ORIGINAL ARTICLE

Journal Section

What can we learn about orographic drag parametrization from high-resolution models? A case study over the Rocky mountains

Simon B. Vosper[1] | **Annelize van Niekerk**[1] | **Andrew Elvidge**[2] | **Irina Sandu**[3] | **Anton Beljaars**[3]

[1]Met Office, FitzRoy Road, Exeter EX1 3PB, UK

[2]School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK

[3]European Centre for Medium Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, UK

Correspondence
Met Office, FitzRoy Road, Exeter EX1 3PB, UK
Email: AnnelizevanNiekerk@MetOffice.gov.uk

Funding information

Comprehensive high resolution numerical weather prediction models provide a virtual laboratory for modelling the atmospheric flow over complex mountain ranges. In this study, global and regional simulations with horizontal grid spacing ranging from 2 km to 32 km, focused over the northern Rocky mountains, are used to assess the orographic blocking and gravity wave drag parametrizations employed in the Met Office Unified Model (UM) and the European Centre for Medium-Range Weather Forecasts Integrated Forecasting System (IFS). The total, resolved and parametrized drag components in coarse resolution simulations are compared with those in high resolution simulations, in which the orographic drag processes are better resolved. The total surface stresses and gravity wave momentum fluxes in the free atmosphere of the global 16km UM and IFS simulations are shown to compare well with 2km regional simulations, in terms of their variability and mean. While the total gravity wave momentum flux is somewhat underestimated by both global models, its vertical distribution is well captured. The 'seamlessness' of the parametrization scheme is then assessed by comparing the total orographic stress, and its components, across several horizontal resolu-

*Equally contributing authors.

tions of the UM. The surface stress remains relatively constant across resolutions, such that the reduction in resolved orographic stress at coarser resolutions is compensated by an almost equivalent increase in parametrized orographic stress. However, the parametrized orographic gravity wave momentum flux in the free atmosphere remains almost constant with resolution, failing to compensate for the lack of resolved flux at coarse resolutions. This leads to an underestimation of the total gravity wave drag at coarser resolutions. Further analysis suggests that this underestimation is due to the monochromatic wave assumption made by the gravity wave drag parametrization scheme.

KEYWORDS
Orographic drag, Gravity wave drag, High resolution modelling, numerical weather prediction, Mountains

1 | INTRODUCTION

The drag imparted on the atmosphere by orography plays an important role in the momentum budget and, correspondingly, the atmospheric circulation. The representation of orographic drag in numerical models, therefore, influences the reliability of predictions across all timescales, from days (Sandu et al., 2019) to climate timescales (Pithan et al., 2016; van Niekerk et al., 2017). However, models used for numerical weather prediction (NWP), seasonal forecasting and climate projections typically only resolve some portion of the orographic scales, resulting in the need for orographic drag parametrization schemes. These schemes generally only represent some of the key orographic drag processes that occur in the real atmosphere, namely orographic blocking and gravity wave drag, and are based typically on theoretical and experimental results from idealised situations: simple flows over single isolated mountains (e.g. Phillips, 1984). Parameters within these schemes are also poorly constrained by observations and, thus, rely on tuning in order to optimise objective measures of forecast skill in, for example, surface pressure, winds and temperatures. Manifestly, the partition between orographic drag and boundary-layer drag, two processes with different scale dependencies and different responses to stratification, varies wildly between different operational NWP models (Zadra et al., 2013).

With the use of grid lengths O(10 km) in global NWP, some of the key orographic processes are becoming better resolved. However, since orography spans multiple scales, all current global NWP and climate models are applied in the orographic drag 'grey zone' (Vosper et al., 2016), in which the drag-producing processes are neither well resolved nor fully parametrized. The handover between parametrized and resolved drag across resolutions is, therefore, an important aspect of atmospheric modelling across time and spatial scales. In order for models to operate 'seamlessly' (e.g. Brown et al., 2012) the parametrizations must accurately account for the unresolved orographic drag over a wide range of resolutions. The extent to which this is true for current orographic drag parametrizations over relatively simple isolated mountains was examined by Vosper (2015). This study showed that, with a well-tuned orographic drag scheme, the hand over between the resolved and parametrized surface pressure drag is well behaved across

resolutions. At high resolution, when the orography is well resolved, the parametrized drag is small and the total drag is dominated by resolved processes. The resolved drag reduces at coarser resolutions, but this is compensated for by an increase in the parametrized drag. Vosper et al. (2016) also showed that this tuning was not globally universal and was dependent on the scale of the sub-grid mountains. While it may have been possible to optimise the orographic drag parametrization for simple isolated mountains, the extent to which they are capable of representing drag across extended and complex mountain ranges is only beginning to emerge. For example, the recent study of Garner (2018) used high resolution simulations over various extended mountain regions to validate and determine the optimal drag coefficients for orographic drag parametrization, thus demonstrating their usefulness for model development.

