), clearly identifiable by their frequencies. In a more realistic setting, a similar sort of separation was identified in simulations of the SWEs on a sphere forced by a large-scale planetary wave (McIntyre and Norton, 2000).

A discrete spectrum of linear waves analogous to that in Figure 2 is visible in the real atmosphere in the manifestation of the trapped global normal modes, known as Lamb waves, seen in surface pressure variations (Sakazaki and Hamilton, 2020). More generally, however, atmospheric spectra are continuous. The global horizontal kinetic energy spatial wavenumber spectrum exhibits a κ^{-3} power law at synoptic scales (Boer and Shepherd, 1983), which is associated with the slow balanced dynamics, but a $\kappa^{-5/3}$ power law on the mesoscales (Nastrom and Gage, 1985). Similar behaviour is seen in atmospheric analyses and models (Figure 3; also Figure 5a below).

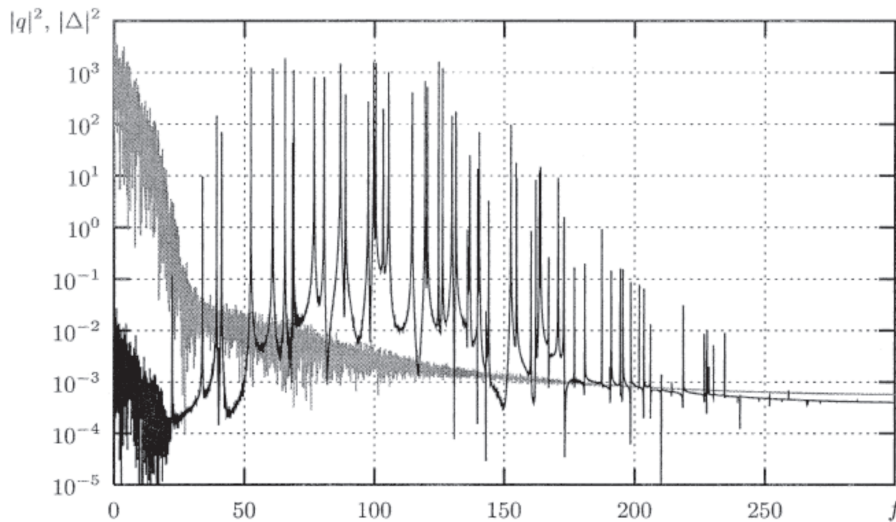


FIGURE 2 Frequency power spectra of potential vorticity (grey line) and divergence (black line) at a single grid point, in a simplified shallow-water model run at 17×17 wavenumbers. The minimum gravity wave frequency is 22.4. From Wirosietisno *et al.* (2002)

Note that, for global spectral analyses, the spherical harmonic index (or global wavenumber) is the analogue of the horizontal wavenumber κ in planar geometry. For local analyses on the mesoscale, as in the aircraft measurements used by Nastrom and Gage (1985), the spectrum is usually assumed to be horizontally isotropic (Figure 4, also Blažica *et al.*, 2013). In that case one can use any directional wavenumber as a proxy for κ , for example, the along-track wavenumber for aircraft measurements, or the zonal wavenumber (Figure 3b), recognizing that in a global spectrum the zonal wavenumber does not correspond to a unique spatial scale. However the spectrum is not isotropic at the large synoptic and planetary scales of the global atmosphere, as discussed in Section 3.1 and illustrated in Figure 3.

The origin of the $\kappa^{-5/3}$ power law spectrum has been the subject of much speculation over the years, but is now generally associated with the fast unbalanced dynamics, because the magnitude of the divergence is comparable to or even exceeds that of vorticity, in contrast to the κ^{-3} regime where the divergence is asymptotically much smaller than vorticity and thus associated with balanced dynamics (Figure 3; also Figure 5a). Both spectral regimes co-exist, and whether the κ^{-3} or the $\kappa^{-5/3}$ regime is dominant at a given scale is a dynamic property which depends on the nature of the flow. Using midlatitude aircraft measurements at tropopause height, Kunkel *et al.* (2019) show examples of this from different flights. Within the middle atmosphere, the balanced spectrum is rapidly attenuated with altitude by Charney–Drazin filtering (Charney and Drazin, 1961), so that the $\kappa^{-5/3}$ spectrum emerges even at fairly large scales (Koshyk and Hamilton, 2001; Burgess *et al.*, 2013). Based on atmospheric simulation models, the spectrum of vertical motion in the tropical troposphere appears to be very shallow, indeed almost white (Schumann, 2019), across a wide range of scales. This is suggestive of a very strong fast unbalanced component

to the flow, since a divergent horizontal kinetic energy power law of $\kappa^{-5/3}$ would imply a horizontal divergence variance power law of $\kappa^{-1/3}$ – that is, close to a (spatially) white spectrum – and horizontal divergence is closely related to vertical velocity. The dynamics of such a strong unbalanced spectrum can be expected to be highly non-linear, meaning that the coherent flows themselves will fall within this spectrum and would be removed by a simple frequency filtering – throwing the baby out with the bathwater.

2.3 | Fast and slow dynamics on the sphere

We now return to the spherical SWEs and explain how their linearized solutions relate to the concepts of fast and slow dynamics discussed above. An arbitrary solution of (1)–(3) can be expressed as

$$(u, v, h)^T = \sum_{k,n,r} W_{n,r}^k(t) \mathbf{H}_{n,r}^k, \quad (12)$$

which can be combined with (4) to transform the nonlinear equations (1)–(3) into the spectral ODE

$$\frac{\partial \mathbf{W}}{\partial t} + i\mathcal{L}\mathbf{W} = \mathbf{N}(\mathbf{W}). \quad (13)$$

Here, \mathbf{W} is a complex vector with the Hough expansion coefficients $W_{n,r}^k(t)$, $i\mathcal{L}$ is a skew-adjoint operator defined by frequencies $\omega_1, \dots, \omega_N$ of eigensolutions of (1)–(3) with the right-hand side equal to zero, $\mathcal{L} = \text{diag}(\omega_1, \dots, \omega_N)$, and \mathbf{N} is a quadratic function induced by nonlinearity of the system with elements $N_{n,k}^k$ defined by

$$N_{n,r}^k = \frac{1}{2\pi} \int_0^{2\pi} \int_{-1}^1 (N_u, N_v, N_h)^T \left(\sum_{k',n',r'} W_{n',r'}^{k'} \mathbf{H}_{n',r'}^{k'} \right)$$

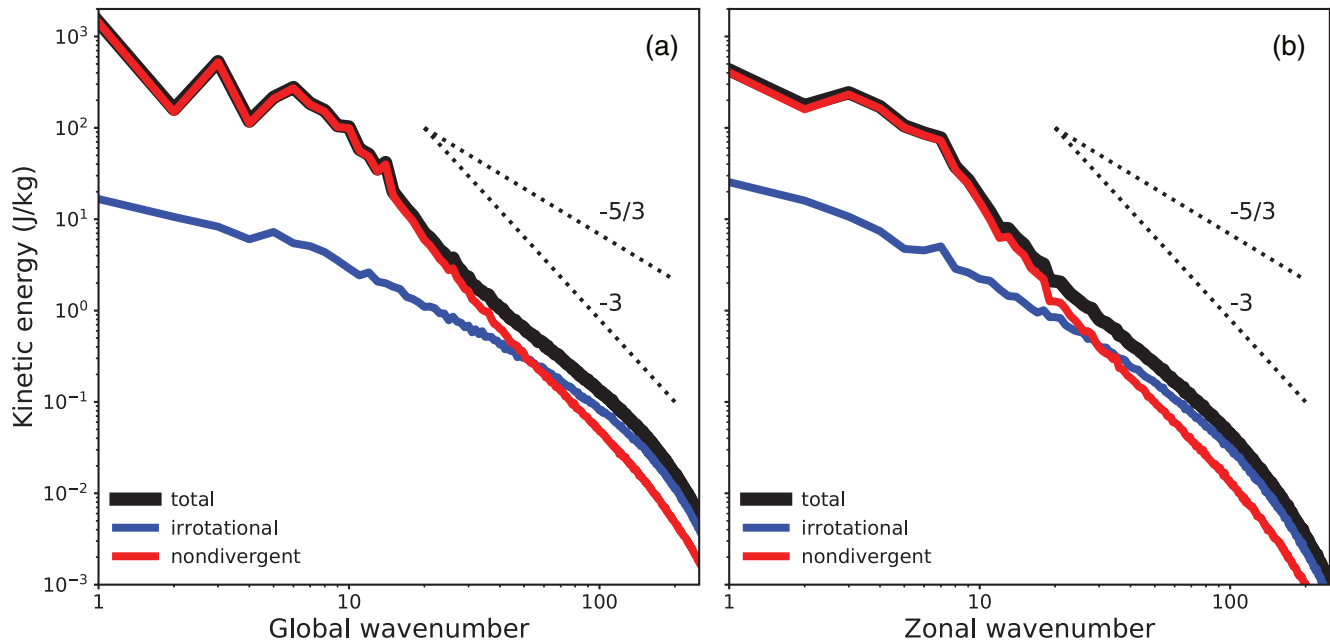


FIGURE 3 Global horizontal wavenumber spectra of horizontal kinetic energy at the 135 hPa level from the ERA5 reanalyses (black), with the spectrum decomposed into its rotational (red) and divergent (blue) components. (a) shows the spectrum as a function of the spherical harmonic index (or global wavenumber), which is the analogue of the horizontal wavenumber κ in planar geometry, and (b) as a function of the zonal wavenumber k . The spectra are averaged over one week of data once per day

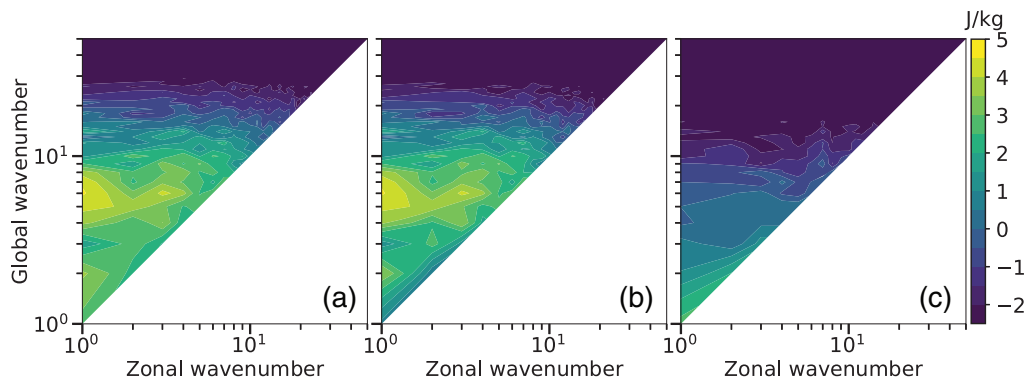


FIGURE 4 Global two-dimensional horizontal kinetic energy spectra at the 135 hPa level from the ERA5 reanalyses decomposed using spherical harmonics defined by the zonal wavenumber and spherical harmonic index (or global wavenumber). (a) Total horizontal kinetic energy, (b) rotational and (c) divergent components. The scale is logarithmic and uses the same data as Figure 3

$$\times (\mathbf{H}_{n,r}^k)^* (\lambda, \mu) d\mu d\lambda, \quad (14) \quad \text{and}$$

where $\mu = \sin(\varphi)$. For details of the derivation see, for example, Kasahara (1977).

To distinguish between the slow and fast dynamics, the state vector \mathbf{W} is split into a partition with the $W_{n,r}^k$ corresponding to the EIG and WIG modes, denoted \mathbf{W}_f , and a slow component made up of Rossby modes, denoted \mathbf{W}_s . Equation (13) is then partitioned into two equations for the IG and Rossby modes:

$$\frac{\partial \mathbf{W}_f}{\partial t} + i\mathcal{L}_f \mathbf{W}_f = \mathbf{N}_f(\mathbf{W}_f, \mathbf{W}_s), \quad (15)$$

$$\frac{\partial \mathbf{W}_s}{\partial t} + i\mathcal{L}_s \mathbf{W}_s = \mathbf{N}_s(\mathbf{W}_f, \mathbf{W}_s), \quad (16)$$

where \mathcal{L}_f and \mathcal{L}_s denote diagonal matrices with frequencies of the IG and Rossby modes, respectively, and \mathbf{N}_f and \mathbf{N}_s consist of elements $N_{n,r}^k$ projecting onto the IG and Rossby modes, respectively. The vectors in (15) have twice as many elements as those in (16). The difference to the fast-slow representation of (9) is the presence of the Kelvin and MRG waves and the non-zero frequencies of the Rossby waves. In the spherical normal-mode

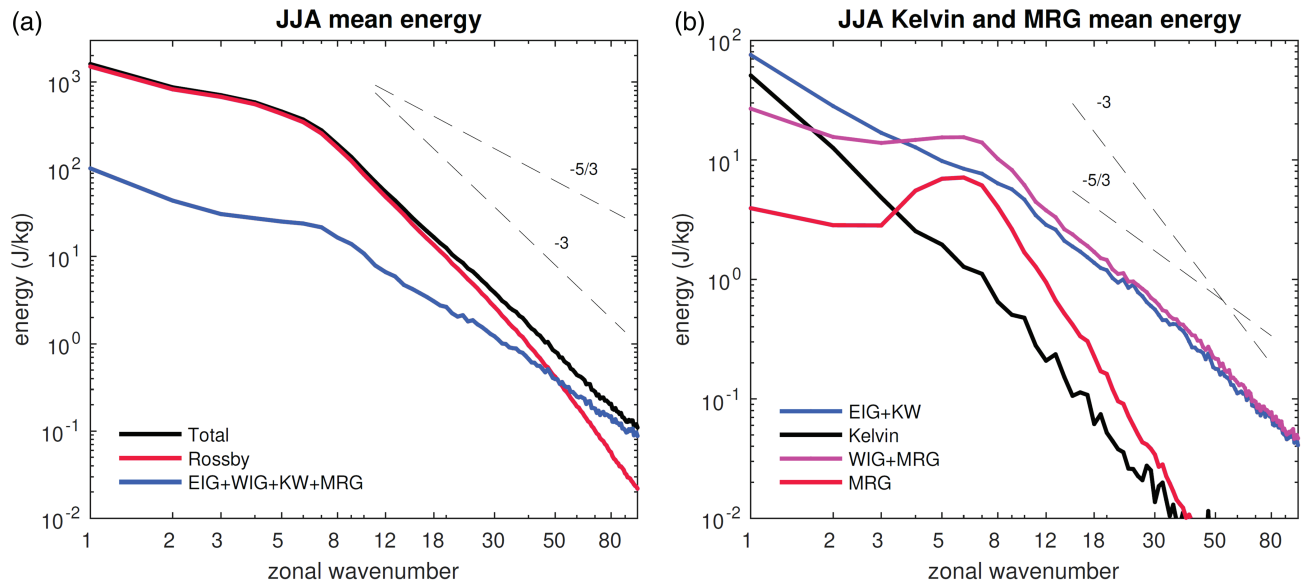


FIGURE 5 Energy spectra of the (a) Rossby and non-Rossby modes, and (b) Kelvin and MRG waves in comparison with the total eastward- and westward-propagating non-Rossby modes. Energy is averaged over a 35-year period and the boreal summer months (JJA). The analysis includes 36 levels from the surface up to 96 hPa in ERA-Interim

decomposition, the MRG wave appears as a solution of the dispersion relationship for slow modes, that is, it is accounted for in (16) whereas the Kelvin wave is a part of (15) and defined as the $n = 0$ EIG mode (Swarztrauber and Kasahara, 1985). The $n = 1$ EIG on the sphere corresponds to the so-called eastward-propagating MRG mode of Matsuno (1966) on the equatorial β -plane.

