Accounting for the three-dimensional nature of mountain waves: parametrizing partial critical level filtering

Article

Accepted Version

Accounting for the three-dimensional nature of mountain waves: parametrizing partial critical level filtering

A. van Niekerk[1,2] | S.B. Vosper[1] | M.A.C. Teixeira[3]

[1]Met Office, Exeter, UK

[2]European Centre for Medium-Range Weather Forecasts, Reading, UK

[3]University of Reading, Berkshire, UK

Correspondence
Annelize van Niekerk, ECMWF
Email: annelize.vanniekerk@ecmwf.int

Present address
ECMWF, Reading, UK

Funding information

Gravity waves generated by mountains are multi-scale and three-dimensional. Current orographic gravity wave drag parametrization schemes assume that the waves are two-dimensional, varying only in the vertical and along one horizontal direction. These schemes, therefore, do not represent the process of partial critical level filtering, whereby a portion of the wave spectrum is saturated where the winds parallel to the wavevectors become small. This results in an unrealistic vertical distribution of the momentum flux and forcing of the waves on the mean flow. In this work, a method of accounting for partial critical level filtering in an orographic gravity wave drag parametrization using the full spectrum of realistic topography is presented. This is achieved through binning of the expression for linear hydrostatic surface stresses, computed using Fourier transforms of the sub-grid orographic heights within model gridboxes, into wavevector directions. The parametrization is compared with idealised non-linear simulations of flow over complex topography and is shown to perform well as the number of wavevector direction bins is increased. Implementation of the scheme into the Met Office Unified model is tested using short-range 5-day forecasts. As is found from idealised simulations, the binned scheme leads to less forcing in the troposphere and increased forcing in the stratosphere within the model. The binned scheme is shown to alleviate biases in the upper stratosphere, between 45 km and 65

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/qj.4421

km, as well as having significant local effects in the troposphere.

Keywords — Gravity wave drag, orography, parametrization, numerical weather prediction, climate modelling, stratosphere

1 | INTRODUCTION AND BACKGROUND

As a result of the steady increase in horizontal resolution of atmospheric models over time, weather and climate centres have seen improved predictability of the large scale circulation across timescales. At first sight, this may appear to suggest that an increase in horizontal resolution is the only way forward for models. However, km-scale modelling requires exa-scale computing, consuming large amounts of energy and producing vasts amounts of data. Even if one could afford the computational cost, comparisons with high resolution observations have shown that even grid-spacings of ~3 km are not sufficient to capture the full spectrum of gravity waves within the atmosphere (Kruse et al., 2022). Another factor in computational cost is the need to run atmospheric models for climate projections using large ensembles at relatively coarse horizontal resolutions, in which mountains are poorly resolved. Many recent studies have demonstrated that the improvements seen in the troposphere when horizontal resolution is increased is dominated by the increase in the resolved orographic detail (Kanehama et al. 2019; Berckmans et al. 2013; Heim et al. 2020). What this implies is that there is significant potential for increased skill from improvements in parametrization of orographic processes at coarser resolutions. The power of this is that it provides a process, while broad, and mechanisms upon which to focus parametrization development (Sandu et al., 2019). It has also been shown that the increase in resolved orographic (Polichtchouk, van Niekerk and Wedi, 2022), as well as nonorographic (Polichtchouk, Wedi and Kim, 2022), gravity wave forcing with increasing resolution plays a significant role in the improved model skill (van Niekerk et al., 2018). To make progress in accurately representing subgrid orographic processes, we focus here on the development of gravity wave drag parametrizations to include directional effects from realistic three-dimensional orography.

Linear orographic gravity wave theory has been relatively well verified against both models and real-world observations. Using this theory to produce analytic solutions for how orographic waves behave in certain flows (Phillips, 1984) and translating these into subgrid parametrizations has led to significant improvements in model fidelity, particularly in the stratosphere (e.g. McFarlane 1987; Palmer et al. 1986). While the bulk effects of these waves may be represented in the current generation of parametrizations, there are many missing processes that become dominant under particular flow regimes that are not accounted for at all. For example: the three-dimensional refraction and advection of waves by the background flow that can lead to lateral propagation, instead of vertical propagation, is not represented due to the columnar approach taken by model parametrizations (Sato et al., 2012; Amemiya and Sato, 2016); the steady state approximation adopted in most gravity wave parametrizations means that the additional non-dissipative mean-flow forcing as well as a transient modulation of the wave field by the resolved flow is neglected (Bölöni et al., 2021; Kruse and Smith, 2018); and the linear approximation itself means that the theory is only valid for small perturbations and nonlinear processes such as self-acceleration are not commonly represented (Scinocca and Sutherland, 2010; Fritts et al., 2020). Even though the theory for some of these unrepresented processes does exist and their impacts have been shown to be important under certain conditions, orographic gravity wave drag parametrizations used by many operational models still rely on many simplifying assumptions, used for ease of both analytical and numerical computation (Plougonven et al., 2020). A significant simplification made by almost all current orographic gravity wave drag schemes is that the waves being generated by mountains are monochromatic (c.f. McFarlane 1987; Palmer et al. 1986; Kim and Arakawa 1995; Lott and Miller 1997). This means that the parametrization assumes that a single wave, with a single wavevector direction and only one horizontal length-scale is launched within each model grid-box, rather than the full three-dimensional spectrum of waves from real orography. The work of van Niekerk and Vosper (2021) (hereafter VV21) was at least able to relax the approximation

that the waves generated by flow over orography consisted of only one lengthscale. This was achieved by computing Fourier transforms of the mountains from an observational dataset and using this directly within the surface stress expression, as opposed to assuming elliptical mountains. They showed that improvement in modelled stratospheric winds at coarser resolutions could be achieved due to the scheme being more 'scale-aware' and, hence, better behaved across grid-lengths. However, this scheme still assumes that these waves have only one wavevector direction.

Several observational, numerical and theoretical studies have shown that three-dimensional gravity waves, made up of several wavevectors, propagating through directionally sheared flow are likely to encounter critical levels at different heights (Shutts and Gadian, 1999; Doyle and Jiang, 2006; Eckermann et al., 2007) depending on the winds relative to the wavevector directions. Building upon this body of work, we investigate here an extension to the scale-aware orographic gravity wave drag parametrization described in VV21 to include the effects of partial critical level filtering in sheared flow. This will provide a more faithful representation of the three-dimensional nature of mountain waves within a model grid-box.

When three-dimensional mountains generate gravity waves, the waves themselves are also three-dimensional, consisting of many scales in many directions. Partial critical level filtering of these waves can be understood through linear theory. The vertical velocity field generated by these waves can be expressed in two-dimensional Fourier space and as a function of height z using (see, for example, Nappo 2013; Teixeira 2014):

$$w'(x,y,z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{w}(k,l,z) exp(i(kx+ly))dkdl \tag{1}$$

where a prime indicates the perturbation from the large-scale mean flow (which is assumed to be horizontally homogeneous), $\hat{(..)}$ indicates the Fourier transform and k and l are the zonal and meridional wave numbers, such that they represent the lengthscale and direction of the wavevectors.

It is common practise to assume that the model's unresolved orographic gravity waves are small enough so that they are unaffected by rotation but large enough so that their horizontal scales are much larger than vertical scales of variation (i.e. in hydrostatic balance). This scale separation is justified based on the fact that the model grid-lengths are generally between 100 km to 10 km, although the hydrostatic scale limit is imposed through filtering of the unresolved orography (see van Niekerk and Vosper (2021) for discussion). With this scale separation in mind, it is worth emphasising that horizontal refraction of waves due to inhomogeneity of the background flow is typically small for orographic waves (Hasha et al., 2008; Amemiya and Sato, 2016), and so this process is also generally neglected and horizontal homogeneity is assumed. Using these approximations, it is possible to derive a simple equation for the vertical velocity induced by the mountains (the Taylor-Goldstein equation):

$$\frac{d^2\hat{w}_B}{dz^2} + \left[\frac{N^2K^2}{(Uk+Vl)^2} - \frac{\frac{d^2U}{dz^2}k + \frac{d^2V}{dz^2}l}{Uk+Vl}\right]\hat{w_B} = 0 \tag{2}$$

where $K = (k^2+l^2)^{1/2}$ is the total wavenumber, N is the Brunt-Vaisala frequency, U and V are the large-scale mean zonal and meridional winds, respectively. Following Bretherton (1966) and Smith (1979), vertical variations in density are accounted for in equation (2) by a change of variable such that $\hat{w_B} = \left(\frac{\rho}{\rho_0}\right)^{1/2}\hat{w}$.

In order to find solutions to this equation, a further assumption that the envelope of the quasi-monochromatic wave packet must be of much larger scale than the vertical wavelength is made, by using a first-order Wentzel–Kramers–Brillouin (WKB) approximation. The WKB approximation does place constraints on the vertical wavelength of the waves and does not account for the interaction of waves, which means it is only valid for small (linear) perturbations. A solution can then be found for $\hat{w}$, following Bretherton (1966); Teixeira and Miranda (2009), given by:

$$\hat{w}(k,l,z) = \hat{w}_0 \left(\frac{\rho_0 N_0 (Uk + Vl)}{\rho N (U_0 k + V_0 l)} \right)^{1/2} \exp\left(iRe\left(\int_0^z m(z')dz' \right) \right) \tag{3}$$

where m is the vertical wavenumber and Re indicates the real part. Subscripts 0 denote that the value is taken at the surface. It is important to note that, under the WKB approximation at this order, this definition of the vertical velocity is valid locally at each height z (below a critical level). To zeroth order $m = \frac{NK}{Uk+Vl}$ (see Teixeira and Miranda 2009 for the solution of equation 3 via asymptotic expansion of m), which means that, when the winds become perpendicular to the wavevector direction (i.e. $(Uk + Vl) \rightarrow 0$), the vertical group velocity tends towards zero and the vertical wavenumber m tends towards infinity. This is known as a critical level and, in reality, leads to wave breaking near the altitude where it occurs. What is also evident is that each particular wavevector (k, l) can have its own critical level, unless the total winds become zero, meaning that some part of the wave spectrum may encounter critical levels while others do not. This is what is known as partial critical level filtering. Since these gravity waves carry momentum, body forces are applied to the mean flow where they encounter critical levels.

The mean momentum transport is dominated by the horizontally averaged vertical flux of horizontal momentum, given by:

$$\tau_x, \tau_y = A^{-1} \rho \int \int_A (u', v') w' dx dy \tag{4}$$

where A is a given area, such as a model grid-box. It is the vertical divergence of this flux, i.e. $\frac{d}{dt}(U,V)_{GWD} = -\frac{1}{\rho}\frac{d}{dz}(\tau_x, \tau_y)$, that leads to forcing from the waves onto the mean flow.

At the surface, it is assumed that the flow follows the terrain and the vertical velocity perturbation is given $w' = U\frac{\partial h}{\partial x} + V\frac{\partial h}{\partial y}$. The linearised kinematic surface boundary condition $\hat{w}_0 = i(U_0 k + V_0 l)\hat{h}$ (Smith, 1980) can then be used in the expression for the surface momentum flux in Fourier space:

$$\tau_x, \tau_y = A^{-1} 4\pi^2 \rho_0 N_0 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{(k,l)}{K} |\hat{h}|^2 (U_0 k + V_0 l) dk dl \tag{5}$$

where $\hat{h}$ is the Fourier transform of the surface height. It is assumed that the waves are forced by an isolated mountain, such that the perturbations, e.g. u' and w', decay to zero at the edges of the domain considered (area A), allowing us to equate (4) and (5). What is more, the mountain amplitude is assumed to be small enough to generate linear waves, see discussion on flow blocking presented in van Niekerk and Vosper (2021).

Broad (1995) emphasised that, owing to the three-dimensional nature of mountain waves, every vertical level has the potential to be a critical level, since it may be perpendicular to any of the infinitely many wavevectors generated by orography. A consequence of this is that the gravity wave momentum flux may be continuously varying with height as more of the wave spectrum is removed by partial critical level filtering and, since the stress vector is determined by the integral over wavevectors, the stress vector direction also changes with height. The reader is referred to Teixeira and Miranda (2009) for a derivation of the expression for the vertically varying momentum flux, which makes this wave filtering more evident. However, current orographic gravity wave drag parametrizations take a two-dimensional approach to determining wave saturation and critical levels. Most schemes used operationally (McFarlane, 1987; Palmer et al., 1986; Kim and Arakawa, 1995; Lott and Miller, 1997; van Niekerk and Vosper, 2021) assume that there is only one wavevector, the direction of which does not vary with height, and that this is parallel to the surface stress vector integrated over all directions. This means that only one critical level is diagnosed for the entire spectrum of waves produced within a model grid-box, which seems like a gross approximation when one considers the fact

that real mountains have a broad spectrum of scales and directions.

With the aim to improve the model accuracy, we relax the assumption of a single dominant horizontal wave vector and account for the effect of partial critical level filtering in the orographic gravity wave drag parametrization of the Met Office Unified Model. While Shutts (1995), Xu et al. (2019) and Teixeira and Yu (2014) were able to formulate parametrization schemes that accounted for directional critical level filtering, their formulations relied on analytic solutions of elliptical or bell-shaped mountains. As pointed out by Guarino et al. (2018), the spectrum of wave directions is inherently connected to the orography generating it and the critical level height and direction of wave forcing is, therefore, strongly modulated by the choice of orographic description. This motivates the use of realistic orography in our parametrization. Here, we make use of observed subgrid orography from the 1 km GMTED dataset (Danielson and Gesch, 2018), 300m Shuttle Radar Topography Mission (SRTM;Van Zyl 2001) and Fourier transforms thereof, to determine the momentum flux from three-dimensional orography and the resultant forcing on the mean flow.

The layout of the paper is as follows. Since one of the main aims of this work is to provide a method of accounting for critical level filtering using realistic orography, we begin with a substantial methodology in section 2. In this section, the surfaces stress calculation, method of wave saturation and launch amplitude are described. Section 3 goes on to describe the testing performed on the scheme, including the models and diagnostics used to do so. Results are then presented in section 4. Conclusions are drawn and discussion is had in section 5.

2 | PARAMETRIZATION METHODOLOGY

This section describes how the directional critical level filtering is handled in the parametrization. Since this scheme builds upon that of VV21, we refer the reader to that paper for details that have been omitted here.

2.1 | Surface stress calculation

The momentum flux at the surface is one of the most important inputs to orographic gravity wave drag parametrization and here we make use of the hydrostatic linear theory, partially described in the introduction, to determine the momentum flux in each wavevector direction. We begin by writing the expression for the surface momentum flux in equation (5) in terms of the total wavenumber K and the wavevector angle $\phi = \arctan(l/k)$, using the transform from Cartesian to polar coordinates $\int_{l_{min}}^{l_{max}} \int_{k_{min}}^{k_{max}} f(k,l)\,dk\,dl = \int_{-\pi}^{\pi} \int_{K_{min}}^{K_{max}} f(K,\phi)K\,dK\,d\phi$

$$\tau_x, \tau_y = 4\pi^2 A^{-1} \rho_0 N \int_{-\pi}^{\pi} \int_{K_{min}}^{K_{max}} K^2(\cos\phi, \sin\phi)|\hat{h}|^2(U_0\cos\phi + V_0\sin\phi)\,dK\,d\phi \tag{6}$$

where we have introduced K_{min} and K_{max} as the total wavenumbers corresponding to the model grid-length and the smallest wavelength cutoff (in VV21 this is 5 km), respectively. Following VV21, the above can be expressed as:

$$\tau_x, \tau_y = \rho_0 N(U_0 F_1 + V_0 F_2, U_0 F_2 + V_0 F_3) \tag{7}$$

where:

$$F_1 = A^{-1} 4\pi^2 \int_{-\pi}^{\pi} \int_{K_{min}}^{K_{max}} K^2 \cos^2\phi |\hat{h}|^2 dK\, d\phi, \tag{8a}$$

$$F_2 = A^{-1} 4\pi^2 \int_{-\pi}^{\pi} \int_{K_{min}}^{K_{max}} K^2 \cos\phi \sin\phi |\hat{h}|^2 dK\, d\phi, \tag{8b}$$

$$F_3 = A^{-1} 4\pi^2 \int_{-\pi}^{\pi} \int_{K_{min}}^{K_{max}} K^2 \sin^2\phi |\hat{h}|^2 dK\, d\phi \tag{8c}$$

These terms can then be separated into discrete wavevector direction bins, so that $F_{1,i}$,$F_{2,i}$ and $F_{3,i}$ become functions of wavevector direction ϕ_i. This leads to an expression for the surface momentum flux as a function of wavevector direction:

$$\tau_{x,i}, \tau_{y,i} = \rho_0 N (U_0 F_{1,i} + V_0 F_{2,i}, U_0 F_{2,i} + V_0 F_{3,i}) \tag{9}$$

where:

$$F_{1,i} = A^{-1} 4\pi^2 \int_{\phi_i - \frac{\Delta\phi}{2}}^{\phi_i + \frac{\Delta\phi}{2}} \int_{K_{min}}^{K_{max}} K^2 \cos^2\phi |\hat{h}|^2 dK\, d\phi, \tag{10a}$$

$$F_{2,i} = A^{-1} 4\pi^2 \int_{\phi_i - \frac{\Delta\phi}{2}}^{\phi_i + \frac{\Delta\phi}{2}} \int_{K_{min}}^{K_{max}} K^2 \cos\phi \sin\phi |\hat{h}|^2 dK\, d\phi, \tag{10b}$$

$$F_{3,i} = A^{-1} 4\pi^2 \int_{\phi_i - \frac{\Delta\phi}{2}}^{\phi_i + \frac{\Delta\phi}{2}} \int_{K_{min}}^{K_{max}} K^2 \sin^2\phi |\hat{h}|^2 dK\, d\phi \tag{10c}$$

The subscript i indicates that the field is representative of a particular wavevector direction where the value of ϕ_i is the middle of the wavevector bin. N_{bins} is the number of bins and $\Delta\phi = \frac{2\pi}{N_{bins}}$ is the bin resolution. $\phi_i - \frac{\Delta\phi}{2}$ and $\phi_i + \frac{\Delta\phi}{2}$ is the range of angles within that bin. The binned fields $F_{1,i}$, $F_{2,i}$ and $F_{3,i}$ are calculated prior to model run-time, since they are a function of the subgrid orography only, and passed to the parametrization at simulation initialisation.

The process for deriving the subgrid orography is illustrated in the Appendix of VV21. Briefly, this involves removing the large-scale mean slope of the orography to reduce aperiodicity of the field, which helps reduce high wavenumber artifacts when a Fourier transform is performed (see Kruse and Smith (2015) for discussion). Then, a two-dimensional discrete Fourier transform of the deplaned gridded latitude-longitude orography within each gridbox is performed and a Gaussian filter is applied in spectral space to damp scales $< 5km$. This produces a two-dimensional field that is a function of zonal and meridional wavenumbers k and l. The expressions inside the integral of equation (10) are then computed, such that we are still in k, l space. The wavevector angle ϕ corresponding to each each $k = K\cos\phi, l = K\sin\phi$ pair is then computed. The two-dimensional arrays of $F_1, F_2, F3$ are then binned (by summation) into one-dimensional fields according to whether or not they fall into the range $\phi_i - \frac{\Delta\phi}{2} < \phi < \phi_i + \frac{\Delta\phi}{2}$ for a particular ϕ_i. As will be discussed later in section 2.3, the binning requires that a sufficient number of points be sampled within each ϕ_i bin, such that there is a limit on the maximum number of bins that can be used, depending on the model grid-length and source data resolution. Note that $F_1 = \Sigma_{i=1}^{N_{bins}} F_{1,i}$, $F_2 = \Sigma_{i=1}^{N_{bins}} F_{2,i}$ and $F_3 = \Sigma_{i=1}^{N_{bins}} F_{3,i}$. The total surface stress is, therefore, given by the sum over all the directional bins and is identical to the values from equation (7), i.e. the scheme described in VV21. As a result, it is only the vertical distribution of this drag that then differs in this extension of the scheme, which is described below.

2.2 | Wave saturation calculation

Vertically propagating gravity waves are subject to critical level filtering as the background winds vary in the vertical. In addition to this, the wave field is subject to convective instability as the background density and stability vary, leading to wave saturation and overturning. Ultimately, the two forms of instability are related, since a wave approaching a critical level will inevitably experience convective or shear instability. For a review of saturation theory see Fritts (1984), in which the onset of convective instability as a wave approaches a critical level is depicted. Analytic solutions and theories for wave saturation are required in every parametrization. Although they do vary somewhat between models, they are based on the same conceptual principles of convective instability (van Niekerk et al., 2020), which is generally accepted as a dominant mechanism for wave saturation in the stratosphere (Fritts, 1984). In order to motivate our choice of saturation method used here, the theoretical background for the current Met Office Unified model parametrization for wave saturation is presented in the Appendix.

Most studies on critical level absorption in directionally sheared flow deal with mountain shapes that have a single descriptive amplitude (such as sinusoidal, e.g. Broad 1995, or bell-shaped hills, e.g. Shutts 1998) and, therefore, can more readily employ convective instability conditions as described in the Appendix. Others deal with Boussinesq fluids (e.g. Doyle and Jiang 2006, Shutts 1998), are not concerned with convective instability criteria (Guarino et al., 2018) and just assume total wave absorption at a critical level. As a result, knowing how to treat critical level filtering and convective instability together in the case of a superposition of waves forced by arbitrary three-dimensional orography is not clear. In order to account for the critical level filtering and convective instability from multiple wavevectors pointing in several directions, we assume that the stability of the individual spectral wave components are independent of each other such that their saturation can be computed independently. The treatment of each direction individually is not strictly consistent with the saturation criterion ($\frac{\partial}{\partial z}\eta > 1$), which is based on the superposition of the waves becoming unstable. It should be emphasised that, because the directional binning is the integral over many total wavenumbers, this method of wave saturation does still represent a superposition of waves to some extent. This assumption may not be valid in a turbulent background with strong vertical gradients, caused by the breaking of a particular wave component, and may violate assumptions of the underlying linear and WKB theory. Nonetheless, these limitations may be justified in the pursuit of progress.

The following describes the method employed in our new parametrization. It is assumed that the vertical displacement of isentropic surfaces in a particular wavevector direction takes the form:

$$\eta_i = \int_{-\pi}^{\pi}\int_{K_{min}}^{K_{max}} \hat{\eta}_i(K)\, exp(iK(\cos\phi_i x + \sin\phi_i y))\, K\, dK\, d\phi, \tag{11}$$

such that there is only one wavevector angle, ϕ_i, representing the wave field. Since we are assuming $\phi = \phi_i$ is a constant, the integral over ϕ can be removed:

$$\eta_i = 2\pi \int_{K_{min}}^{K_{max}} \hat{\eta}_i(K)\, exp(iK(\cos\phi_i x + \sin\phi_i y))\, K\, dK, \tag{12}$$

This is analogous to having only one wavevector, as is described in the Appendix, but here the wave field in a particular direction is the integral over all total wavenumbers (horizontal wavelengths). The total displacement in all directions is the intergal of (12) w.r.t ϕ, $\eta = \int_{-\pi}^{\pi} \eta_i\, d\phi$, but we treat each direction separately here.

Following as in the Appendix, the displacement can be rewritten in terms of the surface orography $\hat{h}$ and large scale flow

variables as:

$$\eta_i = 2\pi \int_{K_{min}}^{K_{max}} \hat{h}_i(K) \left(\frac{\rho_0 N_0 U_{0,i}}{\rho N U_i}\right)^{1/2} \exp\left(iRe\left(\int_0^z m_i(z')dz'\right)\right) exp(iK(\cos\phi_i x + \sin\phi_i y))KdK, \quad (13)$$

where $U_i = U\cos\phi_i + V\sin\phi_i$ is the wind in the direction of the wavevector bin and $m_i(z')$ is the corresponding vertical wavenumber. Again, this is only valid below a critical level. Rearranging gives:

$$\eta_i = 2\pi \left(\frac{\rho_0 N_0 U_{0,i}}{\rho N U_i}\right)^{1/2} \exp\left(iRe\left(\int_0^z m_i(z')dz'\right)\right) \int_{K_{min}}^{K_{max}} \hat{h}_i(K) exp(iK(\cos\phi_i x + \sin\phi_i y))KdK, \quad (14)$$

Note that the integral over total wavenumber in the above equation is independent of height and is just the surface elevation in a particular direction.

Taking the derivative of η_i w.r.t height and its modulus (see Appendix), the saturation criterion for a particular wavevector direction then becomes:

$$|\eta_i| > \left|\frac{U_i}{N}\right| \quad (15)$$

Similarly, the amplitude growth with height is given by:

$$|\eta_i| = \left|\eta_i(z-\Delta z)\left(\frac{\rho(z-\Delta z)N(z-\Delta z)U_i(z-\Delta z)}{\rho(z)N(z)U_i(z)}\right)^{1/2}\right| \quad (16)$$

where Δz is the vertical grid spacing, such that $\eta_i(z-\Delta z)$ is the amplitude at the previous height. As in the current monochromatic scheme, the stress variation with height is:

$$|\tau_i|(z) = |\tau_i|(z-\Delta z)\left(\frac{\eta_i(z)}{\eta_i(z-\Delta z)}\right)^2 \left(\frac{\rho(z)N(z)U_i(z)}{\rho(z-\Delta z)N(z-\Delta z)U_i(z-\Delta z)}\right) \quad (17)$$

We assume that $(\tau_{x,i}, \tau_{y,i}) \parallel (\cos\phi_i, \sin\phi_i)$, such that the wavevector is in the direction of the surface stress for a particular bin and we can use these surface stresses to compute the winds in the direction of the wavevector bins (i.e. U_i). To compute the total forcing on the mean flow, we then sum the contributions from each wavevector bin as such:

$$\frac{d(U,V)}{dt}_{GWD} = -\frac{1}{\rho}\frac{d}{dz}\Sigma_{i=1}^{i=N_{bins}}(\tau_{x,i}, \tau_{y,i}) \quad (18)$$

From equation (15), one sees that the wave saturation criterion will be satisfied if the wind parallel to the wavevector becomes zero and, from equations (17) and (18), that this will lead to forcing on the mean flow along that wavevector direction.

To summarise, we maintain the current saturation method to a large extent but now compute this for every wave vector bin independently using the launch amplitude (described in the following section) and winds in each wavevector direction. This includes the assumption that there is total absorption of the wave if it passes through a critical level (Shutts, 1995; Shutts and Gadian, 1999), as is valid for high Richardson number flow (Teixeira and Miranda, 2009). What is more, all the wave components will saturate at the same height in the presence of a total critical level.

2.3 | Launch amplitude

The launch amplitude is now not just a single monochromatic amplitude but, rather, is the absolute amplitude of the superposition of the waves of different scales K in a given direction (see equation (14)). To bin the amplitude into wavevector directions, the amplitude at the surface, which we refer to as the launch amplitude, should be representative of the orography in that directional bin. This is given by the expression:

$$|\eta_{0,i}| = \left| \int_{\phi_i - \frac{\Delta\phi}{2}}^{\phi_i + \frac{\Delta\phi}{2}} \int_{K_{min}}^{K_{max}} \hat{h} \exp(iK(\cos\phi x + \sin\phi y)) K dK d\phi \right| \tag{19}$$

However, since this expression represents the actual wave field varying in the x and y directions, we need to replace it with something that represents the mean wave amplitude in a given wavevector bin. It is not entirely clear how to set the amplitude for a given direction, especially in the presence of flow blocking.