The problem of solving (15) for $\partial \mathbf{W}_f / \partial t$ in order to suppress unphysical high-frequency oscillations in numerical weather prediction (NWP) was extensively addressed in the late 1970s and 1980s as part of nonlinear normal mode initialisation (Baer and Tribbia, 1977; Machenhauer, 1977; Tribbia, 1979; Temperton and Williamson, 1981). This procedure is represented by (11a) earlier, and represents a slaving of the fast variables to the slow dynamics. The equatorially trapped IG waves with low frequencies coupled with convection posed a major challenge for initialisation based on such an *ad hoc* assumption of the separability of slow and fast modes. As a complex picture of the tropical atmosphere emerged in which the Rossby and IG waves coexist at many scales, along with simpler initialisation methods such as the digital filter initialisation, the nonlinear normal-mode initialisation ceased to be used.

The picture of non-separability of the Rossby and GW dynamics in the Tropics was complemented by the finding that a large part of the analysis and forecast uncertainties in global prediction models projects onto the non-Rossby modes (Žagar *et al.* 2005; 2013; 2015). For example, Žagar *et al.* (2013) found that nearly 50% of the short-range forecast errors in the ECMWF system projected onto the non-Rossby modes, the Kelvin wave in particular. The

growth of short-range forecast uncertainties was also most noticeable in the large-scale equatorial Kelvin waves. Furthermore, the under-dispersiveness of the ensemble prediction system on the medium range was found to be associated with an insufficient variance in the equatorial non-Rossby modes (Žagar *et al.*, 2015). For the purpose of NWP, Žagar *et al.* (2004, 2005) suggested that the mass–wind coupling of both the Rossby and equatorially trapped non-Rossby modes, primarily the Kelvin and MRG waves, be explicitly implemented in the background-error covariance models for data assimilation in a similar way as the linear balance equation provides the geostrophically balanced part of analysis increments in the Extratropics.

2.4 | Gravity waves in the troposphere

The background flow in the middle atmosphere is often slowly varying, so that there is a separation of timescales between the background flow and GWs (Section 2.2). It follows that linear theory is a good approximation for describing the spectral characteristics of GWs, their propagation, and their effects on the background atmosphere (Fritts and Alexander, 2003). In the troposphere and especially near the planetary boundary layer, the atmospheric flow is instead highly variable. The presence of convection and turbulence makes it difficult to extract GWs from the noisy background, whereas several methods of filtering and fitting have successfully been developed for isolating GWs in the middle atmosphere (e.g., Lehmann *et al.*, 2012; Wright *et al.*, 2017; Schoon and Zülicke, 2018).

Much of the description of GWs is based on the Taylor–Goldstein equation, which describes GWs in the 2D space of the horizontal and vertical dimension instead of the shallow-water system introduced in Section 2.1. The Taylor–Goldstein equation can be interpreted as Equation (9) or (15) without the nonlinear terms on their right hand sides, that is, in the case of (9) in the limit $\epsilon \rightarrow 0$. The Taylor–Goldstein equation is not specific to the Tropics, but the examples of wave-convection coupling discussed in this section focus on the Tropics, and on waves of shorter horizontal scales than those discussed in Section 2.3.

We will now solve the two-dimensional irrotational, frictionless, adiabatic Euler equations with the density $\rho = \rho_0 + \rho'$ under the assumption $|\rho'/\rho_0| \ll 1$, which is valid as long as the vertical scale of perturbations is small compared to the density scale height H_s . Compressibility of the background density is retained, $\rho_0 = \rho_{\text{sfc}} e^{-z/H_s}$. The equations are then

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial z} = -\frac{1}{\rho_0} \frac{\partial p}{\partial x}, \quad (17)$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + w \frac{\partial w}{\partial z} = -\frac{1}{\rho_0} \frac{\partial p}{\partial z} - \frac{\rho'}{\rho_0} g, \quad (18)$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \quad (19)$$

$$\frac{\partial \rho}{\partial t} + u \frac{\partial \rho}{\partial x} + w \frac{\partial \rho}{\partial z} = 0. \quad (20)$$

Equations (17) and (18) are the horizontal and vertical, respectively, momentum equations, (19) is the non-divergence condition, and (20) represents the conservation of the density anomaly.

Linearization about a z -dependent background state (U, ρ_0, p_0) in hydrostatic balance (with $w_0 = 0$) with wave solutions for the perturbations, $(u', w', \rho', p') = (\hat{u}(z), \hat{w}(z), \hat{\rho}(z), \hat{p}(z)) e^{i(kx - \omega t)}$, yields the polarization equations

$$i\omega \hat{u} - \hat{w} \frac{dU}{dz} = \frac{i}{\rho_0} k \hat{p}, \quad (21)$$

$$i\omega \hat{w} = \frac{1}{\rho_0} \frac{d\hat{p}}{dz} + \frac{\hat{\rho}}{\rho_0} g, \quad (22)$$

$$ik\hat{u} + \frac{d\hat{w}}{dz} = 0, \quad (23)$$

$$i\omega \hat{\rho} - \hat{w} \frac{\rho_0}{g} N^2 = 0, \quad (24)$$

where the intrinsic frequency

$$\hat{\omega} = \omega - Uk \quad (25)$$

and the Brunt–Väisälä frequency $N = \sqrt{-(g/\rho_0)(\partial \rho_0 / \partial z)}$. (Note however, that this expression for N is only valid under the Boussinesq assumption (local incompressibility) and using this formula for calculating N in the atmosphere would yield incorrect values. Instead, $N = \sqrt{(g/\theta)(\partial \theta / \partial z)}$ must be used, where θ is the potential temperature.) Defining the density-weighted vertical velocity $w^* = e^{z/(2H_s)} \hat{w}$, (21)–(24) can be solved for w^* to yield the Taylor–Goldstein equation

$$\frac{d^2 w^*}{dz^2} + \left[\frac{k^2 N^2}{\hat{\omega}^2} + \frac{k}{\hat{\omega}} \frac{d^2 U}{dz^2} - \frac{k}{\hat{\omega}} \frac{1}{H_s} \frac{dU}{dz} - \frac{1}{4H_s^2} - k^2 \right] w^* = 0, \quad (26)$$

$$\Leftrightarrow \frac{d^2 w^*}{dz^2} + \left[\frac{N^2}{(c - U)^2} + \frac{1}{c - U} \frac{d^2 U}{dz^2} - \frac{1}{H_s(c - U)} \frac{dU}{dz} - \frac{1}{4H_s^2} - k^2 \right] w^* = 0, \quad (27)$$

$$\Leftrightarrow \frac{d^2 w^*}{dz^2} + [\Lambda^2 - k^2] w^* = 0, \quad (28)$$

$$\Leftrightarrow \frac{d^2 w^*}{dz^2} + m^2 w^* = 0, \quad (29)$$

where $c = \omega/k$ is the ground-based phase speed (Nappo 2002, for example, gives more details on GW linear theory). Λ^2 , defined as the first three terms inside the bracket in (27), is also called the Scorer parameter. The general solution $w^* = A e^{imz} + B e^{-imz}$ implies that the fields associated with vertically propagating waves ($m^2 > 0$) vary sinusoidally in the vertical. This has important consequences for the spectrum of tropospheric GWs as discussed in Section 3.3.

2.5 | Slow dynamics in the Tropics

The perspective of fast and slow dynamics has always been problematical in the Tropics, because of the breakdown of quasi-geostrophic scaling. Charney (1963) proposed a low- Fr balance, in accordance with (8), but it is adiabatic and thus neglects the dominant balance in the thermodynamic equation, which is between vertical motion and diabatic heating (Held and Hoskins, 1985). The latter follows from the weakness of the Coriolis term in the Tropics, which means that significant

horizontal temperature gradients cannot be sustained. Imposing this dominant balance as an exact balance (i.e., neglecting the local time derivative of the geopotential height perturbation) in (3) leads to the ‘weak temperature gradient’ (WTG) approximation (Sobel *et al.*, 2001) when using the equatorial β -plane. However, the WTG approximation fails to capture equatorial Kelvin waves.

A widely employed approximation to study low-frequency equatorial waves is the long-wave approximation, which consists of neglecting the local time derivative of the meridional wind in (2) when using the equatorial β -plane (Gill, 1982; Heckley and Gill, 1984; Ogrosky and Stechmann, 2015). The MRG wave is filtered out along with all IG waves. The long-wave approximation can be considered as something like the equatorial equivalent of the quasi-geostrophic approximation as it eliminates the IG waves. It has been helpful in providing a basic picture of low-frequency phenomena including the stationary tropical circulation in response to diabatic heating (e.g., Gill, 1980). One strong reason for the usefulness of the long-wave approximation is that it retains the Kelvin wave, which is of central importance for large-scale tropical motions that are affected by rotation. The long-wave approximation thus captures variability associated with the Walker circulation, but it is not informative for the Hadley cell dynamics and cross-equatorial winds projecting onto MRG waves (Yang and Hoskins, 2017; Hoskins *et al.*, 2020; Hoskins and Yang, 2021).

Chan and Shepherd (2014) combined the long-wave approximation with the slaving approach described in Section 2.2 applied to the diabatically forced SWEs on the equatorial β -plane, using height h as a slaving variable and the anisotropy between meridional and zonal length-scales (L_y , L_x respectively) inherent to the long-wave approximation as the small parameter: $\epsilon = L_y/L_x$. (Potential vorticity does not work as a slaving variable in the Tropics if one wishes to include equatorial Kelvin waves, since the latter have zero potential vorticity. However, vorticity could be used rather than height.) This approach uses the building blocks of the long-wave approximation, but within a fully nonlinear context. Although ϵ is not explicitly a ratio of slow to fast frequencies, it can be interpreted as such when $Fr = 1$ (Chan and Shepherd, 2014). The leading-order system takes the form

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(u_0 h) + \frac{\partial}{\partial y}(v_1 h) = Q, \quad (30)$$

$$u_0 = -\frac{1}{y} \frac{\partial h}{\partial y}, \quad (31)$$

$$v_0 = 0, \quad (32)$$

$$v_1 = \frac{1}{h} \mathcal{L}_0^{-1} \left[y \frac{\partial h}{\partial x} + y u_0 \frac{\partial u_0}{\partial x} + \frac{\partial^2}{\partial x \partial y} (h u_0) - \frac{\partial Q}{\partial y} \right], \quad (33)$$

$$\mathcal{L}_0 \equiv -\frac{\partial^2}{\partial y^2} + \frac{y}{h} \left(y - \frac{\partial u_0}{\partial y} \right). \quad (34)$$

Note that \mathcal{L}_0 is an invertible operator provided the zonal flow $u_0(y)$ is inertially stable. This represents a closed, first-order system (provided the diabatic heating Q is specified somehow), hence it has effectively filtered out the fast waves and is the appropriate slow dynamics in the long-wave equatorial regime. The system reduces to the time-dependent Gill model (Heckley and Gill, 1984) in the small-amplitude limit, but has the merit of being fully nonlinear.

In contrast to the WTG model, the mass Equation (30) is a prognostic equation: the balance between horizontal divergence and diabatic heating is included, but it is not exact. This is what permits the Kelvin wave solution. Thus the model incorporates WTG scaling, but does not impose WTG balance. The balance is rather semi-geostrophic, with the zonal flow in geostrophic balance (31), and the ageostrophic meridional flow (33) being crucial to the dynamics. One appealing feature of this model is that the diagnostic equation for the winds is linear, unlike in WTG. Also – potentially important for applications in data assimilation – the winds can be diagnosed without knowledge of the diabatic heating, so long as the time dependence of the mass field is known. Because MRG waves do not satisfy (31), they are not represented in the system. However, the slow dynamics can nonetheless represent a ‘guiding centre’ about which the MRG oscillations take place (Chan and Shepherd, 2014).

The dynamical balances within this model provide a mathematical framework to relate diabatic heating to meridional convergence (with the diabatic heating potentially derived from the time dependence of the mass field), which could be a useful metric for comparison of the slow tropical dynamics – including phenomena such as the Madden–Julian Oscillation – in kilometre-scale models and in reanalysis, and their evaluation against observations.

3 | INTERPRETING OBSERVED BEHAVIOUR ACROSS SCALES

3.1 | Normal modes in the three-dimensional spherical atmosphere

The rotating SWEs (1)–(3) appear also in the normal-mode decomposition of the three-dimensional equations.

They result from manipulating the standard prognostic variables (u, v, T, p_s) of the hydrostatic primitive equations (Kasahara and Puri, 1981; Kasahara, 2020). The assumption of vertical- and horizontal-time separability of the linearized 3D equations leads to two eigenfunction problems linked through a constant with the dimension of length, the so-called equivalent depth, D . Positive values of D , obtained as a solution of the vertical structure equation for the statically stable, vertically bounded atmosphere (Cohn and Dee, 1989), provide coupling between vertical eigenstructures and the horizontal oscillations described by the linearized SWEs with the average depth D . In this way, the SWEs are used to represent also the internal modes of a continuously stratified atmosphere.

The complete set of three-dimensional normal modes has been used to quantify the response to heating perturbations associated with different vertical modes of the spherical atmosphere. A typical baroclinic structure of the tropical response to transient heating involves a spectrum of IG waves that strongly depends on the time-scale of heating, whereas the Rossby and Kelvin wave responses show little dependence on the heating time-scale (e.g., Geisler and Stevens, 1982; Kasahara, 1984; Salby and Garcia, 1987; Garcia and Salby, 1987). The predicted baroclinic vertical structure has been regularly identified in observations and reanalyses (e.g., Yang *et al.*, 2007; Adames and Wallace, 2014; Žagar and Franzke, 2015).

When heating perturbations occur in a realistic flow including moist feedbacks, the circulation response is still qualitatively similar to that in dry models with imposed temperature perturbations, although nonlinear effects enhance the amplitudes of the induced wave perturbations (Kosovelj *et al.*, 2019). The short-term response to equatorial stationary heating projects predominantly on the Kelvin wave. Its prominent presence in observations, prediction models and their simulated uncertainties, may be due to the horizontal structure of the heating sources resembling that of the Kelvin waves (i.e., maxima of temperature perturbations on the Equator) and the fact that it is the large-scale GW with a frequency closest to that of the Rossby waves (Figure 1).

Note that the frequencies in Figure 1 are for the equations linearized about a motionless mean state. In the applications of the normal-mode function decomposition, these frequencies are related solely with the derivation of the horizontal structures of the Hough harmonics, and not with the discussion of wave propagation properties (Žagar *et al.*, 2017). The wave frequencies differ from Figure 1 when the mean flow is taken into account in the linearization; the frequencies can become unstable, as well as continuous, except for a few of the lowest balanced modes (Kasahara, 1980). However, the structures of the Hough

functions are not significantly different if the linearization is performed around a non-zero mean zonal flow (Corrigendum of Kasahara, 1980; Salby, 1981; Ahlquist, 1982). The Hough harmonics constructed with reference to the basic state at rest are thus a suitable basis for the projection. Evidence for theoretically predicted modes from linear theory in atmospheric observations and reanalyses has a long history (e.g., Madden, 2019, and references therein).

In spherical geometry, the latitudinal variation of β defines the direction of wave activity propagation away from tropical vorticity sources, and thereby teleconnection pathways (Hoskins *et al.*, 1977; Hoskins and Karoly, 1981). The dynamics of thermally direct tropical circulations in response to heating perturbations was elucidated by a one-level vorticity equation model (Sardeshmukh and Hoskins, 1988) and by applying the linear approximation to the primitive equations on the equatorial β -plane (Gill, 1980; Heckley and Gill, 1984). An advantage of the linearized primitive equations over the vorticity equation is that the former supports both Rossby and IG waves and the two already introduced equatorially trapped waves – the Kelvin and the MRG wave.