It is well documented from fluid laboratory experiments, observations and numerical simulations that mountain wave amplitudes are modulated by the near-surface Froude number $F_r = \frac{|U_0|}{N_0 h}$ (see Baines 1987 for an overview). When F_r is small, the winds are too weak or the stability is too high for the flow to go over the mountain, leading to orographic flow-blocking. As F_r increases, the flow more readily traverses the mountain and flow blocking reduces. To account for this, many current orographic gravity wave drag schemes set the launch amplitude to an effective mountain height given by $h_{eff} = min(h, \frac{|U|_{blk}}{N_0 F_c})$, where $h = n\sigma$, with n being a tuning constant and σ being the standard deviation of the orography within a gridbox. $|U|_{blk}$ is the near surface wind in the direction of the major axis of the mountain and F_c is a tuning constant. This single launch amplitude is suitable for a scheme with only one wavevector direction but may not be representative of the wave field in several different directions. As a result, we explore two options for setting the launch amplitude, listed below.

The most practical and readily available method of setting the launch amplitude in the different directions is to set them all to $\eta_{eff,i} = h_{eff} = min(h, \frac{|U|_{blk}}{N_0 F_c})$, which is the effective mountain height that is able to generate a wave. This effective mountain height is computed from the winds in the direction of the major axis of the topography used within the flow blocking parametrization (U_{blk}), and represents a mean amplitude of a wave produced by flow over a single elliptical subgrid mountain.

Another possible choice is to use a representative mean orographic amplitude for a given directional bin:

$$\sigma_i = \left(4\pi^2 A^{-1} \int_{\phi_i - \frac{\Delta\phi}{2}}^{\phi_i + \frac{\Delta\phi}{2}} \int_{K_{min}}^{K_{max}} |\hat{h}(K,\phi)|^2 K dK d\phi \right)^{1/2} \tag{20}$$

This is the standard deviation in a wavevector direction.

During testing of this method, a sampling bias in the number of points within each directional bin was uncovered that resulted from the fact that the orographic source data is discrete and is not in polar coordinates. This meant that there were more points being integrated over directions that were around ~45 degrees from the $x-$axis. The sampling bias led to larger amplitudes in certain wavevector direction bins, since these directions had more data samples and the amplitudes are given as an intergral over the data points. The effect of this was to cause saturation of those wavevectors at lower altitudes. What is more, this sampling bias was exacerbated at higher model resolutions where the number of sub-grid data points within a grid box, and hence within a direction bin, could become very small. To remove this bias, the amplitude is normalised by the ratio of the total number of data points in the model gridbox ($NxNy$) and the number of points within a directional bin (N_{ϕ_i}). This has the effect of giving the same weighting to each bin, independent of the number of points that fall within that wavevector angle range, while maintaining correct magnitudes.

$$\sigma_{dir,i} = \left(4\pi^2 A^{-1} \int_{\phi_{\phi i} - \frac{\Delta\phi}{2}}^{\phi_{\phi i} + \frac{\Delta\phi}{2}} \int_{K_{min}}^{K_{max}} |\hat{h}(K,\phi)|^2 K \, dK \, d\phi \, \frac{Nx Ny}{N_{\phi_i}} \right)^{1/2} \tag{21}$$

where Nx, Ny are the number of discrete data points within a gridbox and N_{ϕ_i} is the number of data points falling within the wavevector bin. The mountain amplitude is then multiplied by some tuning factor n to get the launch amplitude $\eta_{dir,i} = n\sigma_{dir,i}$.

To account for flow blocking, we also multiply this launch amplitude by an empirically derived factor $\beta = \frac{1}{2}\left(1 - \tanh\left(\frac{\frac{z_{blk}}{h_{amp}} - A}{B}\right)\right)$, which reduces when there is flow blocking. Here, z_{blk} is the depth of the blocked layer near the surface, computed by the flow blocking component of the parametrization scheme, h_{amp} is the amplitude of the maximum orography within the model gridbox and $A = 0.42$ and $B = 0.18$ are empirically derived constants. See VV21 for the derivation and motivation behind this empirically derived factor β.

3 | TESTING METHODOLOGY

3.1 | Idealised model setup

High resolution idealised simulations performed with the Met Office Unified Model are used to help guide decisions on design choice and to validate the methods used in this extension to the scale-aware scheme. We aim to answer the following questions using these simulations:

1. How does the near-surface wave amplitude vary with wavevector direction?
2. How is the wave amplitude affected by flow blocking, and do the methods in section 2.3 represent this?
3. How does the gravity wave momentum flux vary with height and wavevector direction?
4. How does the momentum flux from the binned saturation scheme, described above, compare with the idealised resolved flux?
5. How does the number of wavevector direction bins affect the behaviour of the parametrization scheme?

The idealised simulations are performed using the Met Office Unified model, which is fully nonhydrostatic, nonlinear and deep (Wood et al., 2014), at version 12.2. A Smagorinsky-type 3D turbulence scheme, which is blended with the boundary layer turbulence scheme depending on grid-length, is used (Boutle et al., 2014). The surface heat and momentum fluxes are specified to zero.

The horizontal grid-spacing is 500m, using Cartesian coordinates. The simulation domain is bi-cyclic in the horizontal and extends 375km by 250km by 78 km (with 118 levels) in the x, y and z directions, respectively. There is an absorbing layer applied near the top of the model, to prevent spurious wave reflection, which starts at approximately 60 km altitude. The surface orography is taken from a section over the centre of the Alps, since this orography is very multi-scaled and that region does frequently experience directional wind shear (Jiang and Doyle, 2004). The orography is placed at the centre of the domain and is generated from 300m Shuttle Radar Topography Mission (SRTM;Van Zyl 2001) data using the following procedure: the orography is interpolated onto the model grid; a linear-best-fit plane is removed; the minimum value is added on to remove negative regions; this is then multiplied by a Gaussian function $\exp(-((\frac{x}{a})^2 + (\frac{y}{a})^2))$ where $a = 20km$; and it is filtered to remove scales larger than 100km. Finally, the orography is rescaled by dividing by the maximum height and multiplying by 1000m. The reason for the processing is to ensure that the orographic height tends to zero at the edges of the domain and that the

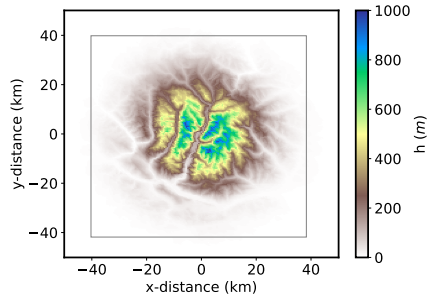

FIGURE 1 The rescaled surface altitude, taken from SRTM data over the European Alps, in a subset of the idealised domain. The grey box indicates the area A, where the orographic height is more than 10m, used in computation of diagnostics from the idealised simulations.

scale of the topography is approximately 40 km in the horizontal and 1000m in the vertical. A zoomed in region of the orography used in the idealised simulations is shown in Figure 1.

The idealised simulations are initialised using a constant Brunt-Vaisala Frequency of $N = 0.03s^{-1}$, from which a hydrostatically balanced pressure and density are computed, resulting in a decreasing density profile. The initial wind speed and direction does not vary over the depth of the idealised orography, so that it is in-keeping with the assumptions made in the calculation of the parametrized surface stress. The Froude number, $Fr = \frac{|U|}{hN}$, computed using $h = 1000m$, is varied by changing the value of the near-surface wind speed, so that the effect of flow blocking on the wave field can be investigated. To understand the impact of directional critical level filtering on the momentum flux, the initial background wind speed, direction and vertical profile above the mountain is also varied in one of the simulations. Simulations are then run for 5 hours. The three cases considered are listed below.

- **Case A: constant wind profile,** $F_r = 1$
  In this case, the zonal and meridional winds are initialised to $U = 30ms^{-1}$ and $V = 0.0ms^{-1}$, respectively, and do not vary with height. The Froude number is $F_r = 1$, such that there is little flow blocking. This case will be used to demonstrate how the waves saturate as a result of amplitude growth with height due to decreasing density alone.
- **Case B: constant wind profile,** $F_r = 0.5$
  In this case, the zonal and meridional winds are initialised to $U = 15ms^{-1}$ and $V = 0.0ms^{-1}$, respectively, such that $F_r = 0.5$ and there is more flow blocking near the surface. Comparing this case with results of Case A will help to reveal the impact of flow blocking on wave amplitude.
- **Case C: wind turns 90 degrees,** $Fr = 1$
  This case is also initialised with zonal and meridional winds of $U = 30ms^{-1}$ and $V = 0.0ms^{-1}$, respectively, such that $F_r = 1$ near the surface. However, above the mountain at approximately 2 km, the windvector begins to turn anticlockwise until it is at 90 degrees from the surface winds at 21 km. The winds then remain constant above this. The change in wind with height is linear and is given by:

$$(U,V)(z) = \begin{cases} (30,0) & \text{if } z \leq 2000 \\ (-\epsilon(z-2000)+30, \epsilon(z-2000)) & \text{if } 2000 < z \leq 21000 \\ (0,30) & \text{if } z > 21000 \end{cases} \quad (22)$$

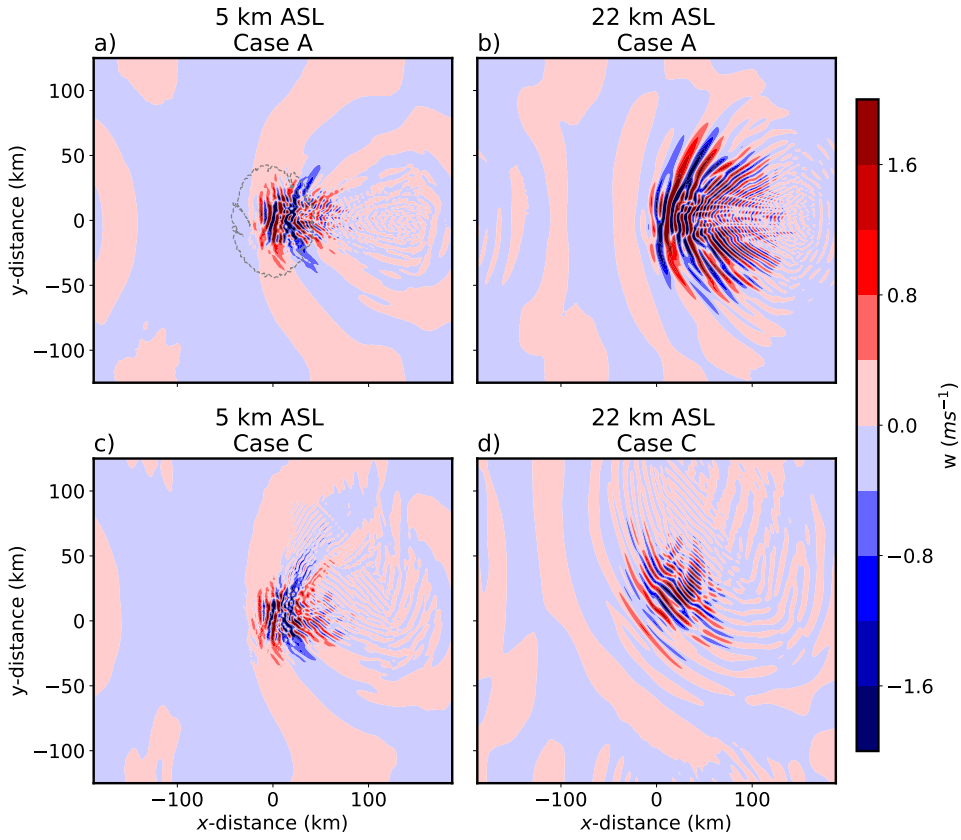

FIGURE 2 Vertical velocity 1.5 hours into the idealised simulations for (a) Case A at 5 km altitude, (b) Case A at 22 km altitude, (c) Case C at 5km altitude and (d) Case C at 22km altitude. The 5 m contour of the orographic height is shown in grey in (a).

where $\epsilon = \frac{30}{21000-2000} \sim 0.00157$ is the windshear. There will be directional critical level filtering for some part of the wave field in this case. Since some of the waves may not encounter critical levels but will experience amplitude growth due to decreasing density, this case will be used to understand both the validity of the linear theory, which states that the stress goes to zero as $(Uk + Vl) \rightarrow 0$, and the amplitude saturation criterion.

To illustrate the concept of partial critical level filtering, Figure 2 shows the vertical velocity at 5km and 22km altitude for the case without wind turning (Case A) and for the case with 90 degree wind turning (Case C). Close to the mountain top (5km altitude) there is clearly a full spectrum of waves in many different directions, with little difference between case A and C. In Case A at 22 km altitude these waves have increased in amplitude and the many wavevector directions and scales become more evident. Comparing Case A and Case C, in which the wind has turned 90 degrees anti-clockwise, one can see that some parts of the wavespectum has been removed in case C at 22 km. This is a result of the fact that those waves with wavevectors that are perpendicular to the winds have encountered critical levels throughout the atmosphere. More importantly, some part of the wave field is still present in Case C and will continue to propagate vertically and eventually deposit momentum. It is this effect that we are trying to account for in the parametrization. In the current parametrization, where there is only one wavevector direction, the

scheme would predict complete saturation of the wave field at this height, since the winds in that wavevector direction would have gone to zero.

3.2 | Diagnostics

To determine which method to use for the launch amplitude, the idealised simulations are performed under varying Froude number conditions and the amplitude above the mountain top is compared with output from the parametrization scheme. The mean amplitude at the launching level (taken to be 2 km) in particular wavevector bins is computed from the idealised simulations using:

$$\eta_{RES,i} = n\left[4\pi^2 A^{-1}\int_{\phi_i-\frac{\Delta\phi}{2}}^{\phi_i+\frac{\Delta\phi}{2}}\int_0^{K_{max}}|\hat{\eta}|^2 K\,dK\,d\phi\,\frac{NxNy}{N_{\phi_i}}\right]^{1/2} \tag{23}$$

where $|\hat{\eta}| = |\frac{\hat{w}}{(Uk+Vl)}|$ is the amplitude computed from the vertical velocity and n is the same tuning constant used in computing the parametrized launch amplitude. The part of the spectrum for which $(Uk + Vl) \rightarrow 0$ is removed from the integral. As is done for the parametrized launch amplitudes, the resolved amplitudes are normalised by $\frac{NxNy}{N_{\phi_i}}$ so that the amplitudes are not biased by the number of data points that fall within a certain wavevector angle bin range.

The parametrized launch amplitude is computed from the idealised orography fed into an offline version of the parametrization scheme, which uses the various methods of setting the launch amplitude described in section 2.3. Since the idealised domain is relatively large and we are only interested in the region of orography, the area A in equations (21) and (23) is taken to be a box around the central region where the orography is higher than 10m in the idealised domain. In generating the orographic inputs for the global model forecast simulations, this area is taken to be the grid-box area.

The saturation method described in section 2 is also validated using the idealised simulations. This is done by computing the idealised resolved momentum fluxes in spectral space, binning these into wavevector directions and then studying the variation of the stress in these different directions with height. To do this, we first compute the mean background density ρ and Fourier transforms of the three wind components, $\hat{u}$, $\hat{v}$ and $\hat{w}$ on constant geometric height surfaces. These are then used to compute the resolved momentum fluxes in wavevector direction bins, analogous to parametrized momentum fluxes of equation (9):

$$\tau_{xRES,i}, \tau_{yRES,i}(z) = A^{-1}4\pi^2\rho(z)\int_{\phi_i-\frac{\Delta\phi}{2}}^{\phi_i+\frac{\Delta\phi}{2}}\int_0^{K_{max}}(\hat{u},\hat{v})\,\hat{w}^*\,K\,dK\,d\phi \tag{24}$$

where $\hat{(...)}^*$ is the complex conjugate of the Fourier transform. The total momentum flux and the forcing of the mean flow (i.e. $\frac{d}{dt}(U,V)_{GWD} = -\frac{1}{\rho}\frac{d}{dz}(\tau_{xRES},\tau_{yRES})$) is then also compared with an offline version of the parametrization scheme, forced by the idealised initial profiles. The parametrization scheme removes scales smaller than $\sim$ 5 km by applying a response function in spectral space to the subgrid orography (see VV21 for details). To make sure the parametrized and resolved wave properties are being compared consistently, the resolved quantities (i.e. $\hat{w}$, $\hat{u}$ and $\hat{v}$) are also multiplied by this response function.

Since the idealised simulations are transient and the parametrization assumes steady-state, we should ensure that the resolved fluxes are analysed at a point during the simulation where the wave propagation and any wake generation has not yet been able to significantly impact the mean flow, but that it has propagated far enough in the vertical, i.e. the model top. The time taken for the wavepacket to reach a particular height can be estimated using the hydrostatic vertical group velocity for stationary waves $\frac{(Uk+Vl)^2}{NK}$ (Kruse and Smith, 2018). For the case where the winds are $30ms^{-1}$, $N = 0.03s^{-1}$ and for a horizontal wavelength of

$40km$ the vertical group velocity is $4.7ms^{-1}$. This means that the time taken to reach $60km$ is approximately 3.5 hours. Note, however, that this is the timescale for the longest waves, which travel the slowest, to reach the model top and so is an upper bound on the propagation time. The time it takes for the flow to traverse the mountain is given by the mountain advection time a/U (Wells et al., 2008), where a is the mountain half-width $20km$, which is approximately 10 mins (20 mins) for Case A (Case B). The time taken for the wake to reach the edge of the domain in the x-direction is approximately 1.7 hours for Case A. As a result, an average of the diagnostics from the idealised simulations are taken between 1.5 - 2.5 hours into the simulations.

3.3 | Global model simulations

We additionally perform short range global numerical weather prediction case studies at a grid-spacing of 130 km (N96) and 40 km (N320) with the binned extension to the scale-aware scheme included, so as to gauge the impact of the scheme on some key circulation metrics. The case studies consist of a set of 5-day forecasts initialised from European Centre for Medium-Range Weather Forecasts (ECMWF) analysis of 24 separate dates distributed throughout the annual cycle over the years 2013 and 2014. These 24 case study dates are those used routinely during the model development process of the Met Office Unified Model (Williams et al., 2020), and sample a variety of weather patterns and seasons over which we have suitable observations for validation. The circulation impacts from these case studies are assessed at a lead-time of 5 days and performance is taken relative to the initialising ECMWF analysis. All simulations are performed using version 12.1 of the Met Office Unified model with the Global Atmosphere 8/ Global Land 9 (GA8/GL9) science configuration. Case studies are performed with 70 vertical levels and a model top at 80 km.

4 | RESULTS

4.1 | Idealised simulation validation

Using the idealised simulations, we first aim to constrain the method for setting the launch amplitude of the waves in binned wavevector directions. Figure 3(a) shows the resolved amplitude ($\eta_{RES,i}$; from eq. (23)) at 2km altitude for Case A with $Fr = 1$ (no flow blocking) and the parametrized launch amplitude ($\eta_{dir,i} = n\sigma_i$; from eq. (21)) binned into 8 wavevector directions. Here, the radius of the bar represents the amplitude of the wave and the angle is the middle of the wavevector direction bin (i.e. ϕ_i). From this plot we can see that the resolved wave amplitude has power in all directions, although some directions have slightly larger amplitudes than others, indicating anisotropy in the wave field. As an aside, a similar plot produced from simulations using an elliptical mountain shows distinctly larger amplitudes in the direction of the major axis of the ellipse, indicating that the Alps orography (Figure 1) truly is multi-directional and complex when compared with a simple ellipse. From panel (a) of Figure 3 alone, we can already answer question 1 posed in section 3.1: How does the near-surface wave amplitude vary with wavevector direction? While the parametrized launch amplitude is slightly larger than the resolved amplitudes, likely due to variation in the background wind and sheltering effects, the resolved wave amplitudes are strongly correlated with the subgrid mountain amplitudes ($n\sigma_i$) in the wavevector directions. The correlation (r) between the resolved and parametrized amplitudes using 32 bins is presented at the top of Figure 3, and demonstrates that the use of $\sigma_{dir,i}$ is an appropriate approach to represent the anisotropy of the launch amplitudes in different directions.

We now investigate question 2 posed in section 3.1: How is the wave amplitude affected by flow blocking, and do the methods in section 2.3 represent this? Figure 3(b),(c) shows the resolved wave amplitude for the idealised Case B with $Fr = 0.5$, along with the parametrized amplitudes computed using the two methods in section 2.3. Note the scale is different between panel (a) with $F_r = 1$ and (b),(c) with $F_r = 0.5$. As is consistent with our current paradigm of orographic flow blocking, the resolved mountain wave amplitude reduces when Fr is reduced, and is approximately halved. Comparing the different methods

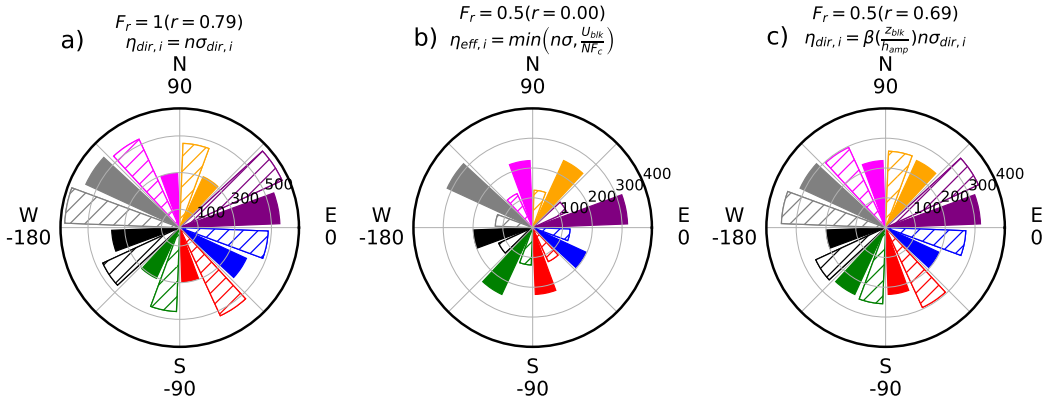

FIGURE 3 (a) Resolved wave amplitude ($\eta_{RES,i}$ solid bars), calculated using eq. (23) from idealised simulations with $F_r = 1$ and the orographic standard deviation ($n\sigma_{dir,i}$ hatched bars), computed using eq. (20); (b) resolved wave amplitude ($\eta_{RES,i}$ solid bars) and the parametrized wave amplitude, computed as $\eta_{eff,i} = \min(n\sigma, \frac{U_{blk}}{NF_c})$ (hatched bars), from idealised simulations with $F_r = 0.5$; (c) resolved wave amplitude ($\eta_{RES,i}$ solid bars) and the parametrized wave amplitude, computed as $\eta_{dir,i} = \beta n\sigma_i$ (hatched bars), from idealised simulations with $F_r = 0.5$. The radius of the bar indicates the amplitude (in m) and the angle represents the wavevector direction range. The r value quoted in brackets is the correlation between the idealised amplitude and the parametrized amplitude when 32 bins are used.

of accounting for flow blocking, it is clear that the bulk effective mountain height ($\eta_{eff,i} = min(n\sigma, \frac{U_{blk}}{NF_c})$: Figure 3(b)), does not capture the anisotropy (or magnitude) of the wave amplitudes in different directions. In contrast, parametrizing the launch amplitude using $\eta_{dir,i} = \beta(z_{blk}/h_{amp}) n\sigma_i$ (Figure 3(c)) shows a remarkable agreement between the resolved and parametrized, in terms of both amplitude and anisotropy. An important comment to make is that, even with tuning of the uncertain parameters, such as F_c and n, to match the parametrized amplitude to the idealised amplitudes, the options $\eta_{eff,i} = min(n\sigma, \frac{U_{blk}}{NF_c})$ would still not produce the correct anisotropy. This is because those tuning parameters only act to switch off the flow blocking or vary the amplitude of $\eta_{eff,i}$, without changing the directional dependence. In the rest of this study, we discount the methods of setting the launch amplitude $\eta_{eff,i}$ and focus further comparison between the idealised simulations and the parametrization using only the method $\eta_{dir,i} = \beta(z_{blk}/h_{amp}) n\sigma_i$.

The next step is to understand the directional dependence of the saturation of the waves and the resultant momentum flux variations with height, as they break due to either convective instability (due to decreasing density) or critical level filtering. Before testing the case with wind turning, we investigate the wave saturation method for a case where there is only decreasing density and, therefore, only saturation due to amplitude growth. Since, in reality, this amplitude growth with height will be occurring simultaneously with the critical level filtering as the wind turns, its directional dependence should be understood in isolation. Figure 4 shows the variation of the resolved and parametrized zonal momentum flux with height for case A. The momentum flux is shown for each bin and, again, we initially only consider 8 bins for visualisation. From the idealised resolved momentum fluxes, we can answer question 3: How does the gravity wave momentum flux vary with height and wavevector direction? As mentioned in section 2.3, the wave saturation with height occurs as a result of convective instability from a superposition of wave modes and it is not given that these waves should saturate independently in wavevector directions, as is assumed by our method. We see that the resolved fluxes do vary differently in each direction, since the rate of momentum flux decrease is more rapid for some directions. A likely reason for this is that the saturation criterion $\eta_i(z) > F_{sat} \frac{U_i}{N}$ is more easily satisfied for those wave directions with winds not aligned with the surface winds (i.e. U_i small) and for those with large launching amplitudes (i.e. η_i large). For example, the bins with $\phi_i = 112.5°, 67.5°$ begin to decrease lower down because the

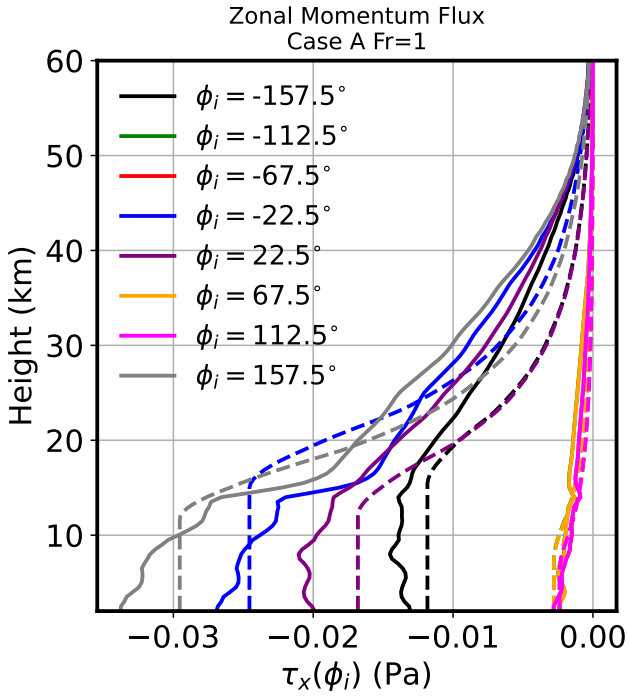

FIGURE 4 Zonal momentum fluxes in 8 wavevector direction bins computed from idealised simulations of Case A using resolved fluxes from equation (24) (solid) and the parametrization (dashed). The legend indicates the middle wavevector direction of the bin. Note that the components $\phi_i = -112.5$ and $\phi_i = -67.5$ are not visible because they are small and overlap with $\phi_i = 112.5$ and $\phi_i = 67.5$.

wind in those directions is small and so they saturate more readily. This suggests that treating the bins individually, even when the waves are breaking only due to changes in density, is plausible.

From Figure 4, we can address question 4: How does the momentum flux from the binned saturation scheme compare with the idealised resolved flux? The magnitude of the parametrized surface fluxes across the 8 wavevector direction bins are remarkably similar to their resolved counterparts. As mentioned in the introduction, many wave processes have been neglected, which could also explain some of the differences seen here. However, as is demonstrated in Scinocca and Sutherland (2010) who were able to show that self-acceleration can be accounted for through tuning of the parameter F_{sat}, some tuning of uncertain parameters may lead to even better agreement. While the variation of the momentum flux with height is similar between the resolved and parametrized when 8 bins are used, the parametrized momentum flux does decrease more rapidly than the resolved at altitudes above ~ 30 km. This becomes more evident when the total flux and forcing summed over the bins is compared with the resolved in Figure 5.