A widely employed method to analyze the role of MRG and Kelvin waves in tropical wave variability is wavenumber–frequency filtering of single-level, single-variable data (e.g., outgoing long-wave radiation) within the tropical belt. Filtering relies on the Fourier series expansion in longitude and time, with the dispersion relations for linear waves on the β -plane overlaid on the wavenumber–frequency variance diagrams to connect the observed power spectra with the Kelvin, MRG and other equatorially trapped waves (e.g., Wheeler and Kiladis, 1999; Kiladis *et al.*, 2009). Spatial filtering using the wave eigenstructures is a vastly different approach; in addition to the Hough harmonics decomposition on the sphere (e.g., Žagar *et al.*, 2009; Castanheira and Marques, 2015; Blaauw and Žagar, 2018), the parabolic cylinder functions can be employed on the equatorial β -plane (e.g., Yang *et al.*, 2003, 2007, Tindall *et al.*, 2006). The procedure can be multivariate, meaning that the geopotential height and winds are analyzed simultaneously (as in (4)). The Hough harmonics decomposition, which is applied in the next subsection, involves many equivalent depths leading to the three-dimensional structure of wave perturbations. The decomposition is time independent, but filtered wave structures nevertheless show coherent propagation properties¹ even for IG waves associated with synoptic-scale processes as demonstrated by the comparison with the hodograph method (Žagar *et al.*, 2017).

¹Examples of real-time equatorial wave filtering can be seen at <http://modes.cen.uni-hamburg.de>.

With the Tropics included, the energy spectra of the large-scale circulation are considerably anisotropic, as illustrated in Figure 4, which is a two-dimensional version of Figure 3. It shows a peak rotational energy at the global wavenumbers 5–10 associated with synoptic eddies. In contrast, divergent energy is greatest at planetary zonal scales and is skewed towards larger zonal wavenumbers due to tropical non-Rossby modes, mainly the Kelvin and MRG waves (Žagar *et al.*, 2009). It is not possible to quantify their role in Figure 4 based on the spherical harmonics, as they are not predominantly vortical or divergent modes. For example, the MRG wave better fits within the Rossby regime at synoptic than it does at planetary scales (Figure 1). A decomposition (12) in terms of the Hough harmonics which isolates the Kelvin and MRG waves within the global framework is thus undertaken in the next section.

3.2 | Spatio-temporal variability of equatorial Kelvin and MRG waves

Under the ergodicity assumption (i.e., time average is equal to average over phase), coupling between spatial and temporal variability associated with different dynamical regimes and waves can be deduced by applying (12) after the vertical decomposition. For the total (horizontal kinetic plus available potential) energy associated with a vertical mode with equivalent depth D at time t , total energy is obtained as

$$I_{n,r}^k(t) = \frac{1}{2} gD \left| W_{n,r}^k(t) \right|^2, \quad (35)$$

where the square of the absolute value is $\left| W_{n,r}^k \right|^2 = W_{n,r}^k (W_{n,r}^k)^*$, and $*$ denotes the complex conjugate (e.g., Kasahara 2020). The time-averaged energy (spatial variance) over time period T spanned by N steps is denoted $\overline{I_{n,r}^k}$.

Figure 5 shows $\overline{I_{n,r}^k}$ for the various modes in ERA-Interim reanalysis data (Dee *et al.*, 2011), with the averaging performed for daily data over the 35 boreal summers in the period 1980–2014. Energy per unit mass is integrated over 36 model levels between the surface and about 100 hPa. Note that the top boundary condition for the 3D normal-mode function decomposition is still applied near the model top ($p \approx 1$ Pa). Including the data above ~ 100 hPa would significantly increase energy levels per unit mass at large scales in the Kelvin and MRG waves and would dilute the spectrum of their tropospheric variance which is of interest here. The impact of vertical data density on the spectra was discussed in Žagar *et al.* (2012, 2017).

In Figure 5a, the MRG waves are counted together with the Kelvin and other IG modes². The continuous total energy spectrum of the IG+KW+MRG modes in Figure 5a has a slope which follows a $k^{-5/3}$ power law for $k \gtrsim 10$, with the non-Rossby energy exceeding the Rossby mode energy at scales around 500 km ($k \approx 50$). Note that the crossing scale is smaller than in Figure 3b which applies to a single level (135 hPa) and to the horizontal kinetic energy in ERA5. Žagar *et al.* (2017) showed that energy spectra of IG modes in more recent, high-resolution analyses follow a $k^{-5/3}$ power law closer than in ERA-Interim data, which is consistent with a better representation of smaller scales and internal GWs in higher-resolution models. Along with the results of energy decomposition into the Rossby and gravity normal modes of the linearized system by other authors (e.g., Dewan, 1979; Kitamura and Matsuda, 2010; Terasaki *et al.*, 2011), this supports the picture of quasi-linear IG waves being associated with the regime transition from a -3 power law at synoptic scales to a $-5/3$ power law at subsynoptic scales, as discussed in Section 2.2.

As discussed above, the MRG and Kelvin waves pose a problem for the partition between fast and slow dynamics on the sphere. The divergence/vorticity decomposition (Figure 3) is not as elucidating here since the Kelvin and MRG wave kinetic energy is spread across the vortical and divergent parts of the spectra. Their less divergent nature is made evident by their steeper energy spectra compared to eastward-propagating and westward-propagating IG modes, with the MRG spectrum more steep than the Kelvin wave spectrum (Figure 5b). There are still significant differences among analysis datasets in the representation of these two special wave solutions (Žagar *et al.*, 2009; Podglajen *et al.*, 2014).

A major contrast between the Kelvin and MRG waves at large scales is seen in their maximal energies at the zonal wavenumbers $k = 1$ and $k = 5$ –7, respectively. Although the peak MRG activity at synoptic and smaller scales exceeds that of the Kelvin wave, the total Kelvin wave energy is almost twice as great. The energy portion of Kelvin waves in all eastward-propagating modes is about 30%, whereas the MRG waves make up around 20% of the energy of the westward-propagating non-Rossby modes. The Kelvin and MRG wave contributions are much smaller when viewed in the context of the total subseasonal variability which is dominated by variability in Rossby modes (Žagar *et al.*, 2020b), but are relevant for the formulation of simplified tropical models as in Section 2.5. Furthermore, the most energetic scales of the two special wave solutions

²Energy spectra in Žagar *et al.* (2017) and related papers count the MRG waves among the balanced modes.

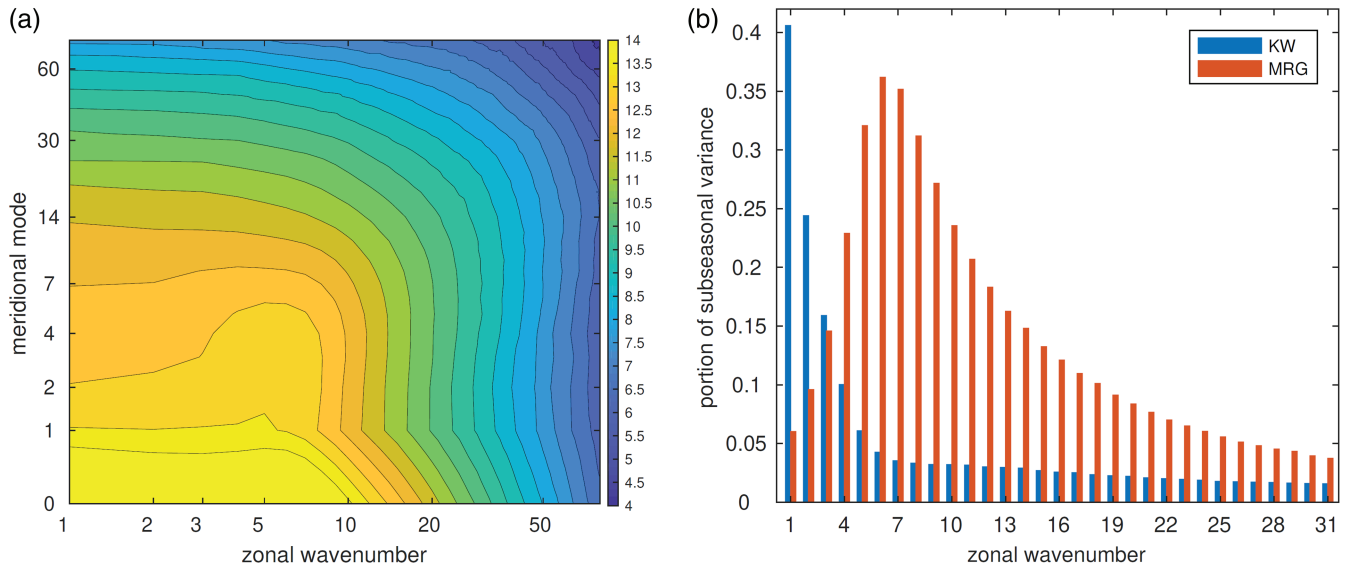


FIGURE 6 (a) Logarithm of subseasonal variance (in $\text{J}\cdot\text{kg}^{-1}$) of the global circulation projecting onto non-Rossby modes. (b) Ratio between the subseasonal variance associated with the Kelvin and mixed Rossby-gravity modes and the total global non-Rossby subseasonal variance in each zonal wavenumber, $(\text{KW}+\text{MRG})/(\text{EIG}+\text{WIG}+\text{KW}+\text{MRG})$. Variance is computed for ERA-Interim data, vertically integrated over 36 model levels from the surface up to ~ 96 hPa, and averaged over all seasons

are characterised by significant interseasonal dynamics, with the largest variance in boreal summer and winter for the Kelvin waves and MRG waves, respectively (not shown), in agreement with previous studies (Yang and Hoskins, 2017; Blaauw and Žagar, 2018).

Time averaging applied to a single mode $W_{n,r}^k$ is the time-mean state $\overline{W_{n,r}^k}$ and the associated energy is $E_{n,r}^k$:

$$E_{n,r}^k = \frac{1}{2}gD\left|\overline{W_{n,r}^k}\right|^2, \quad \text{where} \quad \overline{W_{n,r}^k} = \frac{1}{N}\sum_{t=1}^N W_{n,r}^k(t). \quad (36)$$

The difference between energy of the mean state, $E_{n,r}^k$, and the time mean energy $\overline{I_{n,r}^k}$, is equal to one half of the (biased) temporal variance $V_{n,r}^k$:

$$\overline{I_{n,r}^k} - E_{n,r}^k = \frac{1}{2}V_{n,r}^k, \quad (37)$$

where the biased variance is defined as

$$V_{n,r}^k = \frac{1}{N}\sum_{t=1}^N gD\left|W_{n,r}^k(t) - \overline{W_{n,r}^k}\right|^2. \quad (38)$$

In practice, we use the definition of the unbiased variance, that is, normalization is performed by using $N-1$ instead of N . The unit of energy and variance is $\text{J}\cdot\text{kg}^{-1}$ or $\text{m}^2\cdot\text{s}^{-2}$. In modal space, the horizontally integrated variance $V_{n,r}^k$ is equivalent to the integral in model space, after the vertical projection, of the variance in the wind components and weighted pseudo-geopotential height (Žagar *et al.*, 2020a).

The temporal variance, $V_{n,r}^k$, associated with the Kelvin and MRG modes at subseasonal scales is quantified in Figure 6. It shows the largest variance in the $n=0$ mode which is equatorially trapped and includes the Kelvin wave, the MRG wave and the $n=0$ WIG mode. The larger the zonal scale, the greater is the subseasonal variability, except at synoptic scales and $n=2-6$. A relatively greater variability in this range of zonal scales compared to neighbouring scales is associated with a component of synoptic-scale midlatitude dynamics projecting onto IG modes. We note that neither the annual cycle nor any other component of the signal present in the reanalysis data has been filtered. The variability distribution in ERA-Interim in the troposphere appears isotropic beyond zonal wavenumber $k \approx 50$, the scale where the non-Rossby energy exceeds the Rossby wave energy in Figure 5a.

The portion of variance due to the Kelvin and MRG waves in the total non-Rossby variance in each zonal wavenumber is quantified in Figure 6b. At zonal wavenumber 1, the Kelvin wave contributes a little over 40% of the subseasonal variance below ~ 100 hPa but its portion drops sharply with zonal wavenumber, and beyond wavenumber 5 makes up less than 5%. The MRG variance peaks at synoptic scales where it makes up about one-third of the subseasonal non-Rossby variability. At $k=15$ it is still over 10%. Overall, the two waves are associated with about one-third of the global subseasonal variance not associated with the Rossby-wave dynamics at planetary and synoptic scales (scales larger than 1,000 km). Both Kelvin and MRG waves must therefore be

ingredients of any theory that aims at describing tropical dynamics at large scales.

3.3 | Coupling of GWs to large-scale convective systems

Tropospheric GWs can force coherent or periodic patterns of vertical motion with visible effects on the cloud field. Unlike in the midlatitudes, where the Coriolis force limits the spatial influence of GWs, they can travel large distances in the Tropics. Here we review the mechanisms that have been proposed to describe the coupling between convection, GWs and the large-scale circulation on spatial scales ranging from a few hundred metres to hundreds of kilometres.

The heating or cooling associated with diabatic processes, for example condensation or evaporation, generates density gradients in the atmosphere. Therefore, a fundamental process associated with convective heating is the generation of internal GWs (Bretherton and Smolarkiewicz, 1989). Any vertical heating and cooling profile of depth D can be decomposed as $Q(z) = \sum_{n=1}^N A_n \sin(\pi n z/D)$, which projects onto vertical wavelengths

$$\lambda_z^n = \frac{2D}{n}, \quad (39)$$

(Alexander and Holton, 2004; Stephan *et al.*, 2016). If there is no background wind ($U = 0$) and the compressibility term $1/(4H_s^2)$ is neglected, then the Taylor–Goldstein equation (26) gives the dispersion relation

$$m^2 = k^2 \left(\frac{N^2}{\omega^2} - 1 \right). \quad (40)$$

Therefore, the horizontal phase speed of a wave with vertical wavelength λ_z^n is

$$c_p^n = \frac{\omega}{k} = \frac{1}{k} \frac{N}{\sqrt{1 + \frac{m^2}{k^2}}} \stackrel{m \gg k}{=} \frac{N}{m} = \frac{ND}{\pi n}. \quad (41)$$

For $N = 0.012 \text{ s}^{-1}$ and $D = 11 \text{ km}$, the resulting horizontal phase speeds are $42 \text{ m} \cdot \text{s}^{-1}$ ($n = 1$), $21 \text{ m} \cdot \text{s}^{-1}$ ($n = 2$) and $14 \text{ m} \cdot \text{s}^{-1}$ ($n = 3$). The gravest mode ($n = 1$) is associated with heating spanning the depth of the troposphere. To compensate for the local rising motion associated with this heating, the $n = 1$ mode initially induces deep compensating subsidence throughout the depth of the troposphere (Bretherton and Smolarkiewicz, 1989), which would act to suppress convection. A phase of deep ascending motion follows. This alternating response of

deep rising and sinking motion is similar to the response of the surface of water at rest when it is displaced *upwards* by an initial disturbance, that is, a disturbance opposite to the *downward* displacement a stone thrown into the water would cause. The second mode ($n = 2$) with a vertical wavelength equal to the depth of the troposphere corresponds to lower-tropospheric cooling associated with precipitation. This $n = 2$ mode induces compensating lower-tropospheric ascent, which would act to promote new convection. Shallow convection projects onto the $n = 3$ mode. The spectra of deep convection have been explored in dry simulations with prescribed sources of heating and cooling (Stephan and Alexander, 2015; Stephan *et al.*, 2016), showing that prescribed heating generates very realistic wave amplitudes and propagation characteristics which are consistent with linear theory.