Figure 5 aims to address question 5: How does the number of wavevector direction bins affect the behaviour of the parametrization scheme? The zonal flux and forcing when using $N_{bins} = 1, 8, 16, 32, 64, 128, 256$ in the parametrization are shown. Meridional fluxes are neglected for this case because they are approximately 10 times smaller than the zonal component. With 1 bin, the momentum flux abruptly decreases from about 15 km upwards, causing the forcing to dominate at ~ 20 km. As the number of bins is increased, the forcing, while not entirely systematically, begins to move upwards throughout the stratosphere. This behaviour with increasing bin number can be understood as follows: once the saturation criterion is satisfied

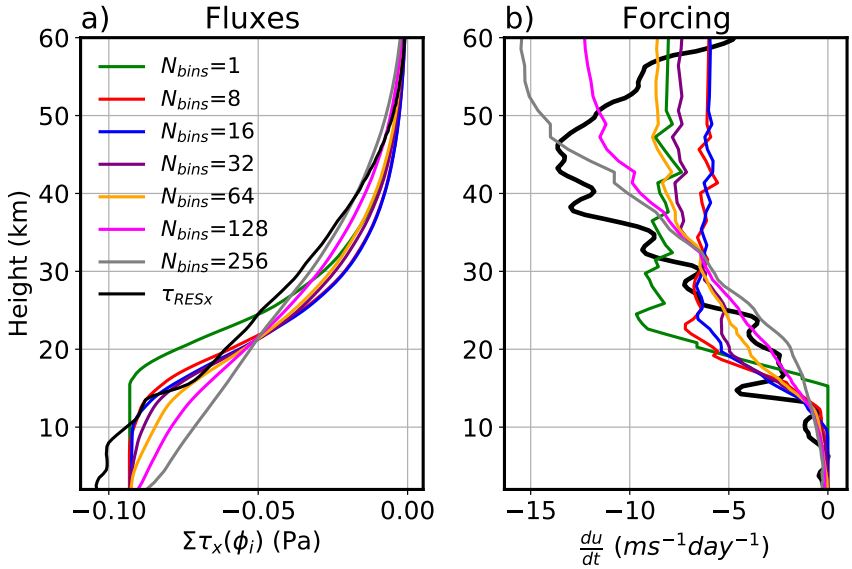

FIGURE 5 Resolved and parametrized zonal momentum (a) fluxes and (b) forcing computed from idealised simulations of Case A. The fluxes and forcing are summed over all bins and use progressively increasing number of bins in the parametrization, as indicated by the legend.

for a particular bin, it will be satisfied at every height above that. Since $\eta_i(z) = F_{sat}\frac{U_i}{N}$ at saturation height, η_i grows according to $\eta(z) = \eta(z - \Delta z)\frac{\rho(z-\Delta z)}{\rho(z)}$ when the winds and stability are constant. This means that, when there is only one bin, the saturation criteria may not be as readily met but once it is the momentum flux can rapidly decrease at a rate determined by the density decrease. On the other hand, when there are several bins, the height at which the saturation criterion is satisfied will be different for each bin. Cumulatively, this leads to a slower saturation with height and more of the momentum flux is able to reach higher altitudes. The behaviour of the parametrization scheme using 128 and 256 bins matches the forcing and flux from the idealised simulations quite well compared with using a smaller number of bins. For this case, the forcing is very sensitive to the number of bins used, particularly in the upper stratosphere. It is likely due to the fact that the launch amplitudes of the different directions decreases as the number of bins is increased, meaning that the waves saturate higher up and have a larger forcing there. While convergence with the number of bins is not yet observed, the need to use such a large number of bins is likely to be dependent on the grid-box size.

We now turn our attention to the case where the winds turn 90 degrees anti-clockwise. Figure 6 shows the binned zonal momentum flux with 8 bins from the idealised simulations and the parametrization for Case C. This case clearly shows how the wavevector bins with directions of $\phi_i = 157.5$ and $\phi_i = -22.5$ saturate more rapidly than those in other directions, in both the idealised simulations and the parametrization. This is because the winds parallel to those components become zero as they turn anticlockwise, and so the waves encounter critical levels causing the momentum flux to go to zero. Those wavevectors that do not encounter critical levels (i.e. $\phi_i = 22.5$ and $\phi_i = -157.5$) are then subject to wave saturation as a result of amplitude growth with decreasing density higher up in the atmosphere. Figure 7 shows the vertical profile of the zonal and meridional momentum flux and wave forcing for the wind turning case using increasingly larger N_{bins}, from 1 to 256. We clearly see that using one directional bin, as is used in the current scheme, leads to all the momentum being deposited between 5km and 15km and none in the stratosphere. This is because the wind parallel to that single wavevector bin becomes zero over this depth. Using

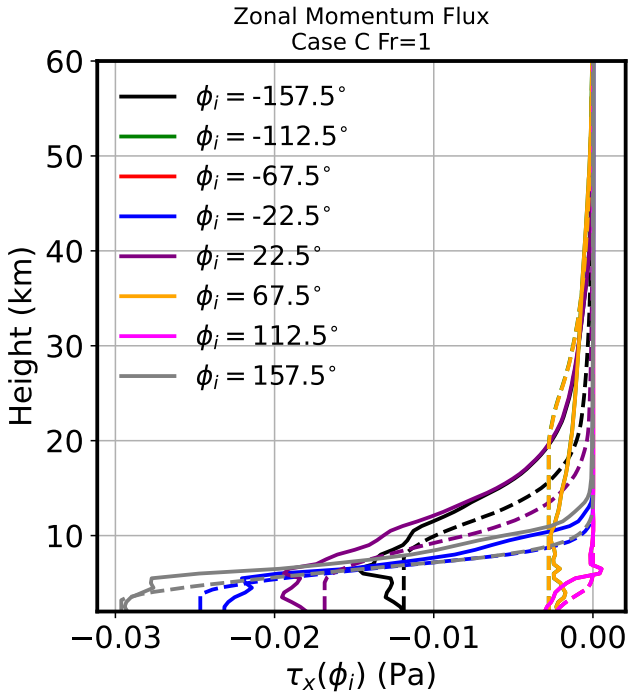

FIGURE 6 As in Figure 4 but for Case C.

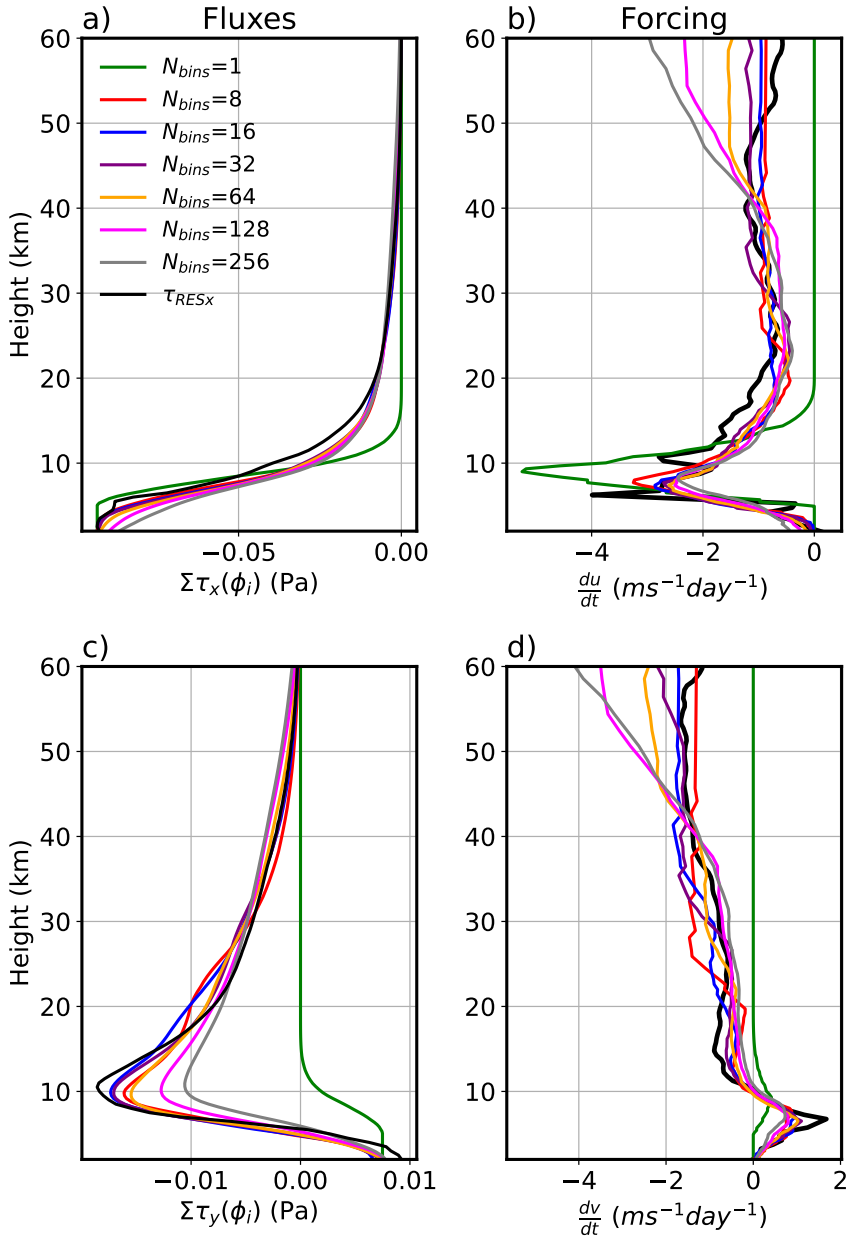

FIGURE 7 Resolved and parametrized zonal and meridional momentum fluxes (a,c) and forcing (b,d) computed from idealised simulations of Case C. The fluxes and forcing are summed over all bins and use progressively increasing number of bins in the parametrization, as indicated by the legend.

an increasingly larger number of bins allows some of the wavevector bins to continue to propagate, reaching higher up and decelerating the flow there. When the number of bins is increased, the zonal forcing between 5km and 20km (the depth over which the wind turns) occurs over a deeper layer and is smaller in magnitude than that with a single bin. What is particularly interesting, however, is the change in the meridional forcing for this case (note that the meridional forcing was small for Case A). When 1 bin is used, the total meridional flux can have only one sign throughout the atmosphere and so leads to a relatively small positive forcing over a small vertical depth. Both the sign, height and amplitude of the forcing and flux with 1 bin does not match the idealised resolved profile well, since the resolved meridional fluxes begin to change sign over the depth of the directional wind shear layer. When the number of bins is increased, the total flux and forcing (summed over the bins) can change sign as directional components of the momentum fluxes encounter critical levels at different height. This leads to much better agreement between the parametrization and the idealised simulations, indicating that the binned scheme is working well for this case. One aspect to note, however, is that increasing the number of bins to 128 or 256, in both Cases A and C, does lead to a larger zonal and meridional forcing above approximately 40 km compared with the resolved idealised forcing. There are several reasons why this might be the case: the damping applied near the top of the model may be acting to reduce the amplitude of the resolved waves in the stratopshere; the parametrization requires tuning of, for example, F_{sat} to get better agreement with the idealised simulations; the transient nature of the idealised simulations is such that the waves acting to decelerate the top of the model have not propagated far enough or the mean state has been altered throughout the simulation, making direct comparison difficult; the sampling bias described in section 2.3 could be worsened as the number of bins increase, leading to larger disagreement between the resolved and parametrized wave amplitudes and, thus, the saturation. Either way, we would argue that the parametrization with 16 or 32 bins performs best in the idealised setup.

4.2 | Global model impacts

The aim of this section is to demonstrate the impact of accounting for partial critical level filtering on the large-scale circulation of the Met Office Unified Model. What is more, the results from the idealised simulations will help us to understand the large scale impacts within the model. The idealised sensitivity experiments indicate that increasing the number of directional bins allows the scheme to increasingly captures the effects of partial critical level filtering of orographic gravity waves. However, the optimal choice for the number of directional bins in a practical application such as global NWP is not clear. We take a pragmatic approach here and do not consider a number of bins greater than 32, in order to avoid significant increases in computation time and memory usage. Estimated increases in memory and wallclock time during model runtime are quantified in the conclusions.

We begin by looking at the change in the zonal mean zonal wind forcing from the parametrization 24 hours into our initialised forecast experiments, averaged over all the different start dates. Simulations where the number of wavevector bins is increased from 1 to 32 (Figure 8) are considered first. At this lead time the large scale mean flow has not yet diverged too much from the initial conditions and statistics remain robust. As was found in the idealised simulations, the forcing in the troposphere (bottom panels) is reduced while the forcing in the upper stratosphere (top panels) is increased (more negative) when the number of bins is increased. This is also consistent with the findings of Xu et al. (2019), who identified a vertical dipole structure when accounting for critical level filtering of gravity waves generated by elliptical mountains. The magnitude of the changes in the parametrized forcing within the stratosphere are on the same order as the mean forcing with 1 bin, and approximately 50% of the forcing in the troposphere. While the values appear small in the troposphere and large in the stratosphere, due to the decrease in density, the relative changes in both these regions are not negligible.