GWs associated with $n = 1, 2, 3$ and higher modes can trigger a sequential initiation of convection, explaining how organized convection can propagate faster than storm outflow (Lac *et al.*, 2002; Tulich and Mapes, 2008). Moreover, Lane and Zhang (2011) argued that quasi-regular cloud systems can result from a quasi-resonant coupling between convection and GWs. They explained that the depth of convection determines the vertical wavelength, while the typical time-scale of convection determines the wave frequency. The horizontal wavelength is then constrained by the dispersion relation and in their case was 100 km, matching the cloud spacing.

The horizontal group velocity is

$$c_{gx} = \frac{\partial \omega}{\partial k} = \frac{Nm^2}{(k^2 + m^2)^{3/2}},$$

and the vertical group velocity is

$$c_{gz} = \frac{\partial \omega}{\partial m} = \frac{Nmk}{(k^2 + m^2)^{3/2}},$$

and thus

$$\frac{c_{gz}}{c_{gx}} = \frac{k}{m}.$$

This implies that the waves will eventually escape into the stratosphere, in particular the leading modes (small m ; (41)), if they are not partially or fully reflected. Vertical variations in wind speed or stability can cause partial or total reflection of upward propagating GWs, so that waves can become trapped. For these waves $m^2 < 0$, which implies that the wave amplitude decays exponentially outside of the trapping region, which is also called the wave duct. If there is a constant background wind U , then (40)

becomes

$$m^2 = k^2 \left(\frac{N^2}{\hat{\omega}^2} - 1 \right), \quad (42)$$

as is readily apparent from (27). One can see that $m^2 \rightarrow 0$ is approached as $\hat{\omega} \rightarrow N$. Since

$$\hat{\omega} = \frac{kN}{(k^2 + m^2)^{1/2}} = N \cos \alpha,$$

where α is the angle between the wave vector and the horizontal, trapped waves have vertically aligned phase lines, that is, the waves are moving only horizontally. In this case energy is not escaping into the stratosphere, but remains within the troposphere.

From (28) and (29), $\Lambda^2 - k^2 = m^2$. Therefore, waves can only propagate upwards if $\Lambda^2 - k^2 > 0$, or equivalently, if $\lambda_x = 2\pi/k$ is greater than the critical wavelength λ_x^* :

$$\lambda_x > \lambda_x^* = \frac{2\pi}{\Lambda}. \quad (43)$$

Conditions conducive to trapping are characterized by λ_x^* increasing with height, or equivalently, by the Scorer parameter decreasing with height. Trapping can occur, for instance, through a Doppler shift of the intrinsic frequency $\hat{\omega}$ when a wave propagates upstream, as this would increase $\hat{\omega}$ (for upstream-propagation $-kU > 0$ in (25)) and decrease the Scorer parameter, since $k^2 N^2 / \hat{\omega}^2$ is the leading term in (26).

Both shallow and deep trapped GWs can promote the growth of shallow convection into banded clouds (Shige and Satomura, 2001), and provide a mechanism for the longevity of convective systems (Stechmann and Majda, 2009). Ducted waves have also been observed to initiate severe convection (Su and Zhai, 2017). The $n = 3$ and higher modes have phase speeds and a vertical structure similar to cold pools. Despite the recognition that cold pools play an important role in the organization of midlatitude convective systems via the Rotunno–Klemp–Weisman theory (Rotunno *et al.*, 1988), Grant *et al.* (2018) showed that not cold pools, but GWs influenced the intensity, mesoscale structure, propagation, and lifetime of their simulated organized tropical oceanic convective systems. They pointed out that the $n = 3$ wave's half-vertical wavelength (~ 5 km) is similar to the altitude of the freezing level. The wave amplitude is reinforced by convective systems via latent cooling from melting and evaporation below the freezing level. Similar arguments were presented by Lane and Moncrieff (2015) and Moncrieff and Lane (2015) to explain the up- and downshear propagation of mesoscale convective systems.

3.4 | Coupling of GWs to the boundary layer

Not only can GWs be forced thermally by heating, they can also be generated mechanically when thermals are subject to wind shear. It is mainly through this so-called obstacle effect that dry and moist thermals in the convective boundary layer generate GWs, which then propagate inside the above-lying stable layer and can be associated with vertical motions spanning the depth of the troposphere (Clark *et al.*, 1986; Balaji and Clark, 1988; Balaji *et al.*, 1993). Unlike the larger-scale waves associated with mesoscale systems (e.g., Lane and Zhang, 2011), the properties of these high-frequency and short-wavelength waves are not so much determined by their source as by the characteristics of the free troposphere (Lane and Clark, 2002). The scale selection is affected by two processes. The first is the trapping mechanism, discussed above, which mainly affects upstream-propagating waves, for which $-kU$ is positive (25). If U varies with height inside the stable layer, then this can cause waves of short horizontal wavelengths to become trapped. The second mechanism is critical-level filtering, which occurs when $\hat{\omega} \rightarrow 0$. A critical level corresponds to a wave propagating downstream, as $\hat{\omega} = 0$ implies $c = U$. For $\hat{\omega}^2 \ll N^2$, the dispersion relation (42) can be approximated as $|m| = N/|\hat{c}|$, where the intrinsic phase speed $\hat{c} = c - U$. Therefore, when a wave approaches a critical level, $\lambda_z \rightarrow 0$. The fluid motions become more and more horizontal, produce turbulence and the wave breaks down before reaching the critical level (Booker and Bretherton, 1967).

Studies using dry or moist models in two or three dimensions, but idealized by using a spatially homogeneous background state, have reported a feedback of such trapped waves on the boundary-layer thermals, whereby the thermals are re-arranged to match the dominant wavelength of the overlying wave field (Clark *et al.*, 1986; Balaji and Clark, 1988; Balaji *et al.*, 1993; Lane and Clark, 2002). Such a feedback is physically plausible because the vertical velocity must be continuous at the interface of the boundary layer and the stably-stratified wave layer. Therefore, wave-induced periodic perturbations are also felt by processes within the boundary layer. Here, it is important to note that, even though the boundary layer does not support GWs, the ducted wave amplitude decays below the height of the boundary layer z_{BL} as $e^{-m(z_{BL} - z)}$. Since m is proportional to k (42), the e-folding scale in the vertical increases with decreasing k . In other words, trapped GWs of longer wavelengths will influence a deeper region than short wavelengths. Since the longest trapped wave is λ_x^* (43), linear theory would predict that this wavelength will become dominant. For $\hat{\omega} = N$ and $\lambda_x = 5$ –15 km, the e-folding scale would be 0.8–2.4 km.

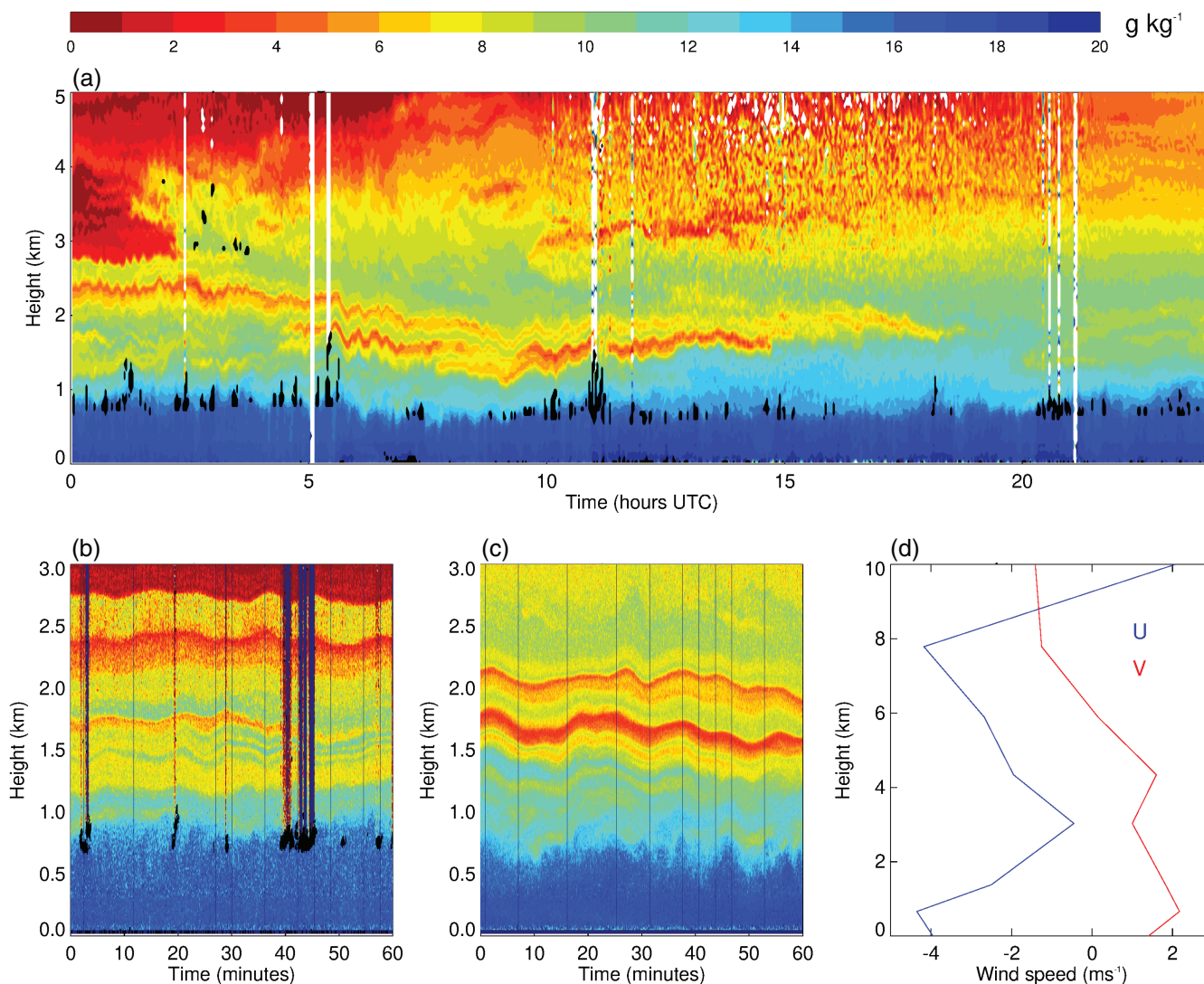


FIGURE 7 (a–c) Water vapour mixing ratio (colours) and cloud mask (black shading) from the CORAL (Cloud Observation with Radar And Lidar) instrument at the Barbados Cloud Observatory on 11 October 2019 (a) over 24 hr and (b) over 0000–0100 UTC and (c) over 0600–0700 UTC. (d) Daily-mean wind profiles on the same day at 300°E, 12.5°N from the NCEP reanalysis

To visualize wave effects inside the boundary layer, Figure 7a–c shows observations from the CORAL (Cloud Observation with Radar And Lidar) instrument at the Barbados Cloud Observatory (Stevens *et al.*, 2016). Since May 2019 CORAL measures high-resolution vertical profiles of water vapour using a 100 Hz 355 nm laser beam. Figure 7 shows a day that is representative of October 2019 in that widespread fields of shallow cumuli with horizontal diameters of up to about 1 km covered substantial fractions of the tropical Atlantic around Barbados. Reanalyzed wind profiles from NCEP (Kalnay *et al.*, 1996) imply advection speeds of about 4–5 m s^{−1} at an altitude just below 1 km (Figure 7d). The most important factors controlling these types of clouds are assumed to be tropospheric stability, surface wind speeds, free tropospheric humidity, and large-scale vertical motion (e.g.,

Stevens and Brenguier, 2009). But CORAL observations also show clear signatures of deep waves in the water vapour mixing ratios. These waves extend up to at least 3 km and they are present above clouds, for instance at 0000–0100 UTC (Figure 7b), but also at times when no clouds are detected, for instance at 0600–0700 UTC (Figure 7c). At 0600–0700 UTC the vertical coherence is particularly evident even down to an altitude of 500 m. These observations are consistent with a coupling between trapped waves and dry or moist boundary-layer thermals. Combining cloud and wind lidar observations could be a promising avenue to enhance our understanding of how background conditions, GWs and their sources in the boundary layer are linked.

Idealized simulations with homogeneous background conditions are a powerful tool to confirm the predictions of

linear theory. However, due to the absence of larger-scale forcings, the feedbacks between the boundary layer and the overlying wave field might be unrealistically enhanced. Recently, Stephan (2020) found that a coupling between clouds and trapped GWs can also be detected in kilometre-scale simulations of full complexity, that is, in numerical simulations similar to weather forecasts (Klocke *et al.*, 2017). In the analyzed simulations of 2.5 km horizontal resolution, feedbacks between trapped waves and the cloud field occurred in regions and at times of strong upper-tropospheric wind shear. These situations were characterized by a highly-organized scattered cloud field, with spectral power at horizontal wavelengths of 19–27 km dominating both the field of vertically integrated cloud water as well as the vertical velocity field throughout the troposphere. The study showed that clouds can be patterned by waves even in the presence of large-scale forcings. Important questions that were not addressed include whether or not such interactions affect the overall amount of cloudiness, favour deeper clouds, or influence cloud lifetime, as suggested by idealized studies (Balaji and Clark, 1988).

Another mechanism through which GWs can couple to the boundary layer is Jeffreys' drag instability mechanism. Jeffreys (1925) described how water flowing down an inclined channel can develop instabilities that result in progressive waves on its surface. The instabilities develop because the surface drag is proportional to the square of the horizontal wind speed. Thus, variations in wind speed create variations in drag, which, due to turbulent mixing, are felt by the entire layer of water. These in turn result in convergence or divergence of water, respectively, eventually forming visible waves. Under certain circumstances, which are governed by nonlinear equations, this system can achieve a resonant state. Chimonas (1993) applied this concept to the atmospheric boundary layer by substituting the gravitational acceleration on slanted terrain with the acceleration due to the horizontal pressure gradient force. Figure 8a illustrates this case. In the undisturbed state the pressure gradient force is balanced by surface drag (d) which, due to turbulent mixing, acts on all fluid elements inside the boundary layer. In the disturbed state, drag and layer depth vary in space and time. Stephan (2021) showed that this mechanism contributes to the formation of shallow-convective cloud lines perpendicular to the near-surface wind, which are often observed by satellites over the tropical oceans. An example of such cloud lines is shown in Figure 8b.