To identify which geographic regions these changes are occurring over, Figure 9 shows maps of the change in the zonal mean zonal wind forcing at three different heights, focusing on the northern hemisphere wintertime case study dates. At 12 km altitude, the negative forcing is reduced almost everywhere when 32 bins are used, but with the Southern Himalayas and Rockies dominating the decrease. This decrease in the forcing at lower altitudes is observed as a result of the fact that, while

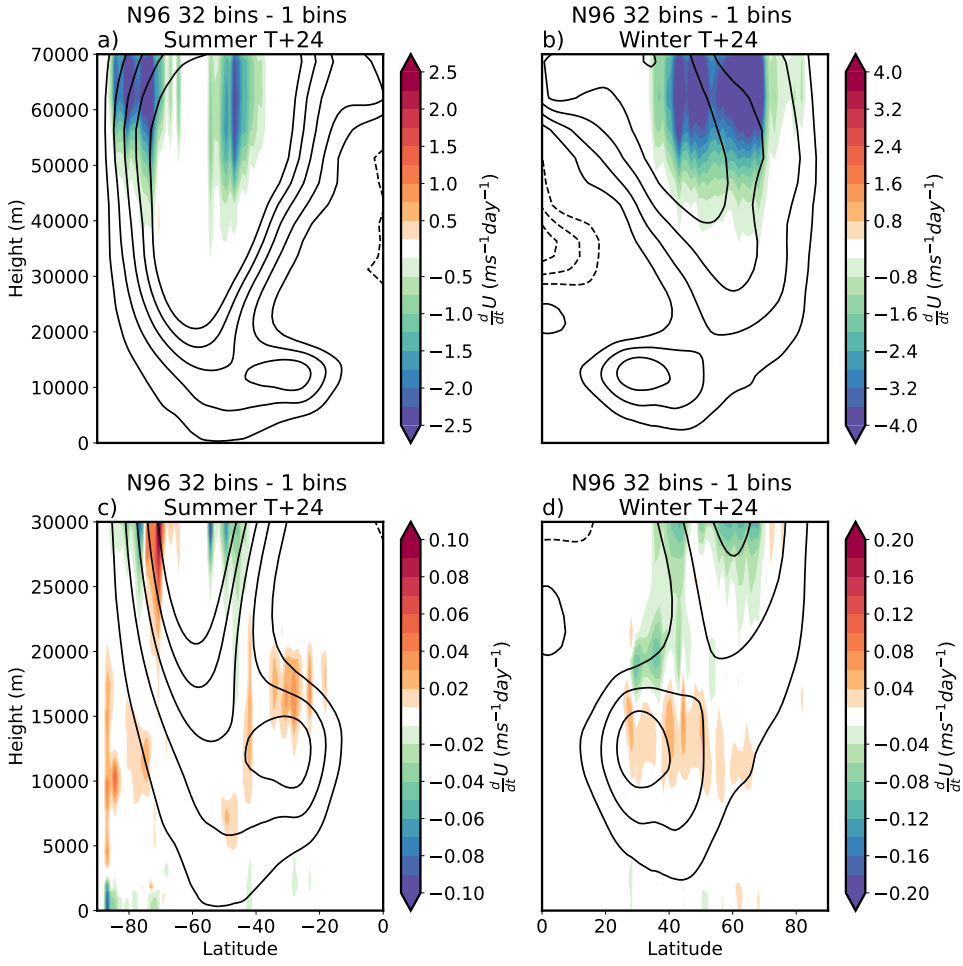

FIGURE 8 Change in zonal mean zonal wind forcing from parametrized orographic gravity wave drag ($\frac{\partial}{\partial t}U_{GWD}$) when increasing the number of wavevector direction bins from 1 bin to 32. (a) and (b) are focused on the stratosphere, (c) and (d) are focused on the troposphere. Differences are shown 24 hours into the initialised forecasts, averaged over Northern Hemisphere summer dates between 90S to 0S (a, c) and Northern Hemisphere winter dates between 0N and 90N (b, d). Colours indicate the change in the forcing and black line contours are the zonal mean zonal winds from the simulation with 1 bin using a $10ms^{-1}$ interval.

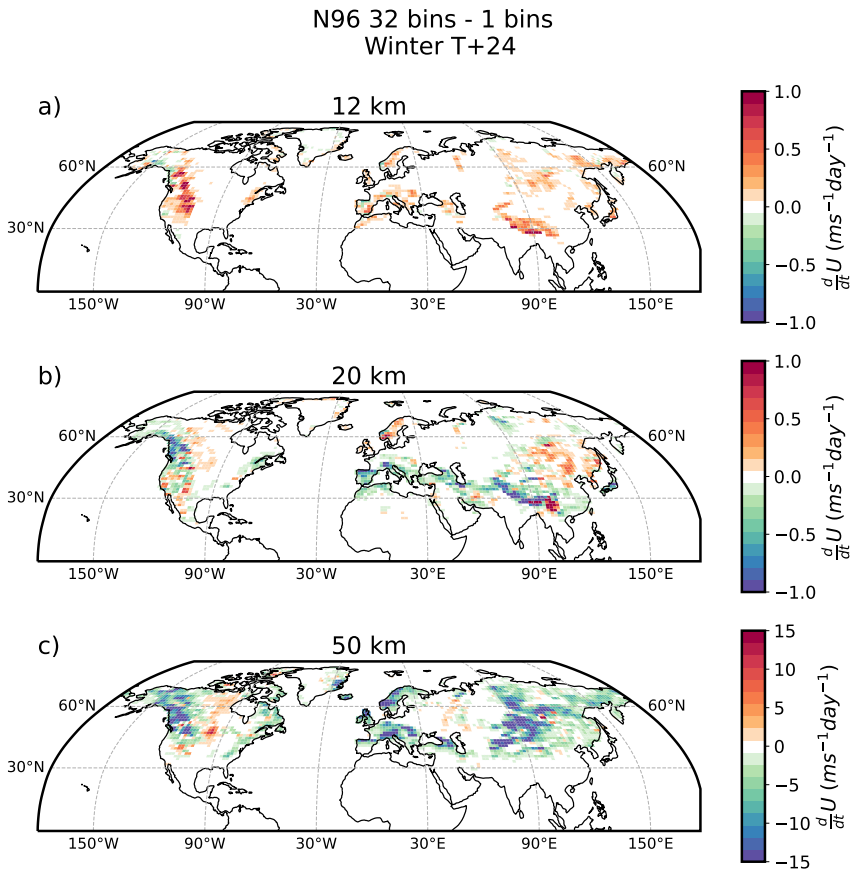

FIGURE 9 Change in zonal mean zonal wind forcing from parametrized orographic gravity wave drag ($\frac{\partial}{\partial t}U_{GWD}$) when increasing the number of wavevector direction bins from 1 bin to 32 in the N96 simulations. Altitudes of (a) 12 km, (b) 20 km and (c) 50 km are shown. Differences are shown 24 hours into the initialised forecasts, averaged over all Northern Hemisphere wintertime dates.

some of the momentum flux from particular bins may saturate at ~ 12 km altitude, some portion of the wave field from the other wavevector directions would not have saturated, allowing more momentum to leave the troposphere. This is similar to what is observed from the idealised simulations. Moving up through the atmosphere to 20 km, we see that the drag is now increased over the Southern Himalayas, some parts of the Rockies and Southern Europe. This is due to there being a total critical level for almost all wavevectors at this altitude within the mid-latitudes, resulting in most of the momentum that was able to leave the troposphere being deposited there. Northern regions such as Norway and the Mongolian mountains still exhibit reduced drag at 20 km, indicating significant filtering of the waves throughout the atmosphere over these regions. At 50 km altitude, where the density has decreased sufficiently such that most of the gravity waves would have saturated, there is a large increase in drag over the Mongolian mountains, Norway and the Northern Rockies. Interestingly, there is also an increase in forcing over the European Alps, where deep propagating waves carrying large momentum have been observed (Dörnbrack, 2021). Again, this is consistent with our understanding from the idealised simulations, in which the complex orography and wind profile led to partial critical level filtering. There is little change over the mid-latitudes at 50 km, likely due to the fact that there is generally a total critical level above the mid-latitude jet and, as a result, partial filtering of the wave field does not play a significant role above ~ 30 km there.

An important question then is: does the change in the vertical and horizontal distribution of the gravity wave forcing lead to improved atmospheric circulation? To help answer this, Figure 10 shows the zonal mean zonal wind change and error relative to ECMWF analysis used to initialise the forecasts at a lead time of 5 days when using 1, 8 and 32 bins. While there is a small decrease in the wind errors over the SH stratosphere (not shown), this error reduction is very small compared with the mean error, particularly at 60S, which is indicative of additional missing processes there, such as three-dimensional propagation of waves (Amemiya and Sato, 2016) and transient effects (Bölöni et al., 2021). As a result, we focus on the NH wintertime errors in Figure 10, since these dominate the overall response. There is a reduction in the positive zonal mean zonal wind error between 50N to 75N at altitudes of 40 km to 65 km. There is also a decrease in the negative bias further South, at about 25N to 40N. It is likely that the acceleration towards the south of the negative gravity wave forcing is due to a change in the structure and position of the meridional circulation, which reacts to the changes in gravity wave drag. It is clear that errors still remain in the upper stratosphere, potentially due to other missing gravity wave or radiative processes. While the error in the upper stratosphere is slightly smaller at N320 (Figure 11), the location and structure of this positive wind error is very similar to that at N96. The impact of increasing the number of bins is also similar at N320, namely the deceleration in the upper stratosphere leads to improved zonal winds there.

One distinction between the N320 and N96 simulations with the binned scheme is their sensitivity to increasing the number of bins from 8 to 32. At N96, there is a large difference when the bins are increased, whereas the impact appears almost identical when the same is done at N320. Through inspection of the orographic fields that are input to the parametrization, we are able to explain this behaviour. A gridbox with a grid-length of 40 km (N320), compared with 130 km (N96), will have significantly fewer orographic source data points to bin into. Increasing the number of bins in a gridbox of 40 km leads to little additional information being added, compared with 130 km, resulting in weaker sensitivity. This conclusion is also transferable to the idealised simulations, where strong sensitivity to increasing the number of bins was seen. The sensitivity observed there is likely due to the fact that the idealised domain, treated as one grid-box, had many data points and a lot more information is added as the number of bins is doubled. We, therefore, conclude that the optimal number of bins may vary according to the model grid-length and source data grid-length.

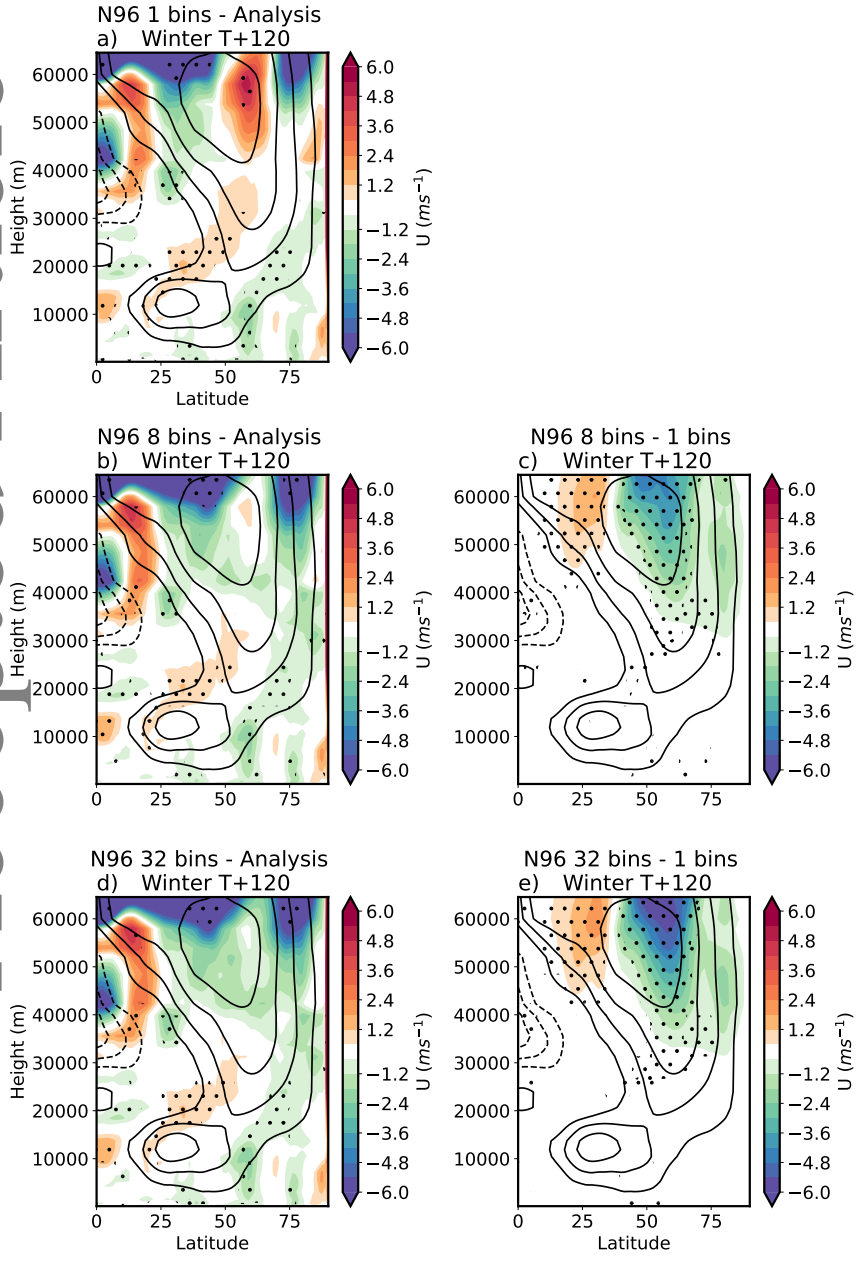

FIGURE 10 Zonal mean zonal wind error relative to the ECMWF initialising analysis, 5 days into the forecasts from N96 simulations using (a) 1 bins (b) 8 bins and (d) 32 bins, averaged over all the Northern Hemisphere wintertime dates. Also shown are the corresponding differences in zonal mean zonal winds 5 days into the forecast when the number of bins is increased from 1 to (c) 8 bins and (d) 32 bins. Colours indicate the difference and black line contours are the zonal mean zonal winds from the simulation with 1 bin using a $10ms^{-1}$ interval. The stippling indicates statistically significant differences, using a paired t-test at a 95% confidence interval.