The governing equation system was derived by Chimonas (1993). The simplest model consists of two layers of incompressible fluids: a turbulent boundary layer of depth h with constant density ρ_0 and constant wind speed U , topped by a neutrally or stably stratified semi-infinite

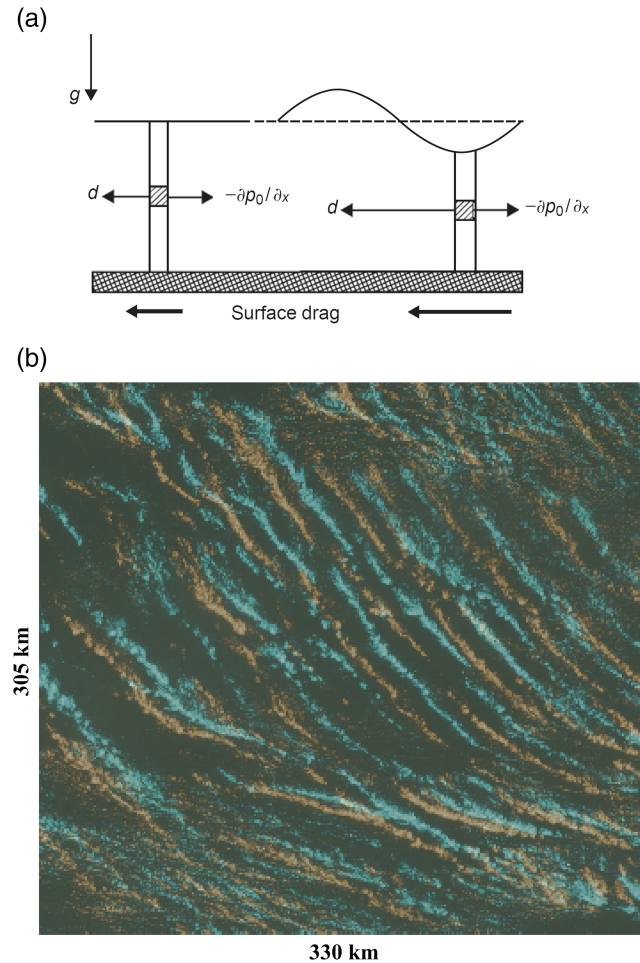


FIGURE 8 (a) Jeffreys' drag instability mechanism in the atmospheric boundary layer. Instabilities develop because variations in the surface drag d result in convergence or divergence (taken from Chimonas, 1993). Here, the symbol d corresponds to the term ρd in (44). (b) Two superimposed visible GOES-E images colorized by time (cyan: 1730 UTC and beige: 1800 UTC). The images show the area 51–48°W, 7.25–10°N on 5 April 2020

Boussinesq fluid with constant density ρ_1 and constant wind u_1 . In the boundary layer, the mass continuity equation in the long-wavelength limit is

$$\frac{\partial h}{\partial t} = -\frac{\partial(uh)}{\partial x}$$

(Jeffreys, 1925), and the horizontal momentum equation is

$$\rho \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} \right) u + \frac{\partial p}{\partial x} + \rho d = 0. \quad (44)$$

The key assumption is that for the constant background flow U with $d(U) = d_0$,

$$\frac{\partial p_0}{\partial x} + \rho d_0 = 0,$$

where the surface friction

$$\int_0^{h_0} \rho_0 d_0 \, dz = 0.5 \rho_0 C_D U^2.$$

Here, C_D is a constant drag coefficient and h_0 is the unperturbed height of the boundary layer. In the upper layer we have a harmonic wave field with the perturbation vertical velocity

$$w' = A(z) e^{i(kx - \omega t)}. \quad (45)$$

This equation system can be solved by linearizing and requiring that the field of pressure and height perturbations must be continuous at the interface of the two fluids. Then, in the long-wavelength approximation $\lambda_h \gg h_0$, the growth rate of the perturbation vertical velocity is

$$\tau_{\text{double}}^{-1} = \frac{C_D}{2} \frac{U}{h_0} \left(\frac{\frac{3}{2}U - c}{c - U} \right), \quad (46)$$

where c is the horizontal phase speed. Thus, perturbations moving downstream have positive growth rates for $U < c < 3U/2$. Stephan (2021) showed that the frequently observed cloud patterns with wavelengths of several tens of kilometres form over the course of about half a day when the mixed-layer depth (h_0) is about 1 km and the mean wind speed in this layer (U) is greater than 7 m s^{-1} . For slower wind speeds, the growth rates of the drag instability decrease significantly.

The fact that GWs can be about as slow, in terms of phase speed, as the coherent flows in their surrounding environment implies great difficulty in separating them. For Jeffreys' drag instability, the wave speed is close to the speed of the background flow. A second example, discussed earlier in Section 3.3, is that GWs and cold pools have nearly the same phase speed.

3.5 | Gravity waves and mesoscale divergence

So far, all mechanisms discussed in sections 3.3 and 3.4 have focused on discrete wave modes with horizontal wavelengths matching the scales of the respective organized convective systems. Recent findings suggest that mesoscale GWs may influence cloud patterns more generally, at least in convectively suppressed conditions, by modulating coherent vertical motion on horizontal scales of a few tens to several hundreds of kilometres. Weather and climate modelling often adopts the notion that a slowly-varying large-scale vertical motion, such as the sinking at the edge of a presumed Hadley cell, sets

the stage for convection. Yet, Stevens *et al.* (2020) showed that small and intermediate scales of motion substantially pattern cloud fields. Bony and Stevens (2019) observed strong long-lived wave-like variability in vertical profiles of mesoscale divergence. These profiles were computed from dropsonde measurements of horizontal winds over the tropical Atlantic. Stephan *et al.* (2020) confirmed such fine-scale structures in divergence profiles derived from the Tropical Warm Pool-International Cloud Experiment (TWP-ICE; May *et al.*, 2008).

In an attempt to explain these features, Stephan *et al.* (2020) analyzed radiosonde observations from TWP-ICE for the presence of low-frequency GWs, for which $\hat{\omega} \ll N$. GWs that may explain the divergence patterns need to have periods consistent with the temporal autocorrelation scales of the divergence profiles, horizontal wavelengths consistent with their spatial autocorrelation scales, and vertical wavelengths consistent with the vertical variability of the divergence profiles. The divergence $\text{Div}(x, y, z, t)$ is directly related to the wind perturbations associated with a plane wave with wave vector (k, l, m) , frequency ω and phase ϕ ,

$$(u, v, w) = (A_x, A_y, A_z) \sin(kx + ly + mz - \omega t + \phi), \quad (47)$$

as

$$\text{Div} \sim (kA_x + lA_y) \cos(kx + ly + mz - \omega t + \phi). \quad (48)$$

The typical characteristics of these low-frequency GWs can be extracted from radiosonde soundings using hodograph methods (e.g., Eckermann and Vincent, 1989). For a zonally propagating wave, the meridional velocity perturbation is $\hat{v} = -i(\hat{\omega}/f)\hat{u}$, that is, \hat{v} is 90° out of phase with \hat{u} . Therefore, $\hat{\omega}$ can be obtained from the polarization ellipse that is traced by the horizontal wind components. The vertical wavelength is found from a power spectral analysis and the horizontal wavelength follows from the dispersion relations. Stephan *et al.* (2020) found waves with typical vertical wavelengths of about 4 km, horizontal wavelengths of about 600 km and intrinsic periods of about 12 hr. The observed vertical, horizontal and temporal variability of divergence inside the area they considered are related, as one would expect from these GW properties.

The dispersion relation of low-frequency GWs can be approximated as

$$\hat{\omega}^2 = N^2 \frac{\kappa^2}{m^2} + f^2, \quad \text{where } \kappa^2 = k^2 + l^2,$$

$$\text{and thus } \left| \frac{c_{gz}}{c_{gh}} \right| = \frac{\kappa}{m} = \frac{\sqrt{\hat{\omega}^2 - f^2}}{N}.$$

Since for these waves $\hat{\omega}^2 \approx f^2$, they do not quickly escape into the stratosphere, but propagate over far distances within the troposphere.

The studies by Bony and Stevens (2019) and Stephan *et al.* (2020) sampled constant-sized areas with diameters of about 200 km, which did not permit studying how divergence characteristics depend on area size. The EUREC⁴A campaign of 2020, which took place over the western tropical Atlantic (Stevens *et al.*, 2021) included an extensive radiosounding network composed of a station on the island of Barbados and four ships (Stephan *et al.*, 2021). The polygon spanned by the island station and the ships varied in size, sampling areas with equivalent radii of 100 km up to about 400 km. Stephan and Mariaccia (2021) showed that observed divergence magnitudes during EUREC⁴A scaled approximately inversely with the area equivalent radius. Moreover, they confirmed this result in the ERA5 reanalysis and in a 2.5 km global numerical simulation. By applying the normal mode decomposition described in Section 3.1 to the numerical data, they derived an analytic expression that explains the scaling law of divergence magnitudes based on the wavenumber dependence of the global IG wave energy spectrum (the blue line of Figure 5a computed for their 2.5 km simulation). This finding clearly demonstrates that some properties of mesoscale circulation depend on all scales and that care must be taken when studying mesoscale processes in limited-area domains.

3.6 | Cause and effect

A common approach to develop conceptual models of atmospheric variability is to consider the steady form of a tendency equation derived from the underlying governing equations. Kilometre-scale models, for instance, can be used to study how the energy budget changes with cloud organization. Other examples of budgets include the zonal momentum budget and the moist static energy budget. The issue here is that, by definition, budgets must balance, and it is not immediately apparent which terms in the budget should be considered the cause, and which the effect. For example, in the context of climate change, a poleward shift of the Southern Hemisphere (SH) westerlies is a robust global climate model (GCM) response to global warming (Kushner *et al.*, 2001), but the mechanisms behind the shift remain unclear (Shaw, 2019). The proximate explanation of the jet shift is a shift in the regions of eddy momentum flux convergence, which can be explained by critical-layer control (Randel and Held, 1991), but this is merely a diagnostic balance and the causality of the relation remains unclear (Chen and Held, 2007).

The moist static energy budget is often used as a diagnostic framework for the zonally averaged tropical circulation (Neelin and Held, 1987). The framework has been used to explain the location and potential shifts in the ITCZ (e.g., Bischoff and Schneider, 2016). However, in their study of the response of the zonal-mean circulation to changes in surface roughness using a comprehensive GCM, Polichtchouk and Shepherd (2016) found that whilst the circulation response was thermally rather than dynamically driven (by changes in tropical surface sensible and latent heat fluxes), the moist static energy budget provided only a descriptive rather than a predictive framework in which to understand the circulation shifts.

Another common approach, applied both to observations and to the output of model simulations, is lagged correlations. This relies on the principle that causality introduces an arrow of time and thus that, if a perturbation in variable A occurs after a perturbation in variable B, then B is the cause of A. There are many pitfalls in such an approach, however (Runge *et al.*, 2014). An obvious one is that both perturbations might be caused by an unobserved common driver, with different lags in the responses of A and B. Another is the role of autocorrelation, since atmospheric variability generally has a red spectrum. For example, Lorenz and Hartmann (2001) performed cross-correlation analysis of the zonal momentum equation in the SH, and found not only the expected correlation with the momentum flux convergence leading the zonal wind acceleration, but also a weak oppositely lagged correlation which they interpreted as a positive eddy feedback. However, Byrne *et al.* (2016) showed that such a feature can be produced in a synthetic model of an eddy-driven mean flow with, by construction, no eddy feedback, through autocorrelation.

In the stratosphere, causality can sometimes be easier to untangle because the eddy forcing arises from waves propagating up from below, rather than from *in situ* instability, and thus can be considered exogenous to the stratosphere. For example, the robust strengthening of the Brewer–Dobson circulation under climate change found in GCMs is a direct consequence of the fact that tropospheric warming extends to higher altitudes in the Tropics than the Extratropics. By thermal-wind balance, this implies a strengthening of the upper flank of the subtropical westerlies, which (all things being equal) induces an upward shift in the critical layers of the baroclinic eddies, allowing more of the wave forcing to be exerted within the stratosphere (Shepherd and McLandress, 2011). The fact that the stratospheric response cannot nullify the first-order nature of tropospheric warming makes the causality in this case straightforward. On the other hand, it has been common to partition the wave forcing of the Brewer–Dobson circulation into its

component parts, with a particular focus on the role of IG waves whose representation in stratosphere-resolving GCMs is mainly parametrized and subject to considerable uncertainty (Butchart *et al.*, 2010). However, there is a very strong, practically buffering interaction between the wave forcing from Rossby waves and parametrized GWs in the midlatitude lower stratosphere (Sigmond and Shepherd, 2014), challenging the causality of such a partitioning.

In the Tropics, the dominant balance between diabatic heating and vertical motion naturally invites an interpretation of the diabatic heating driving the vertical motion. However, that is not as obvious as it might seem. From a transient perspective, the diabatic heating had to arise from some dynamical event or process, and thus depends on the past history of the circulation structure that induced it. From a steady perspective, the rising air must descend somewhere else, and there could be constraints on the horizontal motion or on the descent; and the diabatic heating requires a moisture supply, which depends on the atmospheric circulation. All this complicates understanding of cause and effect.

4 | OUTLOOK

In this article we have reviewed some aspects of the present understanding of waves and coherent flows in the atmosphere, and the interactions between them. The motivation for doing so is that most of the theoretical frameworks for understanding how the atmosphere works rest, in one way or another, on such concepts, which need to be applied in meaningful ways to observations and model output to take full advantage of the new generation of kilometre-scale global models. We examined classical approaches to this challenge, represented by conceptual models targeting specific dynamical regimes, and assessed their successes and limitations in providing causal accounts of the behaviour seen in observations and in comprehensive simulation models.

A common theme in our assessment is that the limitations are especially pronounced in the Tropics as quasi-geostrophic scaling ceases to apply. The equatorially trapped Kelvin wave and MRG wave, two special solutions provided by linear wave theory, obliterate the distinction between ‘fast’ and ‘slow’ dynamics which otherwise exists (at least asymptotically) in the Extratropics, and which underpins so much of extratropical dynamical theory. We presented new results which quantitatively estimate the role of the Kelvin and MRG waves in the spatio-temporal spectrum of tropical variability, clearly demonstrating the necessity of including both the Kelvin and MRG waves in simplified models of the Tropics.

On the tropical mesoscales, the lack of a clean separation of dynamical regimes becomes even more apparent in the complex interplay between convection, GWs, and coherent flows, where ‘slow’ trapped GWs can provide a large-scale constraint on ‘fast’ convective coherent flows. In this final section we provide an outlook on potential ways forward.

4.1 | Building causal interpretations

Pearl and Mackenzie (2018), in their discussion of Artificial Intelligence, identify three levels in their ‘Ladder of Causation’. The most primitive level is *Association*, which identifies relationships between variables (correlations). This assumes statistical stationarity and is not causal. Causality can be inferred from the second level, *Intervention*, but that is generally not an option in atmospheric science. Although information can be gleaned from particular interventions, such as volcanic eruptions, it is not straightforward to map it onto the other sources of non-stationarity in which we may be interested. The highest level is that of *Counterfactuals*, meaning an imagined outcome. By definition, counterfactuals cannot be observed, thus are unverifiable in a traditional sense. However they constitute the essential evidence that informs decision-making, and are anchored in understanding.

As has been discussed in Section 3.6, all too often what passes for an explanation in atmospheric science is no more than a description, without predictive power. Thus it would be helpful to be more explicit about causality in our scientific reasoning. For the real system, physical causality is probably not a well-defined concept for dynamical interactions since the system is strongly coupled, but from a statistical perspective causality can be defined as conditional dependence (e.g., Kretschmer *et al.*, 2016), that is, how one thing would be different if something else earlier in time were different, all else being equal. The rub here is that for the real system, it will never be the case that all else is equal, and we do not even know what ‘all else’ has to be. Nevertheless, causal networks provide a framework for articulating causal hypotheses, and testing them against data (Pearl and Mackenzie, 2018).

Although budgets have their interpretative limitations, they are anchored in the fundamental physical principles (conservation of energy, momentum, mass and moisture) that govern atmospheric dynamics. As a result, they consist of well-defined and in principle measurable quantities, which can be used to bridge between models and measurements, and within model hierarchies spanning different levels of complexity, including now the global kilometre-scale models. It would therefore be useful to better understand how different balances between terms can

manifest themselves in the steady limit, through study of the evolution from transient to steady-state balances. One way of doing this is through cross-spectral analysis, as has been used for the zonal momentum equation

$$\frac{\partial \bar{u}}{\partial t} = -\frac{\partial}{\partial y}(\overline{u'v'}) - r\bar{u} \quad (49)$$

in midlatitudes, applied to the barotropic component of the flow (Lorenz and Hartmann, 2001). Here the bar denotes a zonal average, the prime denotes deviations from the zonal average, and r is an effective damping rate. Taking the Fourier transform in time yields

$$\frac{Z^*M}{Z^*Z} = i\omega + r, \quad (50)$$

where Z is the transform of the zonal-mean wind, Z^* its complex conjugate, M the transform of the eddy momentum flux convergence, and ω the frequency. This relationship turns out to hold remarkably well across time-scales for describing the variability of the SH midlatitude jet, and allows a clear identification of transient, intermediate, and quasi-steady regimes (Boljka *et al.*, 2018). The intermediate regime (periods between about 10 and 50 days) corresponds to the subseasonal-to-seasonal (S2S) time-scales that are of much current interest in climate and weather prediction. Somewhat surprisingly, Boljka *et al.* (2018) also found that a similar relationship holds for the eddy kinetic energy equation, which can be considered a proxy for baroclinic wave activity, on intermediate to quasi-steady time-scales, though not for transient time-scales, despite the presence of additional terms that are ignored in such a simplified treatment. This shows that simplified versions of budgets may well hold to good approximation in practice, and help enable understanding of the transition to quasi-steady balances.