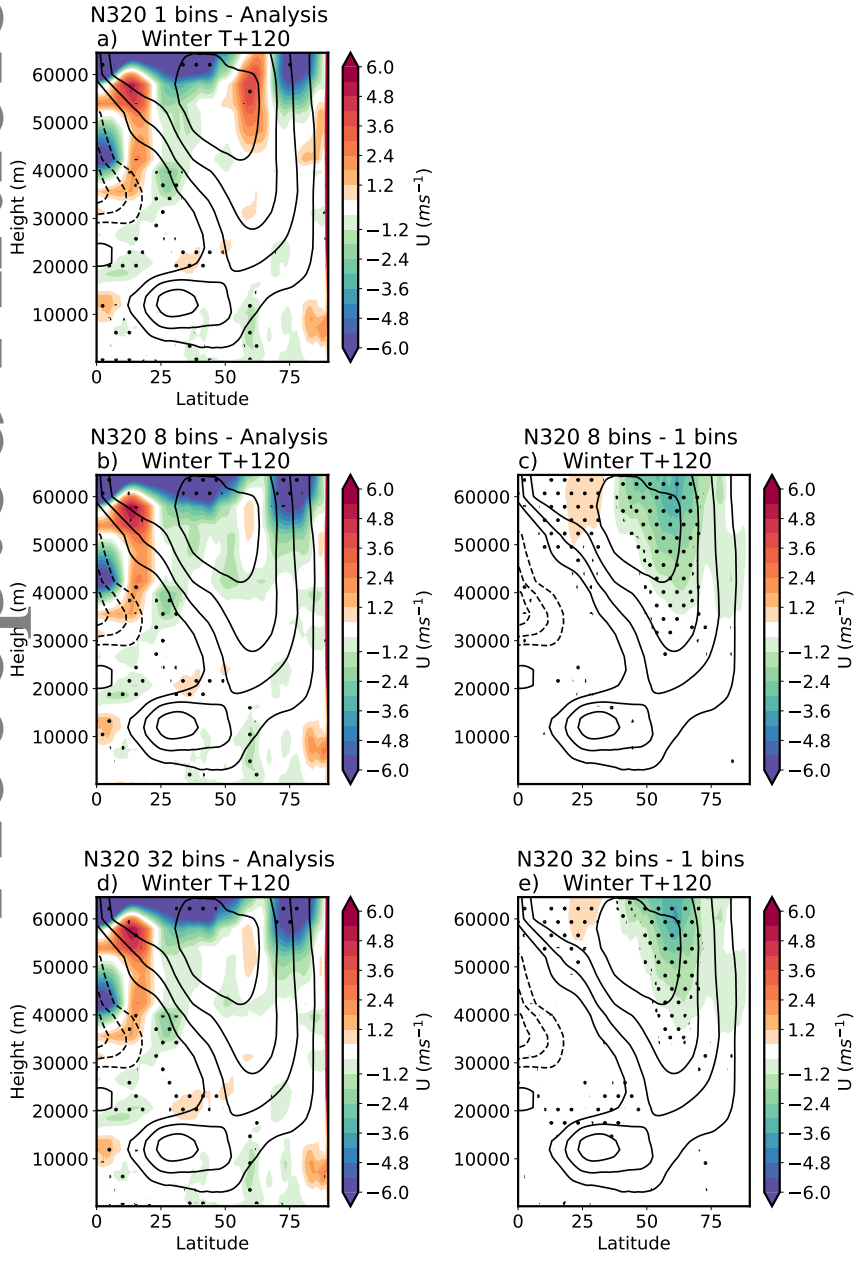

FIGURE 11 As in Figure 10 but at a resolution of N320.

5 | CONCLUSIONS AND DISCUSSION

A method for incorporating the effects of partial critical level filtering in orographic gravity wave drag parametrizations, an extension to the scheme described in VV21, has been presented. This work signifies a move towards a more faithful representation of the three-dimensional nature of unresolved mountain waves in atmospheric models on coarse grids. The method computes the expression for the linearised surface momentum flux for hydrostatic mountain waves using Fourier transforms of the sub-grid orography, and then bins this into wavevector directions. The approach taken is to treat the wavevector directions independently and allow them to saturate according to their launch amplitude, winds parallel to the wavevector direction and static stability variation with height. Since there is no theory for exactly how to define the wave launch amplitudes in different directions when low level flow blocking is present, we make use of idealised simulations over complex orography to provide guidance on the choice of method to define the amplitudes in the parametrization. What is more, the idealised simulations are used to validate the wave saturation method and the sensitivity to the number of wavevector direction bins under different flow regimes. The parametrization is shown to perform relatively well compared with the idealised resolved momentum fluxes, with some convergence with number of bins under certain flow conditions. In particular, it is shown that increasing the number of bins allows more wave forcing to the reach the stratosphere when there is directional wind shear within the troposphere, making the parametrization more consistent with the idealised simulations and known physical behaviour from observations.

Results from implementing the partial critical level filtering parametrization in the Met Office Unified Model are shown for short-range (5 days) forecast simulations at horizontal grid-spacings of ~130 km and ~ 40 km. As was found in other studies that parametrized the effects of critical level filtering (c.f. Xu et al. 2019) and the results from the idealised simulations, the parametrized gravity wave forcing decreases within the troposphere and increases in the stratosphere. Maps of the change in the gravity wave forcing indicate that the additional negative forcing in the upper stratosphere comes from the high northern latitudes, particularly the Mongolian mountains, Northern Rockies and Europe. The impact on the zonal winds of increasing the number of wavevector bins is broadly beneficial at both N96 (130 km grid-spacing) and N320 (40 km grid-spacing) at high altitudes. The response is largest in the upper stratosphere between altitudes of approximately 45 km to 65 km, where increasing the number of bins removes a systematic westerly bias between the latitudes of approximately 50N to 75N, while slightly worsening errors south of this. Since the effects of stratosphere-troposphere coupling take some time to communicate downwards, there is little evident impact on the troposphere from these changes in our relatively short forecast simulations. However, at longer timescales, the increased forcing in the stratosphere is likely to affect the surface more significantly (Sigmond et al., 2008; Kidston et al., 2015). The short-range impacts seen here may lead to changes in the stratosphere-troposphere coupling at both seasonal and climate timescales, a topic which could be investigated in the context of sudden stratospheric warmings and/or stratopsheric teleconnections. This more detailed, process based verification will be left for future investigation.

While the large-scale zonal mean impacts appear to be mostly confined to the upper stratosphere, the gravity wave forcing from increasing the number of bins show regional effects within the troposphere. The local impacts of including this effect in parametrizations, although not fully investigated here, may be important for regional skill over complex orography even at finer grid-spacing of 9 km. For example, comparing realistic high resolution (~ 2.5 km grid-spacing) gravity wave momentum fluxes with output from parametrizations at ~9 km grid-spacing over the Drake's Passage (see Figure 12 and 14 of Kruse et al. 2022) showed that the parametrizations, which employ the monochromatic wave approximation, had too little wave forcing at high altitudes and a peak at lower altitudes. This vertical profile is consistent with what is seen in Figure 6, in which the parametrization with 1 bin deposits all its momentum over a relatively shallow depth. This supports the addition of partial critical level filtering in orographic gravity wave drag parametrizations even at relatively high resolution (9 km) and motivates future studies looking at this phenomena in more detail. To that end, the observations and modelling activities planned for TEAMx (Serafin et al., 2020) would be ideal for gaining a better understanding of these local processes over the European Alps.

Increases in parametrization complexity are often accompanied by increases in computational cost. It is, therefore, worth

stating that in the tests presented here with the Met Office Unified Model, increasing the number of bins from 1 to 32 leads to approximately a 6% increase in wall-clock time and less than a 1% increase in memory used at a resolution of 40 km. This is a moderate increase in memory and compute time and would be acceptable for operational use. It should be emphasised, however, that this code has not been optimised in any way and the authors are confident that this overhead can be significantly reduced with elicitation from expert scientific software engineers.

A | APPENDIX: WAVE SATURATION THEORY

This appendix provides an overview of the theory and assumptions made in deriving the wave saturation criterion used in the current Met Office Unified Model orographic gravity wave drag parametrization, and many parametrizations employed in other models. Following McFarlane (1987), convective instability is diagnosed when the gradient of the total potential temperature (mean plus wave perturbation) becomes negative $\frac{\partial}{\partial z}(\theta + \theta') < 0$. The potential temperature perturbation can be written in terms of the vertical displacement of isentropic surfaces, η:

$$\theta' = -\eta \frac{d\theta}{dz} \tag{A.1}$$

It then follows that the criterion for saturation is:

$$\frac{\partial}{\partial z}(\theta' + \theta) = \left(1 - \frac{\partial \eta}{\partial z}\right)\frac{d\theta}{dz} - \eta \frac{d^2\theta}{dz^2} < 0 \tag{A.2}$$

In order to diagnose where wave saturation occurs, the current scheme assumes that the term involving $\frac{d^2}{dz^2}\theta$ is small compared with the others, such that the criterion becomes $\frac{\partial}{\partial z}\eta > 1$. More significantly, the scheme assumes that the wave produced by the orography is monochromatic, such that the streamline displacement takes the form:

$$\eta = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\eta}(k_1, l_1) exp(i(kx + ly))dkdl = \hat{\eta}\cos(k_1 x + l_1 y) \tag{A.3}$$

where the integrals over k, l have been performed because $\hat{\eta}(k_1, l_1)$ is a monochromatic wave and, therefore, acts as a delta function in k, l. This means that the wave has only one wavevector (given by (k_1, l_1)) and one coefficient describing its amplitude ($\hat{\eta}$). The parametrization also takes the wavevector (k_1, l_1) to be in the direction of the surface stress (i.e. $(\tau_x, \tau_y) \parallel (k_1, l_1)$), which would be correct for the case of a monochromatic wave (see equation (5)).

The amplitude of the streamline displacement can then be written in terms of the vertical velocity in Fourier space:

$$\hat{\eta} = \frac{\hat{w}}{i(Uk_1 + Vl_1)} \tag{A.4}$$

Using the linear kinematic bottom boundary condition $\hat{w}_0 = i(U_0 k_1 + V_0 l_1)\hat{h}$ (Smith, 1980) and equation (3), an expression for the wave amplitude (below a critical level) can be derived:

$$\hat{\eta}(z) = \hat{h}\left(\frac{\rho_0 N_0 (U_0 k_1 + V_0 l_1)}{\rho N (Uk_1 + Vl_1)}\right)^{1/2} exp\left(iRe\left(\int_0^z m(z')dz'\right)\right) \tag{A.5}$$

The saturation criterion for the monochromatic wave can then be expressed in terms of large scale mean flow variables. First, by taking the vertical derivative of $\hat{\eta}$ and then only retaining terms up to zeroth order (i.e. neglecting terms involving the derivatives of background fields U, V and N):

$$\frac{d\hat{\eta}}{dz} = \frac{iNK_1}{(Uk_1 + Vl_1)}\hat{h}\left(\frac{\rho_0 N_0 (U_0 k_1 + V_0 l_1)}{\rho N (Uk_1 + Vl_1)}\right)^{1/2} exp\left(iRe\left(\int_0^z m(z')dz'\right)\right) \tag{A.6}$$

which can be written in terms of the amplitude itself as:

$$\frac{d\hat{\eta}}{dz} = \frac{iNK_1}{(Uk_1 + Vl_1)}\hat{\eta} \tag{A.7}$$

Since we are only interested in the height at which the peak of the wave becomes convectively unstable, we are only concerned with the wave amplitude at a particular height and not the phase of the wave, i.e. $\mathsf{max}(|\eta(z)|) = |\hat{\eta}(z)|$ where $\mathsf{max}(...)$ is the maximum taken over x and y. We can express the saturation condition as:

$$|\hat{\eta}| > \left|\frac{(Uk_1 + Vl_1)}{NK_1}\right| \tag{A.8}$$

Taking the ratio of amplitude at a particular height, $\hat{\eta}(z)$, and at a height level below, $\hat{\eta}(z - \Delta z)$, using equation (A.5), we can express the amplitude growth with height:

$$|\hat{\eta}(z)| = \left|\hat{\eta}(z - \Delta z)\left(\frac{\rho(z - \Delta z)N(z - \Delta z)(U(z - \Delta z)k_1 + V(z - \Delta z)l_1)}{\rho(z)N(z)(U(z)k_1 + V(z)l_1)}\right)^{1/2}\right| \tag{A.9}$$

To employ this saturation criterion within the parametrization scheme, a launch amplitude $\hat{\eta}_0$ at the surface is required. Since the VV21 scheme, like most orographic gravity wave drag schemes, uses only one wavevector direction, only a single launch amplitude is needed to describe the wave. This launch amplitude attempts to represent the mean amplitude over all orographic scales by assuming that it is proportional to the standard deviation of the orography within a grid box, such that $\hat{\eta}_0 = n\sigma$. Here, $\sigma^2 = A^{-1}4\pi^2 \int\int |\hat{h}|^2 dk dl$ is the mean variance of the orography within the gridbox and n is a tuning constant. The wave amplitude then grows according to equation (A.9), until the criterion $\hat{\eta}(z) > F_{sat}|\frac{Uk_1+Vl_1}{NK_1}|$ (i.e. equation A.8) is met. A tuning parameter, F_{sat}, is introduced. The amplitude is set to $F_{sat}|\frac{(Uk_1+Vl_1)}{NK_1}|$ where saturation is diagnosed.

In the monochromatic wave assumption, the absolute momentum flux is proportional to the square of the wave amplitude, i.e. $|\tau|(z) \propto \hat{\eta}^2(z)$ (Teixeira, 2014; McFarlane, 1987). This means that the momentum flux variation with height can be expressed as:

$$|\tau|(z) = |\tau|(z - \Delta z)\left(\frac{\hat{\eta}(z)}{\hat{\eta}(z - \Delta z)}\right)^2\left(\frac{\rho(z)N(z)(U(z)k_1 + V(z)l_1)}{\rho(z - \Delta z)N(z - \Delta z)(U(z - \Delta z)k_1 + V(z - \Delta z)l_1)}\right) \tag{A.10}$$

The momentum fluxes are then resolved into their x and y components using:

$$\tau_x(z), \tau_y(z) = |\tau|(z)\frac{\tau_{x,0}}{|\tau|_0}, |\tau|(z)\frac{\tau_{y,0}}{|\tau|_0} \tag{A.11}$$

where $\tau_{x,0}, \tau_{y,0}$ are the surface momentum fluxes.

The resultant force from the gravity waves onto the mean flow is then computed as:

$$\frac{d}{dt}(U,V)_{GWD} = -\frac{1}{\rho}\frac{d}{dz}(\tau_x(z),\tau_y(z)) \tag{A.12}$$

When $(U(z-\Delta z)k_1+V(z-\Delta z)l_1)*(U(z)k_1+V(z)l_1) <= 0$, i.e. the wind in the direction of the wavevector has encountered a critical level across two vertical levels, all of the gravity wave momentum is deposited. This method of saturation ensures that both convective instability and total critical level filtering is accounted for in the scheme. However, since there is only one wave vector, there is only one critical level for the scheme.

ACKNOWLEDGEMENTS

A van Niekerk was supported by the Met Office Weather and Climate Science for Service Partnership (WCSSP) India as part of the Newton Fund. We would like to acknowledge the members of the 'New Quantitative Constraints on Orographic Gravity Wave Stress and Drag' team at the International Space Science Institute in Bern for their helpful discussion. The authors would also like to thank Steve Derbyshire and Keith Williams for reading a draft of the manuscript and for their support. The authors would like to thank Petr Šácha and one anonymous reviewer for their careful reading of the manuscript and their very useful comments.

DATA AVAILABILITY

Data used in the production of plots are available for download at `https://doi.org/10.5281/zenodo.6668952`.

REFERENCES

Amemiya, A. and Sato, K. (2016), 'A new gravity wave parameterization including three-dimensional propagation', *Journal of the Meteorological Society of Japan. Ser. II* **94**(3), 237–256.

Baines, P. G. (1987), 'Upstream blocking and airflow over mountains', *Annual Review of Fluid Mechanics* **19**(1), 75–95.

Berckmans, J., Woollings, T., Demory, M.-E., Vidale, P.-L. and Roberts, M. (2013), 'Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing', *Atmospheric Science Letters* **14**(1), 34–40.

Boutle, I. A., Eyre, J. E. J. and Lock, A. P. (2014), 'Seamless stratocumulus simulation across the turbulent gray zone', *Monthly Weather Review* **142**(4), 1655 – 1668.

Bretherton, F. P. (1966), 'The propagation of groups of internal gravity waves in a shear flow', *Quarterly Journal of the Royal Meteorological Society* **92**(394), 466–480.

Broad, A. S. (1995), 'Linear theory of momentum fluxes in 3-d flows with turning of the mean wind with height', *Quarterly Journal of the Royal Meteorological Society* **121**(528), 1891–1902.

Bölöni, G., Kim, Y.-H., Borchert, S. and Achatz, U. (2021), 'Toward transient subgrid-scale gravity wave representation in atmospheric models. part i: Propagation model including nondissipative wave–mean-flow interactions', *Journal of the Atmospheric Sciences* **78**(4), 1317 – 1338.
URL: *https://journals.ametsoc.org/view/journals/atsc/78/4/JAS-D-20-0065.1.xml*

Danielson, J. and Gesch, D. (2018), 'Global multi-resolution terrain elevation data 2010 (gmted2010)', *U.S. Geo- logical Survey Open-File Report* p. 2011–1073.

Doyle, J. D. and Jiang, Q. (2006), 'Observations and numerical simulations of mountain waves in the presence of directional wind shear', *Quarterly Journal of the Royal Meteorological Society* **132**(619), 1877–1905.

Dörnbrack, A. (2021), 'Stratospheric mountain waves trailing across northern europe', *Journal of the Atmospheric Sciences* **78**(9), 2835 – 2857.

Eckermann, S. D., Ma, J., Wu, D. L. and Broutman, D. (2007), 'A three-dimensional mountain wave imaged in satellite radiance throughout the stratosphere: Evidence of the effects of directional wind shear', *Quarterly Journal of the Royal Meteorological Society* **133**(629), 1959–1975.

Fritts, D. C. (1984), 'Gravity wave saturation in the middle atmosphere: A review of theory and observations', *Reviews of Geophysics* **22**(3), 275–308.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG022i003p00275*

Fritts, D. C., Dong, W., Lund, T. S., Wieland, S. and Laughman, B. (2020), 'Self-acceleration and instability of gravity wave packets: 3. three-dimensional packet propagation, secondary gravity waves, momentum transport, and transient mean forcing in tidal winds', *Journal of Geophysical Research: Atmospheres* **125**(3), e2019JD030692. e2019JD030692 2019JD030692.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD030692*

Guarino, M.-V., Teixeira, M. A. C., Keller, T. L. and Sharman, R. D. (2018), 'Mountain-wave turbulence in the presence of directional wind shear over the rocky mountains', *Journal of the Atmospheric Sciences* **75**(4), 1285 – 1305.

Hasha, A., Bühler, O. and Scinocca, J. (2008), 'Gravity wave refraction by three-dimensionally varying winds and the global transport of angular momentum', *Journal of the Atmospheric Sciences* **65**(9), 2892 – 2906.
URL: *https://journals.ametsoc.org/view/journals/atsc/65/9/2007jas2561.1.xml*

Heim, C., Panosetti, D., Schlemmer, L., Leuenberger, D. and Schär, C. (2020), 'The influence of the resolution of orography on the simulation of orographic moist convection', *Monthly Weather Review* **148**(6), 2391 – 2410.

Jiang, Q. and Doyle, J. D. (2004), 'Gravity wave breaking over the central alps: Role of complex terrain', *Journal of the Atmospheric Sciences* **61**(18), 2249 – 2266.

Kanehama, T., Sandu, I., Beljaars, A., van Niekerk, A. and Lott, F. (2019), 'Which orographic scales matter most for medium-range forecast skill in the northern hemisphere winter?', *Journal of Advances in Modeling Earth Systems* **11**(12), 3893–3910.

Kidston, J., Scaife, A. A., Hardiman, S. C., Mitchell, D. M., Butchart, N., Baldwin, M. P. and Gray, L. J. (2015), 'Stratospheric influence on tropospheric jet streams, storm tracks and surface weather', *Nature Geoscience* **8**, 433 – 440.

Kim, Y.-J. and Arakawa, A. (1995), 'Improvement of orographic gravity wave parameterization using a mesoscale gravity wave model', *Journal of Atmospheric Sciences* **52**(11), 1875 – 1902.

Kruse, C. G., Alexander, M. J., Hoffmann, L., van Niekerk, A., Polichtchouk, I., Bacmeister, J. T., Holt, L., Plougonven, R., Šácha, P., Wright, C., Sato, K., Shibuya, R., Gisinger, S., Ern, M., Meyer, C. I. and Stein, O. (2022), 'Observed and modeled mountain waves from the surface to the mesosphere near the drake passage', *Journal of the Atmospheric Sciences* **79**(4), 909 – 932.

Kruse, C. G. and Smith, R. B. (2015), 'Gravity wave diagnostics and characteristics in mesoscale fields', *Journal of the Atmospheric Sciences* **72**(11), 4372 – 4392.

Kruse, C. G. and Smith, R. B. (2018), 'Nondissipative and dissipative momentum deposition by mountain wave events in sheared environments', *Journal of the Atmospheric Sciences* **75**(8), 2721 – 2740.

Lott, F. and Miller, M. J. (1997), 'A new subgrid-scale orographic drag parametrization: Its formulation and testing', *Quarterly Journal of the Royal Meteorological Society* **123**(537), 101–127.

McFarlane, N. A. (1987), 'The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere', *Journal of Atmospheric Sciences* **44**(14), 1775 – 1800.

Nappo, C. J. (2013), *An introduction to atmospheric gravity waves*, Academic press.

Palmer, T. N., Shutts, G. J. and Swinbank, R. (1986), 'Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization', *Quarterly Journal of the Royal Meteorological Society* **112**(474), 1001–1039.

Phillips, D. S. (1984), 'Analytical surface pressure and drag for linear hydrostatic flow over three-dimensional elliptical mountains', *Journal of Atmospheric Sciences* **41**(6), 1073 – 1084.

Plougonven, R., de la Cámara, A., Hertzog, A. and Lott, F. (2020), 'How does knowledge of atmospheric gravity waves guide their parameterizations?', *Quarterly Journal of the Royal Meteorological Society* **146**(728), 1529–1543.

Polichtchouk, I., van Niekerk, A. and Wedi, N. (2022), 'Resolved gravity waves in the extra-tropical stratosphere: Effect of horizontal resolution increase from o(10 km) to o(1 km)', *Journal of the Atmospheric Sciences* .
URL: *https://journals.ametsoc.org/view/journals/atsc/aop/JAS-D-22-0138.1/JAS-D-22-0138.1.xml*

Polichtchouk, I., Wedi, N. and Kim, Y.-H. (2022), 'Resolved gravity waves in the tropical stratosphere: Impact of horizontal resolution and deep convection parametrization', *Quarterly Journal of the Royal Meteorological Society* **148**(742), 233–251.

Sandu, I., van Niekerk, A. and Shepherd, T. e. a. (2019), 'Impacts of orography on large-scale atmospheric circulation', *npj Clim Atmos Sci* **2**(10), 2397–3722.

Sato, K., Tateno, S., Watanabe, S. and Kawatani, Y. (2012), 'Gravity wave characteristics in the southern hemisphere revealed by a high-resolution middle-atmosphere general circulation model', *Journal of the Atmospheric Sciences* **69**(4), 1378 – 1396.

Scinocca, J. F. and Sutherland, B. R. (2010), 'Self-acceleration in the parameterization of orographic gravity wave drag', *Journal of the Atmospheric Sciences* **67**(8), 2537 – 2546.
URL: *https://journals.ametsoc.org/view/journals/atsc/67/8/2010jas3358.1.xml*

Serafin, S., Rotach, M., Arpagaus, M., Colfescu, I., Cuxart, J., De Wekker, S., Evans, M., Grubišić, V., Kalthoff, N., Karl, T., Kirshbaum, D., Lehner, M., Mobbs, S., Paci, A., Palazzi, E., Raudzens Bailey, A., Schmidli, J., Wohlfahrt, G. and Zardi, D. (2020), *Multi-scale transport and exchange processes in the atmosphere over mountains - Programme and experiment*.

Shutts, G. (1995), 'Gravity-wave drag parametrization over complex terrain: The effect of critical-level absorption in directional wind-shear', *Quarterly Journal of the Royal Meteorological Society* **121**(525), 1005–1021.

Shutts, G. J. (1998), 'Stationary gravity-wave structure in flows with directional wind shear', *Quarterly Journal of the Royal Meteorological Society* **124**(549), 1421–1442.

Shutts, G. J. and Gadian, A. (1999), 'Numerical simulations of orographic gravity waves in flows which back with height', *Quarterly Journal of the Royal Meteorological Society* **125**(559), 2743–2765.

Sigmond, M., Scinocca, J. F. and Kushner, P. J. (2008), 'Impact of the stratosphere on tropospheric climate change', *Geophysical Research Letters* **35**(12).

Smith, R. B. (1979), The influence of mountains on the atmosphere, Vol. 21 of *Advances in Geophysics*, Elsevier, pp. 87–230.

Smith, R. B. (1980), 'Linear theory of stratified hydrostatic flow past an isolated mountain', *Tellus* **32**(4), 348–364.
URL: *https://doi.org/10.3402/tellusa.v32i4.10590*

Teixeira, M. A. C. (2014), 'The physics of orographic gravity wave drag', *Frontiers in Physics* **2**, 43.

Teixeira, M. A. C. and Miranda, P. M. A. (2009), 'On the momentum fluxes associated with mountain waves in directionally sheared flows', *Journal of the Atmospheric Sciences* **66**(11), 3419 – 3433.

Teixeira, M. A. and Yu, C. L. (2014), 'The gravity wave momentum flux in hydrostatic flow with directional shear over elliptical mountains', *European Journal of Mechanics - B/Fluids* **47**, 16–31. Enok Palm Memorial Volume.
URL: *https://www.sciencedirect.com/science/article/pii/S0997754614000259*

van Niekerk, A., Sandu, I. and Vosper, S. B. (2018), 'The circulation response to resolved versus parametrized orographic drag over complex mountain terrains', *Journal of Advances in Modeling Earth Systems* **10**(10), 2527–2547.

van Niekerk, A., Sandu, I., Zadra, A., Bazile, E., Kanehama, T., Köhler, M., Koo, M.-S., Choi, H.-J., Kuroki, Y., Toy, M. D., Vosper, S. B. and Yudin, V. (2020), 'Constraining orographic drag effects (coorde): A model comparison of resolved and parametrized orographic drag', *Journal of Advances in Modeling Earth Systems* **12**(11), e2020MS002160. e2020MS002160 10.1029/2020MS002160.

van Niekerk, A. and Vosper, S. (2021), 'Towards a more "scale-aware" orographic gravity wave drag parametrization: Description and initial testing', *Quarterly Journal of the Royal Meteorological Society* **147**(739), 3243–3262.

Van Zyl, J. J. (2001), 'The shuttle radar topography mission (srtm): a breakthrough in remote sensing of topography', *Acta Astronautica* **48**(5-12), 559–565.

Wells, H., Vosper, S. B., Ross, A. N., Brown, A. R. and Webster, S. (2008), 'Wind direction effects on orographic drag', *Quarterly Journal of the Royal Meteorological Society* **134**(632), 689–701.

Williams, K. D., van Niekerk, A., Best, M. J., Lock, A. P., Brooke, J. K., Carvalho, M. J., Derbyshire, S. H., Dunstan, T. D., Rumbold, H. S., Sandu, I. and Sexton, D. M. H. (2020), 'Addressing the causes of large-scale circulation error in the met office unified model', *Quarterly Journal of the Royal Meteorological Society* **146**(731), 2597–2613.

Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., Melvin, T., Smith, C., Vosper, S., Zerroukat, M. and Thuburn, J. (2014), 'An inherently mass-conserving semi-implicit semi-lagrangian discretization of the deep-atmosphere global non-hydrostatic equations', *Quarterly Journal of the Royal Meteorological Society* **140**(682), 1505–1520.

Xu, X., Xue, M., Teixeira, M. A. C., Tang, J. and Wang, Y. (2019), 'Parameterization of directional absorption of orographic gravity waves and its impact on the atmospheric general circulation simulated by the weather research and forecasting model', *Journal of the Atmospheric Sciences* **76**(11), 3435 – 3453.