In this respect, the use of asymptotic models to define physically based causal networks with which to study particular phenomena would seem a fruitful avenue to pursue. Here the goal is not to solve the asymptotic equations to simulate the phenomena in question, which is the traditional use of asymptotic models. Rather, the goal is to use the asymptotic models to motivate the choice of dynamical variables and causal linkages with which to develop a causal network that can be applied to observations and comprehensive simulation models. This could be a powerful way of testing physical hypotheses. Indeed, this philosophy lies at the heart of the approach to tropical waves described in Sections 3.1–3.2, which uses linear theory (which is an asymptotic theory) to define the wave structure but observations or high-resolution models to determine the wave dynamics (evolution in time). In the context of the nonlinear asymptotic model described

in Section 2.5, for example, one might use (33) to motivate a spatial relationship between the diabatic heating and the meridional velocity, but use observations or high-resolution models to infer the quantitative nature of the relationship.

With respect to simulation models, causality probably is a well-defined concept when it comes to model biases. Certainly, one can perturb a parameter or a process in a model, and run the simulation sufficiently long to be sure that the difference is robust; the perturbation will then unambiguously be the cause of the difference. The problem is that this sort of causality is generally not very useful. Sensitivity to model parameters usually takes some time to manifest itself, undergoes feedbacks, and leads to changes in emergent properties of the solution, such as coherent flows, which can be difficult to interpret physically. The zonal mean circulation response to altered surface roughness discussed earlier (Polichtchouk and Shepherd, 2016) is an example of this. On the other hand, an abrupt switch-on of an exogenous forcing, such as increased CO₂, can sometimes be useful for separating chicken from egg (Chemke and Polvani, 2019). A common approach to overcome this issue is to intervene in the model directly, for example, through mechanism denial or nudging, but this violates physical self-consistency and, whilst useful for testing hypotheses such as the role of eddy forcing (O’Gorman and Schneider, 2007), can only be an intermediate step. The use of causal networks to formalize the sort of reasoning involved in such experimentation would help to make the logic more transparent and make clear how the results should be interpreted.

4.2 | Tropical dynamics and model biases

Linearized equations and asymptotic models have so far provided a gross understanding of tropical large-scale variability in response to heating, predominantly represented in terms of the Kelvin waves and Rossby waves with the lowest meridional mode $n = 1$. With a factor of three difference in their phase speeds and their opposite mass–wind coupling near the Equator (positive and negative for the Kelvin and $n = 1$ Rossby wave, respectively), the combined effect of diabatic forcing is a weak gradient in the mass field as expected due to the weakness of the Coriolis effect close to the Equator. The cross-equatorial flow is instead associated with the MRG mode and IG modes, even for the solstitial Hadley circulation (Hoskins *et al.*, 2020). The absence of the MRG wave from the Matsuno–Gill type models (or their nonlinear extensions as discussed in Section 2.5) is likely the most important missing building block in asymptotic tropical models, as evidenced by reanalysis data in Section 3.2.

The Kelvin and MRG waves generated by convective forcing are responsible for the loss of separability between the slow and fast dynamics close to the Equator. This makes wave-wave interactions involving these waves and their contribution to the tropical vorticity sources timely to study. Studies of resonant triads involving the equatorial Rossby, MRG and Kelvin waves are just beginning (Raupp and Dias, 2009). A model describing dynamics in terms of Rossby and IG eigensolutions of primitive linearized equations (i.e., Hough harmonics), as envisaged by Kasahara (1978), provides a suitable framework for evaluating tropical wave interactions on low-frequency (intraseasonal or longer) time-scales. A high-resolution, high-accuracy version of such a model also offers an attractive framework for untangling GW dynamics in high-resolution climate simulations (Vasylkevych and Žagar, 2021).

Deficiencies in simulated variability by the climate models are related to model biases which may originate in remote regions. Their improved understanding is possible within the normal-mode function framework discussed in Sections 3.1–3.2. Applying (37) to numerical model output and its verifying dataset, for example reanalysis, which is assumed unbiased, we obtain the following relationship between the spatial variability (energy spectra), temporal variability (variance spectra) and the model bias (Žagar *et al.*, 2020a):

$$\overline{I^c} - \overline{I^a} = \frac{1}{2} (V^c - V^a) + \frac{1}{2} B + P(\overline{\Delta W}, \overline{W^a}). \quad (51)$$

The indices *c* and *a* stand for the climate model simulation and reanalysis data, respectively, while other indices (*k*, *r* and *n*) have been dropped. The spectrum of the bias is denoted *B* and is defined as the variance of the time-averaged difference between the model and reanalysis:

$$B = gD \left| \overline{\Delta W} \right|^2, \quad \text{where} \quad \overline{\Delta W} = \overline{W^c} - \overline{W^a}. \quad (52)$$

The covariance term $P(\overline{\Delta W}, \overline{W^a})$ describes the covariance between the mean state of the verifying reanalysis and bias and is defined as

$$P(\overline{\Delta W}, \overline{W^a}) = \frac{1}{2} gD \left[\overline{\Delta W} [\overline{W^a}]^* + \left(\overline{\Delta W} [\overline{W^a}]^* \right)^* \right]. \quad (53)$$

Using (37) for the model and reanalysis, (51) can be rewritten as

$$E^c(\overline{W}) - E^a(\overline{W}) = \frac{1}{2} B + P(\overline{\Delta W}, \overline{W^a}). \quad (54)$$

Equation (54) states that the misrepresentation of the climatological spatial variance (energy of the mean state)

in the model with respect to the reanalysis is described by the two terms that account for the amplitude and phase of the bias. As an example, let us look at the case with a variability mode *W* (with all indices dropped) simulated with a perfect amplitude and a phase shift of $\Delta\phi$ with respect to the reanalysis. The time-mean state of the reanalysis is given by $\overline{W^a} = W_o e^{i\phi}$ and the simulated mean state is $\overline{W^c} = W_o e^{i(\phi+\Delta\phi)}$. The two components of the bias are $B = 2gDW_o^2(1 - \cos \Delta\phi)$ and $P = gDW_o^2(1 - \cos \Delta\phi)$. The simulated variance spectrum is perfect, $E^c = E^a = 0.5gDW_o^2$. The total bias is $0.5B + P = 2gDW_o^2(1 - \cos \Delta\phi)$ and, depending on the phase error, it may well exceed the variance. For example, for $\Delta\phi = \pi/2$, the bias is four times the variance.

Both terms on the right-hand side of (54) need to be evaluated for full information about bias associated with a particular mode, whereas filtering of $\Delta W_{n,r}^k$ per scale (indices *k* and *n*) and per mode (index *r*) to physical space may provide causal links between the biases in simulations and energy transfers within, and possibly also outside, the tropical atmosphere.

4.3 | A better understanding of upscale links

An important remaining question is to what level the background of coherent vertical motions is saturated or composed of discrete modes. Is the sky really white? There is a decent understanding of how large scales affect smaller scales, but the effects of mesoscale and sub-mesoscale dynamics on larger scales has presented a notorious challenge for theoretical understanding as well as for numerical modelling, as global models are needed to include large-scale motions, whereas a very fine horizontal mesh is required to resolve convection.

Convective processes, such as diabatic heating associated with cloud formation, act on scales that global models are beginning to resolve. There are good observations of clouds, but understanding the connection between moist processes and motion fields has been a long-standing goal, as the three-dimensional wind field is difficult to observe directly. Global kilometre-scale modelling makes the link between small-scale convection and the large-scale circulation a new frontier.

Conceptual models based on simplifying assumptions targeted to specific dynamical regimes are useful for conditional process understanding, but cannot account for the large number of observed scale interactions. The era of kilometre-scale global climate models that can account for many, albeit not all, scale interactions and the nonlinear behaviour of the atmosphere, is just beginning. Even global integrations of hectometre scale will soon become

available. Cloud scales have first-order effects on larger scales and mean aspects of climate. Clever ways of linking observations to hectometre simulations and to coarser but longer integrations harbour a great potential for advancing Earth system science. In this regard, improving the understanding of the large-scale effects of mesoscale and sub-mesoscale dynamics is a key step.

A fundamental requirement are observations that capture small temporal and spatial scales. We have given examples of how combining local measurement techniques, such as radiosondes and lidar, with remote-sensing measurements that cover the mesoscale (satellite and radar) is proving to be an efficient approach for quantifying the effects of cloud-scale dynamics on the larger-scale circulation. Thinking of new ways of using observations synergistically, to decompose the spectrum of variability, in order to isolate small-scale features and their link to the larger-scale dynamics, can reveal deep insights into physically and dynamically consistent atmospheric phenomena.

5 | FINAL REMARKS


It is sometimes argued that traditional ‘theory’ is highly idealized and that the atmospheric science community’s theoretical understanding is best represented in its most sophisticated atmospheric models. Whilst this might be a tenable position if models did not have systematic errors and deficiencies, and made highly reliable predictions, this is not the case. The new generation of kilometre-scale global atmospheric models will simulate an unprecedented range of interactions across scales, reaching into uncharted dynamical territory, and offers much hope for a more direct comparison with observations that span this range of scales than has hitherto been the case. Yet there will still be systematic errors and deficiencies, and limited sampling of the full range of natural variability. The challenge is thus to understand what we can reliably learn from the models, and what are the reasons for their remaining systematic errors and deficiencies. Especially for application to climate change, which is fundamentally unverifiable, the only sound basis for prediction is physical understanding of causal mechanisms.

The line of argument developed here is that it is fruitful to start from the observations, armed with the toolkit of theory. We have given many examples of how rather simple theory is surprisingly explanatory of observed atmospheric behaviour, despite the complexity of the latter. Theory provides the qualitative framework – the causal narrative – within which the quantitative information from models and observations can be meaningfully compared and usefully combined.

ACKNOWLEDGEMENTS

C.C.S. was supported by the Minerva Fast Track Program of the Max Planck Society. BCO data are freely accessible by a broad community, as detailed in Stevens *et al.* (2016), with the primary data provided at <https://www.mpimet.mpg.de/en/science/the-atmosphere-in-the-earth-system/working-groups/tropical-cloud-observation/barbadosstation1/instrumentation-and-data/>. NCEP reanalysis data are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at <https://www.esrl.noaa.gov/psd/>. The ERA-Interim reanalysis data are available from <http://www.ecmwf.int>. Visible (0.64 μm , Band 2) images of the GOES-E Advanced Baseline Imager (ABI; Schmit *et al.*, 2017) are provided by NASA’s Worldview application (<https://worldview.earthdata.nasa.gov>). We thank participants of the Cluster of Excellence Climate, Climatic Change and Society (CliCCS) Workshop on Waves, Bjorn Stevens who suggested the workshop, and Wenlin Xiao for drawing Figures 3 and 4, as well as Fiona Fix, Brian Hoskins, Julia Mindlin, Hauke Schmidt, Joe Tribbia, and two anonymous reviewers, for comments on the manuscript. We are especially grateful to the Associate Editor, Ed Gerber, for his patient and perceptive guidance on the substantial revisions of the manuscript. Open access funding enabled and organized by Projekt DEAL.

ORCID

Claudia C. Stephan  <https://orcid.org/0000-0001-5736-1948>

REFERENCES

- Adames, A.F. and Wallace, J.M. (2014) Three-dimensional structure and evolution of the MJO and its relation to the mean flow. *Journal of the Atmospheric Sciences*, 71, 2007–2026.
- Ahlquist, J.E. (1982) Normal-mode global Rossby waves: theory and observations. *Journal of the Atmospheric Sciences*, 39, 193–202.
- Alexander, M.J. and Holton, J.R. (2004) On the spectrum of vertically propagating gravity waves generated by a transient heat source. *Atmospheric Chemistry and Physics*, 4, 923–932.
- Andrews, D.G., Holton, J.R. and Leovy, C.B. (1987) *Middle Atmosphere Dynamics*. Academic Press.
- Baer, F. and Tribbia, J. (1977) On complete filtering of gravity modes through nonlinear initialization. *Monthly Weather Review*, 105, 1536–1539.
- Balaji, V. and Clark, T.L. (1988) Scale selection in locally forced convective fields and the initiation of deep cumulus. *Journal of the Atmospheric Sciences*, 45, 3189–3211.
- Balaji, V., Redelsperger, J.-L. and Klaassen, G.P. (1993) Mechanisms for the mesoscale organization of tropical cloud clusters in GATE Phase III. Part I. shallow cloud bands. *Journal of the Atmospheric Sciences*, 50, 3571–3589.
- Beven, K. (2016) Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, 61, 1652–1665.

- Bischoff, T. and Schneider, T. (2016) The equatorial energy balance, ITCZ position, and double-ITCZ bifurcations. *Journal of Climate*, 29, 2997–3013.
- Blaauw, M. and Žagar, N. (2018) Multivariate analysis of Kelvin wave seasonal variability in ECMWF L91 analyses. *Atmospheric Chemistry and Physics*, 18, 8313–8330.
- Blažica, V., Žagar, N., Strajnar, B. and Cedilnik, J. (2013) Rotational and divergent kinetic energy in the mesoscale model ALADIN. *Tellus A*, 65. <https://doi.org/10.3402/tellusa.v65i0.18918>.
- Boer, G.J. and Shepherd, T.G. (1983) Large-scale two-dimensional turbulence in the atmosphere. *Journal of the Atmospheric Sciences*, 40, 164–184.
- Bokhove, O. and Shepherd, T.G. (1996) On Hamiltonian balanced dynamics and the slowest invariant manifold. *Journal of the Atmospheric Sciences*, 53, 276–297.
- Boljka, L., Shepherd, T.G. and Blackburn, M. (2018) On the coupling between barotropic and baroclinic modes of extratropical atmospheric variability. *Journal of the Atmospheric Sciences*, 75, 1853–1871.
- Bony, S. and Stevens, B. (2019) Measuring area-averaged vertical motions with dropsondes. *Journal of the Atmospheric Sciences*, 76, 767–783.
- Booker, J.R. and Bretherton, F.P. (1967) The critical layer for internal gravity waves in a shear flow. *Journal of Fluid Mechanics*, 27, 513–539.
- Bretherton, C.S. and Smolarkiewicz, P.K. (1989) Gravity waves, compensating subsidence and detrainment around cumulus clouds. *Journal of the Atmospheric Sciences*, 46, 740–759.
- Burgess, B.H., Erler, A.R. and Shepherd, T.G. (2013) The troposphere-to-stratosphere transition in kinetic energy spectra and nonlinear spectral fluxes as seen in ECMWF analyses. *Journal of the Atmospheric Sciences*, 70, 669–687.
- Butchart, N., Cionni, I., Eyring, V., Shepherd, T.G., Waugh, D.W., Akiyoshi, H., Austin, J., Brühl, C., Chipperfield, M.P., Cordero, E., Dameris, M., Deckert, R., Dhomse, S., Frith, S.M., Garcia, R.R., Gettelman, A., Giorgetta, M.A., Kinnison, D.E., Li, F., Mancini, E., McLandress, C., Pawson, S., Pitari, G., Plummer, D.A., Rozanov, E., Sassi, F., Scinocca, J.F., Shibata, K., Steil, B. and Tian, W. (2010) Chemistry–climate model simulations of twenty-first century stratospheric climate and circulation changes. *Journal of Climate*, 23, 5349–5374.
- Byrne, N.J., Shepherd, T.G., Woollings, T. and Plumb, R.A. (2016) Annular modes and apparent eddy feedbacks in the Southern Hemisphere. *Geophysical Research Letters*, 43, 3897–3902.
- Castanheira, J. and Marques, C. (2015) Convectively coupled equatorial-wave diagnosis using three-dimensional normal modes. *Quarterly Journal of the Royal Meteorological Society*, 141, 2776–2792.
- Chan, I.H. and Shepherd, T.G. (2014) Diabatic balance model for the equatorial atmosphere. *Journal of the Atmospheric Sciences*, 71, 985–1001.
- Charney, J.G. (1963) A note on large-scale motions in the tropics. *Journal of the Atmospheric Sciences*, 20, 607–609.
- Charney, J.G. and Drazin, P.G. (1961) Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *Journal of Geophysical Research*, 66, 83–109.
- Chemke, R. and Polvani, L.M. (2019) Exploiting the abrupt 4x CO₂ scenario to elucidate tropical expansion mechanisms. *Journal of Climate*, 32, 859–875.
- Chen, G. and Held, I.M. (2007) Phase speed spectra and the recent poleward shift of Southern Hemisphere surface westerlies. *Geophysical Research Letters*, 34, L21805.
- Chimonas, G. (1993) Surface drag instabilities in the atmospheric boundary layer. *Journal of the Atmospheric Sciences*, 50, 1914–1924.
- Clark, T.L., Hauf, T. and Kuettner, J.P. (1986) Convectively forced internal gravity waves: results from two-dimensional numerical experiments. *Quarterly Journal of the Royal Meteorological Society*, 112, 899–925.
- Cohn, S.E. and Dee, D.P. (1989) An analysis of the vertical structure equation for arbitrary thermal profiles. *Quarterly Journal of the Royal Meteorological Society*, 115, 143–171.
- Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Holm, E., Isaksen, L., Kallberg, P., Kahler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.
- Dewan, E. (1979) Stratospheric wave spectra resembling turbulence. *Science*, 204, 832–835.
- Eckermann, S.D. and Vincent, R.A. (1989) Falling sphere observations of anisotropic gravity wave motions in the upper stratosphere over Australia. *Pure and Applied Geophysics*, 130, 509–532.
- Fritts, D.C. and Alexander, M.J. (2003) Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics*, 41. <https://doi.org/10.1029/2001RG000106>.
- Garcia, R.R. and Salby, M.L. (1987) Transient response to localized episodic heating in the tropics. Part II: far-field behavior. *Journal of the Atmospheric Sciences*, 44, 499–532.
- Geisler, J.E. and Stevens, D.E. (1982) On the vertical structure of damped steady circulation in the Tropics. *Quarterly Journal of the Royal Meteorological Society*, 108, 87–93.
- Gill, A.E. (1980) Some simple solution for heat-induced tropical circulation. *Quarterly Journal of the Royal Meteorological Society*, 106, 447–462.
- Gill, A.E. (1982) *Atmosphere–Ocean Dynamics*. New York, NY: Academic Press.
- Grant, L.D., Lane, T.P. and van den Heever, S.C. (2018) The role of cold pools in tropical oceanic convective systems. *Journal of the Atmospheric Sciences*, 75, 2615–2634.
- Heckley, W.A. and Gill, A.E. (1984) Some simple analytical solutions to the problem of forced equatorial long waves. *Quarterly Journal of the Royal Meteorological Society*, 110, 203–217.
- Held, I.M. and Hoskins, B.J. (1985) Large-scale eddies and the general circulation of the troposphere, in *Issues in Atmospheric and Oceanic Modeling*, Part A. Saltzman, B. (ed.), *Advances in Geophysics*, vol. 28, Amsterdam, Netherlands: Elsevier.
- Hoskins, B.J. and Karoly, D.J. (1981) The steady linear response of a spherical atmosphere to thermal and orographic forcing. *Journal of the Atmospheric Sciences*, 38, 1179–1196.
- Hoskins, B.J. and Yang, G.-Y. (2021) The detailed dynamics of the Hadley Cell. Part II: December–February. *Journal of Climate*, 34, 805–823.

- Hoskins, B.J., Simmons, A. and Andrews, D.G. (1977) Energy dispersion in a barotropic atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 103, 553–567.
- Hoskins, B.J., McIntyre, M.E. and Robertson, A.W. (1985) On the use and significance of isentropic potential vorticity maps. *Quarterly Journal of the Royal Meteorological Society*, 111, 877–946.
- Hoskins, B.J., Yang, G.-Y. and Fonseca, R. (2020) The detailed dynamics of the June–August Hadley cell. *Quarterly Journal of the Royal Meteorological Society*, 146, 557–575.
- Jeffreys, H. (1925) On fluid motions produced by differences of temperature and humidity. *Quarterly Journal of the Royal Meteorological Society*, 51, 347–356.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. (1996) The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77, 437–472.
- Kasahara, A. (1977) Numerical integration of the global barotropic primitive equations with Hough harmonic expansions. *Journal of the Atmospheric Sciences*, 34, 687–701.
- Kasahara, A. (1978) Further studies on a spectral model of the global barotropic primitive equations with Hough harmonic expansions. *Journal of the Atmospheric Sciences*, 35, 2043–2051.
- Kasahara, A. (1980) Effect of zonal flows on the free oscillations of a barotropic atmosphere. *Journal of the Atmospheric Sciences*, 37, 917–929. Corrigendum, *Journal of the Atmospheric Sciences*, 38 (1981), 2284–2285.
- Kasahara, A. (1984) The linear response of a stratified global atmosphere to a tropical thermal forcing. *Journal of the Atmospheric Sciences*, 41, 2217–2237.
- Kasahara, A. (2020). 3D Normal Mode Functions (NMFs) of a global baroclinic atmospheric model, in *Modal View Of Atmospheric Variability: Applications of Normal-Mode Function Decomposition in Weather and Climate Research*, Žagar, N, Tribbia, J. (eds). Berlin: Springer.
- Kasahara, A. and Puri, K. (1981) Spectral representation of three-dimensional global data by expansion in normal mode functions. *Monthly Weather Review*, 109, 37–51.
- Kiladis, G.N., Wheeler, M.C., Haertel, P.T., Straub, K.H. and Roundy, P.E. (2009) Convectively coupled equatorial waves. *Reviews of Geophysics*, 47. <https://doi.org/10.1029/2008RG000266>.
- Kitamura, Y. and Matsuda, Y. (2010) Energy cascade processes in rotating stratified turbulence with application to the atmospheric mesoscale. *Geophysical Research Letters*, 115, 3013–3030.
- Klocke, D., Brueck, M., Hohenegger, C. and Stevens, B. (2017) Rediscovery of the doldrums in storm-resolving simulations over the tropical Atlantic. *Nature Geoscience*, 10, 891–896.
- Koshyk, J.N. and Hamilton, K. (2001) The horizontal kinetic energy spectrum and spectral budget simulated by a high-resolution troposphere-stratosphere-mesosphere GCM. *Journal of the Atmospheric Sciences*, 58, 329–348.
- Kosovelj, K., Kucharski, F., Molteni, F. and Žagar, N. (2019) Modal decomposition of the global response to tropical heating perturbations resembling MJO. *Journal of the Atmospheric Sciences*, 76, 1457–1469.
- Kretschmer, M., Coumou, D., Donges, J.F. and Runge, J. (2016) Using causal effect networks to analyze different Arctic drivers of midlatitude winter circulation. *Journal of Climate*, 29, 4069–4081.
- Kunkel, D., Hoor, P., Kaluza, T., Ungermann, J., Kluschat, B., Giez, A., Lachnitt, H.-C., Kaufmann, M. and Riese, M. (2019) Evidence of small-scale quasi-isentropic mixing in ridges of extratropical baroclinic waves. *Atmospheric Chemistry and Physics*, 19, 12607–12630.
- Kushner, P., Held, I.M. and Delworth, T. (2001) Southern Hemisphere atmospheric circulation response to global warming. *Journal of Climate*, 14, 2238–2249.
- Lac, C., Lafore, J.-P. and Redelsperger, J.-L. (2002) Role of gravity waves in triggering deep convection during TOGA COARE. *Journal of the Atmospheric Sciences*, 59, 1293–1316.
- Lane, T.P. and Clark, T.L. (2002) Gravity waves generated by the dry convective boundary layer: two-dimensional scale selection and boundary-layer feedback. *Quarterly Journal of the Royal Meteorological Society*, 128, 1543–1570.
- Lane, T.P. and Moncrieff, M.W. (2015) Long-lived mesoscale systems in a low-convective inhibition environment. Part I: upshear propagation. *Journal of the Atmospheric Sciences*, 72, 4297–4318.
- Lane, T.P. and Zhang, F. (2011) Coupling between gravity waves and tropical convection at mesoscales. *Journal of the Atmospheric Sciences*, 68, 2582–2598.
- Lehmann, C.I., Kim, Y.-H., Preusse, P., Chun, H.-Y., Ern, M. and Kim, S.-Y. (2012) Consistency between Fourier transform and small-volume few-wave decomposition for spectral and spatial variability of gravity waves above a typhoon. *Atmospheric Measurement Techniques*, 5, 1637–1651.
- Longuet-Higgins, M.S. (1968) The eigenfunctions of Laplace's tidal equations over a sphere. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 262, 511–607.
- Lorenz, D.J. and Hartmann, D.L. (2001) Eddy-zonal flow feedback in the Southern Hemisphere. *Journal of the Atmospheric Sciences*, 58, 3312–3327.
- Machenhauer, B. (1977) On the dynamics of gravity oscillations in a shallow-water model, with applications to normal mode initialization. *Beiträge zur Physik der Atmosphäre*, 50, 253–271.
- Madden, R.A. (2019) How I learned to love normal-mode Rossby–Haurwitz waves. *Bulletin of the American Meteorological Society*, 100, 503–511.
- Matsuno, T. (1966) Quasi-geostrophic motions in the equatorial area. *Journal of the Meteorological Society of Japan*, 44, 25–42.
- May, P.T., Mather, J.H., Vaughan, G., Jakob, C., McFarquhar, G.M., Bower, K.N. and Mace, G.G. (2008) The Tropical Warm Pool International Cloud Experiment. *Bulletin of the American Meteorological Society*, 89, 629–645.
- McIntyre, M. (2002) Some fundamental aspects of atmospheric dynamics, with a solar spinoff. Pearce, R.P., *Meteorology at the Millennium. International Geophysics*. 83 Academic Press, 283–305.
- McIntyre, M.E. and Norton, W.A. (2000) Potential-vorticity inversion on a hemisphere. *Journal of the Atmospheric Sciences*, 57, 1214–1235.
- Moncrieff, M.W. and Lane, T.P. (2015) Long-lived mesoscale systems in a low-convective inhibition environment. Part II: downshear propagation. *Journal of the Atmospheric Sciences*, 72, 4319–4336.
- Nappo, C. (2002) *An Introduction to Atmospheric Gravity Waves*. Cambridge, MA: Academic Press. International Geophysics Series, vol. 85.

- Nastrom, G.D. and Gage, K.S. (1985) A climatology of aircraft wavenumber spectra observed by commercial aircraft. *Journal of the Atmospheric Sciences*, 42, 950–960.
- Neelin, J.D. and Held, I.M. (1987) Modeling tropical convergence based on the moist static energy budget. *Monthly Weather Review*, 115, 3–12.
- O’Gorman, P.A. and Schneider, T. (2007) Recovery of atmospheric flow statistics in a general circulation model without nonlinear eddy–eddy interactions. *Geophysical Research Letters*, 34(22). <https://doi.org/10.1029/2007GL031779>.
- Ogrosky, H.R. and Stechmann, S.N. (2015) Assessing the equatorial long-wave approximation: asymptotics and observational data analysis. *Journal of the Atmospheric Sciences*, 72, 4821–4843.
- Pearl, J. and Mackenzie, D. (2018) *The Book of Why: The New Science of Cause and Effect*. London: Penguin Books.
- Pedlosky, J. (1987) *Geophysical Fluid Dynamics* (2nd edn). Berlin: Springer.
- Podglajen, A., Hertzog, A., Plougonven, R. and Žagar, N. (2014) Assessment of the accuracy of (re)analyses in the equatorial lower stratosphere. *Journal of Geophysical Research*, 119, 11166–11188.
- Polichtchouk, I. and Shepherd, T.G. (2016) Zonal-mean circulation response to reduced air–sea momentum roughness. *Quarterly Journal of the Royal Meteorological Society*, 142, 2611–2622.
- Randel, W.J. and Held, I.M. (1991) Phase speed spectra of transient eddy fluxes and critical layer absorption. *Journal of the Atmospheric Sciences*, 48, 688–697.
- Raupp, C. and Dias, P.S. (2009) Resonant wave interactions in the presence of a diurnally varying heat source. *Journal of the Atmospheric Sciences*, 66, 3165–3183.
- Rotunno, R., Klemp, J.B. and Weisman, M.L. (1988) A theory for strong, long-lived squall lines. *Journal of the Atmospheric Sciences*, 45, 463–485.
- Runge, J., Petoukhov, V. and Kurths, J. (2014) Quantifying the strength and delay of climatic interactions: the ambiguities of cross-correlation and a novel measure based on graphical models. *Journal of Climate*, 27, 720–739.
- Sakazaki, T. and Hamilton, K. (2020) An array of ringing global free modes discovered in tropical surface pressure data. *Journal of the Atmospheric Sciences*, 77, 2519–2539.
- Salby, M. (1981) Rossby normal modes in non-uniform background configurations, Part II: equinox and solstice conditions. *Journal of the Atmospheric Sciences*, 38, 1827–1840.
- Salby, M.L. and Garcia, R.R. (1987) Transient response to localized episodic heating in the Tropics. Part I: excitation and short-time near-field behavior. *Journal of the Atmospheric Sciences*, 44, 458–498.
- Sardeshmukh, P.D. and Hoskins, B.J. (1988) The generation of global rotational flow by steady idealized tropical divergence. *Journal of the Atmospheric Sciences*, 45, 1228–1251.
- Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W.M. and Düben, P. (2019) Global cloud-resolving models. *Current Climate Change Reports*, 5, 172–184.
- Saujani, S. and Shepherd, T.G. (2006) A unified theory of balance in the extratropics. *Journal of Fluid Mechanics*, 569, 447–464.
- Schär, C., Fuhrer, O., Arteaga, A., Ban, N., Charpillot, C., Girolamo, S.D., Hentgen, L., Hoefler, T., Lapillonne, X., Leutwyler, D., Osterried, K., Panosetti, D., Rüdisühli, S., Schlemmer, L., Schulthess, T.C., Sprenger, M., Ubbiali, S. and Wernli, H. (2019) Kilometer-scale climate models: prospects and challenges. *Bulletin of the American Meteorological Society*, 101, 567–587.
- Schmit, T.J., Griffith, P., Gunshor, M.M., Daniels, J.M., Goodman, S.J. and Lebar, W.J. (2017) A closer look at the ABI on the GOES-R series. *Bulletin of the American Meteorological Society*, 98, 681–698.
- Schoon, L. and Zülicke, C. (2018) A novel method for the extraction of local gravity wave parameters from gridded three-dimensional data: description, validation, and application. *Atmospheric Chemistry and Physics*, 18, 6971–6983.
- Schumann, U. (2019) The horizontal spectrum of vertical velocities near the tropopause from global to gravity wave scales. *Journal of the Atmospheric Sciences*, 76, 3847–3862.
- Shaw, T. (2019) Mechanisms of future predicted changes in the zonal mean mid-latitude circulation. *Current Climate Change Reports*, 5, 345–357.
- Shepherd, T.G. (2019) Storyline approach to the construction of regional climate change information. *Proceedings of the Royal Society A*, 475, 20190013.
- Shepherd, T.G. and McLandress, C. (2011) A robust mechanism for strengthening of the Brewer–Dobson circulation in response to climate change: critical-layer control of subtropical wave breaking. *Journal of the Atmospheric Sciences*, 68, 784–797.
- Shige, S. and Satomura, T. (2001) Westward generation of eastward-moving tropical convective bands in TOGA COARE. *Journal of the Atmospheric Sciences*, 58, 3724–3740.
- Sigmond, M. and Shepherd, T.G. (2014) Compensation between resolved wave driving and parameterized orographic gravity wave driving of the Brewer–Dobson circulation and its response to climate change. *Journal of Climate*, 27, 5601–5610.
- Sobel, A., Polvani, N. and Nilsson, J. (2001) The weak temperature gradient approximation. *Journal of the Atmospheric Sciences*, 58, 1473–1489.
- Stechmann, S.N. and Majda, A.J. (2009) Gravity waves in shear and implications for organized convection. *Journal of the Atmospheric Sciences*, 66, 2579–2599.
- Stephan, C.C. (2020) Seasonal modulation of trapped gravity waves and their imprints on trade wind clouds. *Journal of the Atmospheric Sciences*, 77, 2993–3009.
- Stephan, C.C. (2021) Mechanism for the formation of arc-shaped cloud lines over the tropical oceans. *Journal of the Atmospheric Sciences*, 78, 817–824.
- Stephan, C.C. and Alexander, M.J. (2015) Realistic simulations of atmospheric gravity waves over the continental U.S. using precipitation radar data. *Journal of Advances in Modeling Earth Systems*, 7, 823–835.
- Stephan, C.C. and Mariaccia, A. (2021) The signature of the tropospheric gravity wave background in observed mesoscale motion. *Weather and Climate Dynamics*, 2, 359–372. <https://doi.org/10.5194/wcd-2-359-2021>.
- Stephan, C.C., Alexander, M.J., Hedlin, M., de Groot-Hedlin, C.D. and Hoffmann, L. (2016) A case study on the far-field properties of propagating tropospheric gravity waves. *Monthly Weather Review*, 144, 2947–2961.
- Stephan, C.C., Lane, T.P. and Jakob, C. (2020) Gravity wave influences on mesoscale divergence: an observational case study. *Geophysical Research Letters*, 47(1). <https://doi.org/10.1029/2019GL086539>.
- Stephan, C.C., Schnitt, S., Schulz, H., Bellenger, H., de Szoeke, S.P., Acquistapace, C., Baier, K., Dauhut, T., Laxenaire, R.,

- Morfa-Avalos, Y., Person, R., Quiñones Meléndez, E., Bagheri, G., Böck, T., Daley, A., Güttler, J., Helfer, K.C., Los, S.A., Neuberger, A., Röttenbacher, J., Raeke, A., Ringel, M., Ritschel, M., Sadoulet, P., Schirmacher, I., Stolla, M.K., Wright, E., Charpentier, B., Doerenbecher, A., Wilson, R., Jansen, F., Kinne, S., Reverdin, G., Speich, S., Bony, S. and Stevens, B. (2021) Ship- and island-based atmospheric soundings from the 2020 EUREC⁴a field campaign. *Earth System Science Data*, 13, 491–514. <https://essd.copernicus.org/articles/13/491/2021/>.
- Stevens, B. and Brenguier, J.-L. (2009) Cloud-controlling factors: low clouds, *Clouds in the Perturbed Climate System*. Cambridge, MA: MIT Press, pp. 173–196.
- Stevens, B., Farrell, D., Hirsch, L., Jansen, F., Nuijens, L., Serikov, I., Brüggmann, B., Forde, M., Linne, H., Lonitz, K. and Prospero, J.M. (2016) The Barbados cloud observatory: anchoring investigations of clouds and circulation on the edge of the ITCZ. *Bulletin of the American Meteorological Society*, 97, 787–801.
- Stevens, B., Bony, S., Brogniez, H., Hentgen, L., Hohenegger, C., Kiemle, C., L'Ecuyer, T.S., Naumann, A.K., Schulz, H., Siebesma, P.A., Vial, J., Winker, D.M. and Zuidema, P. (2020) Sugar, gravel, fish and flowers: mesoscale cloud patterns in the trade winds. *Quarterly Journal of the Royal Meteorological Society*, 146, 141–152.
- Stevens, B., Bony, S., Farrell, D., Ament, F., Blyth, A., Fairall, C., Karstensen, J., Quinn, P.K., Speich, S., Acquistapace, C., Aemisegger, F., Albright, A.L., Bellenger, H., Bodenschatz, E., Caesar, K.-A., Chewitt-Lucas, R., de Boer, G., Delanoë, J., Denby, L., Ewald, F., Fildier, B., Forde, M., George, G., Gross, S., Hagen, M., Hausold, A., Heywood, K.J., Hirsch, L., Jacob, M., Jansen, F., Kinne, S., Klocke, D., Kölling, T., Konow, H., Lothon, M., Mohr, W., Naumann, A.K., Nuijens, L., Olivier, L., Pincus, R., Pöhlker, M., Reverdin, G., Roberts, G., Schnitt, S., Schulz, H., Siebesma, A.P., Stephan, C.C., Sullivan, P., Touzé-Peiffer, L., Vial, J., Vogel, R., Zuidema, P., Alexander, N., Alves, L., Arix, S., Asmath, H., Bagheri, G., Baier, K., Bailey, A., Baranowski, D., Baron, A., Barrau, S., Barrett, P.A., Batier, F., Behrendt, A., Bendinger, A., Beucher, F., Bigorre, S., Blades, E., Blossy, P., Bock, O., Böing, S., Bosser, P., Bourras, D., Bouruet-Aubertot, P., Bower, K., Branellec, P., Branger, H., Brennek, M., Brewer, A., Brilouet, P.-E., Brüggmann, B., Buehler, S.A., Burke, E., Burton, R., Calmer, R., Canonici, J.-C., Carton, X., Cato, Jr.G., Charles, J.A., Chazette, P., Chen, Y., Chilinski, M.T., Choularton, T., Chuang, P., Clarke, S., Coe, H., Cornet, C., Coutris, P., Couvreux, F., Crewell, S., Cronin, T., Cui, Z., Cuypers, Y., Daley, A., Damerell, G.M., Dauhut, T., Deneke, H., Desbios, J.-P., Dörner, S., Donner, S., Douet, V., Drushka, K., Dütsch, M., Ehrlich, A., Emanuel, K., Emmanouilidis, A., Etienne, J.-C., Etienne-Leblanc, S., Faure, G., Feingold, G., Ferrero, L., Fix, A., Flamant, C., Flatau, P.J., Foltz, G.R., Forster, L., Furtuna, I., Gadian, A., Galewsky, J., Gallagher, M., Gallimore, P., Gaston, C., Gentemann, C., Geyskens, N., Giez, A., Gollop, J., Gouirand, I., Gourbeyre, C., de Graaf, D., de Groot, G.E., Grosz, R., Güttler, J., Gutleben, M., Hall, K., Harris, G., Helfer, K.C., Henze, D., Herbert, C., Holanda, B., Ibanez-Landeta, A., Intrieri, J., Iyer, S., Julien, F., Kalesse, H., Kazil, J., Kellman, A., Kidane, A.T., Kirchner, U., Klingebiel, M., Körner, M., Kremper, L.A., Kretzschmar, J., Krüger, O., Kumala, W., Kurz, A., L'Hégaret, P., Labaste, M., Lachlan-Cope, T., Laing, A., Landschützer, P., Lang, T., Lange, D., Lange, I., Laplace, C., Lavik, G., Laxenaire, R., Le Bihan, C., Leandro, M., Lefevre, N., Lena, M., Lenschow, D., Li, Q., Lloyd, G., Los, S., Losi, N., Lovell, O., Luneau, C., Makuch, P., Malinowski, S., Manta, G., Marinou, E., Marsden, N., Masson, S., Maury, N., Mayer, B., Mayers-Als, M., Mazel, C., McGeary, W., McWilliams, J.C., Mech, M., Mehlmann, M., Meroni, A.N., Mieslinger, T., Minikin, A., Minnett, P., Möller, G., Morfa Avalos, Y., Muller, C., Musat, I., Napoli, A., Neuberger, A., Noisel, C., Noone, D., Nordsiek, F., Nowak, J.L., Oswald, L., Parker, D.J., Peck, C., Person, R., Philippi, M., Plueddemann, A., Pöhlker, C., Pörtge, V., Pöschl, U., Pologne, L., Posyniak, M., Prange, M., Quiñones Meléndez, E., Radtke, J., Ramage, K., Reimann, J., Renault, L., Reus, K., Reyes, A., Ribbe, J., Ringel, M., Ritschel, M., Rocha, C.B., Rochetin, N., Röttenbacher, J., Rollo, C., Royer, H., Sadoulet, P., Saffin, L., Sandiford, S., Sandu, I., Schäfer, M., Schemann, V., Schirmacher, I., Schlenczek, O., Schmidt, J., Schröder, M., Schwarzenboeck, A., Sealy, A., Senff, C.J., Serikov, I., Shohan, S., Siddle, E., Smirnov, A., Späth, F., Spooner, B., Stolla, M.K., Szkółka, W., de Szoeko, S.P., Tarot, S., Tetoni, E., Thompson, E., Thomson, J., Tomassini, L., Totems, J., Ubele, A.A., Villiger, L., von Arx, J., Wagner, T., Walther, A., Webber, B., Wendisch, M., Whitehall, S., Wiltshire, A., Wing, A.A., Wirth, M., Wiskandt, J., Wolf, K., Worbes, L., Wright, E., Wulfmeyer, V., Young, S., Zhang, C., Zhang, D., Ziemann, F., Zinner, T. and Zöger, M. (2021) EUREC⁴a. *Earth System Science Data Discussion*, 2021, 1–78. <https://doi.org/10.5194/essd-2021-18>.
- Su, T. and Zhai, G. (2017) The role of convectively generated gravity waves on convective initiation: a case study. *Monthly Weather Review*, 145, 335–359.
- Swartztrauber, P.N. and Kasahara, A. (1985) The vector harmonic analysis of Laplace tidal equations. *SIAM Journal on Scientific Computing*, 6, 464–491.
- Temperton, C. and Williamson, D.L. (1981) Normal mode initialization for a multilevel grid-point model. Part I: linear aspects. *Monthly Weather Review*, 109, 729–743.
- Terasaki, K., Tanaka, H. and Žagar, N. (2011) Energy spectra of Rossby and gravity waves. *SOLA*, 7, 45–48.
- Tindall, J.C., Thuburn, J. and Highwood, E.J. (2006) Equatorial waves in the lower stratosphere. I: a novel detection method. *Quarterly Journal of the Royal Meteorological Society*, 132, 177–194.
- Tribbia, J. (1979) Non-linear initialization on an equatorial beta plane. *Monthly Weather Review*, 107, 704–713.
- Tulich, S.N. and Mapes, B.E. (2008) Multiscale convective wave disturbances in the tropics: insights from a two-dimensional cloud-resolving model. *Journal of the Atmospheric Sciences*, 65, 140–155.
- Vasylkevych, S. and Žagar, N. (2021) A high-accuracy global prognostic model for the simulation of Rossby and gravity wave dynamics. *Quarterly Journal of the Royal Meteorological Society*, 147, 1989–2007.
- Žagar, N. and Franzke, C. (2015) Systematic decomposition of MJO into balanced and inertio-gravity modes. *Geophysical Research Letters*, 42, 6829–6835.
- Žagar, N., Gustafsson, N. and Källén, E. (2004) Variational data assimilation in the tropics: the impact of a background-error constraint. *Quarterly Journal of the Royal Meteorological Society*, 130, 103–125.
- Žagar, N., Andersson, E. and Fisher, M. (2005) Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *Quarterly Journal of the Royal Meteorological Society*, 131, 987–1011.
- Žagar, N., Tribbia, J., Anderson, J.L. and Raeder, K. (2009) Uncertainties of estimates of inertio-gravity energy in the atmosphere.

- Part II: large-scale equatorial waves. *Monthly Weather Review*, 137, 3858–3873. Corrigendum, *Monthly Weather Review*, 138, 2476–2477.
- Žagar, N., Terasaki, K. and Tanaka, H. (2012) Impact of the vertical discretization of analysis data on the estimates of atmospheric inertia-gravity energy. *Monthly Weather Review*, 140, 2297–2307.
- Žagar, N., Isaksen, L., Tan, D. and Tribbia, J. (2013) Balance and flow-dependency of background-error variances in the ECMWF 4D-Var ensemble. *Quarterly Journal of the Royal Meteorological Society*, 139, 1229–1238.
- Žagar, N., Buizza, R. and Tribbia, J. (2015) A three-dimensional multivariate modal analysis of atmospheric predictability with application to the ECMWF ensemble. *Journal of the Atmospheric Sciences*, 72, 4423–4444.
- Žagar, N., Jelić, D., Blaauw, M. and Bechtold, P. (2017) Energy spectra and inertia-gravity waves in global analyses. *Journal of the Atmospheric Sciences*, 74, 2447–2466.
- Žagar, N., Kosovelj, K., Manzini, E., Horvat, M. and Castanheira, J. (2020a) An assessment of scale-dependent variability and bias in global prediction models. *Climate Dynamics*, 54, 287–306. <https://doi.org/10.1007/s00382-019-05001-x>.
- Žagar, N., Zlotnik, Ž. and Karami, K. (2020b) Atmospheric sub-seasonal variability and circulation regimes: spectra, trends, and uncertainties. *Journal of Climate*, 33, 9375–9390.
- Warn, T., Bokhove, O., Shepherd, T.G. and Vallis, G.K. (1995) Rossby number expansions, slaving principles, and balance dynamics. *Quarterly Journal of the Royal Meteorological Society*, 121, 723–739.
- Wheeler, M.C. and Kiladis, G.N. (1999) Convectively coupled equatorial waves: analysis of clouds and temperature in the wavenumber–frequency domain. *Journal of the Atmospheric Sciences*, 56, 374–399.
- Wirosoetisno, D. and Shepherd, T.G. (2000) Averaging, slaving and balance dynamics in a simple atmospheric model. *Physica D*, 141, 37–53.
- Wirosoetisno, D., Shepherd, T.G. and Temam, R.M. (2002) Free gravity waves and balanced dynamics. *Journal of the Atmospheric Sciences*, 59, 3382–3398.
- Wright, C.J., Hindley, N.P., Hoffmann, L., Alexander, M.J. and Mitchell, N.J. (2017) Exploring gravity wave characteristics in 3-D using a novel S-transform technique: AIRS/Aqua measurements over the southern Andes and Drake passage. *Atmospheric Chemistry and Physics*, 17, 8553–8575.
- Yang, G.-Y. and Hoskins, B.J. (2017) The equivalent barotropic structure of waves in the tropical atmosphere in the western hemisphere. *Journal of the Atmospheric Sciences*, 74, 1689–1704.
- Yang, G.-Y., Hoskins, B.J. and Slingo, J. (2003) Convectively coupled equatorial waves: a new methodology for identifying wave structures in observational data. *Journal of the Atmospheric Sciences*, 60, 1637–1654.
- Yang, G.-Y., Hoskins, B.J. and Slingo, J. (2007) Convectively coupled equatorial waves. Part I: horizontal and vertical structures. *Journal of the Atmospheric Sciences*, 64, 3406–3423.

How to cite this article: Stephan, C.C., Žagar, N. & Shepherd, T.G. (2021) Waves and coherent flows in the tropical atmosphere: New opportunities, old challenges. *Quarterly Journal of the Royal Meteorological Society*, 147(738), 2597–2624. Available from: <https://doi.org/10.1002/qj.4109>