Constraining orographic drag parametrization schemes and assessing their realism is a major challenge. Surface pressure drag arises as a consequence of correlations between small (~1 hPa) flow induced pressure perturbations and terrain slope and, although historically there have been attempts to quantify this through observational studies (e.g. Smith, 1978; Vosper and Mobbs, 1997), these correlations are very difficult to measure at the level of accuracy required for model evaluation. Similarly, mountain-wave momentum fluxes, which vary spatially in three dimensions and are intermittent in time (Plougonven et al., 2013), require comprehensive precision observations in order to be reliably measured. Aircraft studies (e.g. Fritts et al., 2016) provide the most accurate observations, but these are expensive and do not sample all of the wave field. While such measurements are very useful for evaluating whether gravity waves are explicitly resolved by models (e.g. Shutts and Broad, 1993), the combination of sampling issues and three-dimensionality of the wave fields mean they are not suitable for reliable assessment of the grid-box average fluxes represented by parametrization schemes. The spatial and temporal coverage afforded by satellite remote-sensing techniques (Alexander and Barnet, 2007; Wright and Mitchell, 2017) do offer considerable promise in this respect, although their usefulness is somewhat limited by a lack of resolution (in either the horizontal, or vertical, for limb- and nadir-sounders, respectively (Preusse et al., 2008)) and the fact that reliable measurements are generally limited to the middle atmosphere.

An alternative, indirect approach to constraining parametrized orographic drag is to make use of high-resolution km-scale simulations. Simulations with horizontal grid spacings smaller than a few kilometres are able to resolve many, though not all, orographic drag processes, and can be applied to realistic time-evolving flows over complex terrain. van Niekerk et al. (2018) used such simulations to compare the total impact on the circulation from resolved orographic drag at high resolution and parametrized orographic drag at coarser resolution. In contrast, this study compares the stress and drag from the orographic drag parametrizations in low-resolution simulations to the corresponding stress and drag from resolved orography in the high-resolution simulations. To this end, high-resolution regional simulations over the northern Rocky mountains (domain shown in Figure 1) are used to assess the performance of drag parametrizations at different resolutions.

This paper has three primary objectives, each addressed in sections as follows:

- By comparing the resolved and parametrized drag from two different global NWP models with a high-resolution regional model, assess the ability of the global models to realistically represent the orographic flow-blocking and gravity-wave drag. To the extent that the high-resolution model can be regarded as a 'truth', this allows us to assess the realism of the parametrizations used in operational NWP models. Both the total surface stress and the vertical distribution of the orographic gravity wave momentum fluxes and drag will be considered. Results are presented in section 3.1.
- Extend the Vosper et al. (2016) study from relatively simple isolated mountainous islands to a more complex major continental mountain range, determining the behaviour of orographic drag parametrizations across a range of horizontal resolutions. Since contributions to to the total drag exerted on the large-scale circulation from these

larger mountain chains will dominate over those from small islands, the behaviour of the parametrization schemes in this case is of greater relevance to global NWP, seasonal and climate projections. This is investigated in section 3.2.

- Quantify the scale contributions of the total orographic gravity wave momentum fluxes, and how this varies across resolutions. Identifying both the smallest scale that the model is able to resolve at a particular resolution and which scales contribute most to the total momentum flux will help to inform the development of more 'seamless' orographic drag parametrizations. This is discussed in section 3.3.

Conclusions are drawn in section 4.

2 | METHODOLOGY

Global model simulations are performed with the Met Office Unified Model (UM) at version 10.9 and the European Centre for Medium-Range Weather Forecasts Integrated Forecasting System (IFS) at cycle43r1. These have an approximate horizontal grid spacing of 16 km, i.e. resolution N768 of the UM and TL1279 of the IFS, which are typical of the resolutions used until recently for global weather forecasts. The global UM has 70 vertical levels reaching 80km above sea level (ASL) and the global IFS has 137 vertical levels reaching 1 Pa. The scientific configuration of the global UM is described in Walters et al. (2019) and the IFS is described by ECMWF (2015). The global simulations are made up of a series of 24 hour forecasts initialised from their respective operational analyses at 00UTC spanning the month long period 01 December 2015 - 01 January 2016. Both models use a variant of the Lott and Miller (1997) orographic drag parametrization to represent low-level flow blocking and mountain-wave drag, but vary in certain aspects of its implementation. They both also include a parametrization for turbulent form drag associated with sub-grid hills whose horizontal scale is smaller than ~ 5 km (and hence stability effects are considered generally unimportant). These schemes differ in that the UM uses an effective roughness approach (Wood and Mason, 1993), whereby the vegetative roughness length is increased to represent the effects of the hills, whereas the IFS uses an explicit stress representation (Beljaars et al., 2004). Previous studies have shown that the partitioning between the orographic and boundary layer drag is very different between these two models (Sandu et al., 2016; Zadra et al., 2013). It is likely that these differences result from of a combination of subjective tuning, differences in formulation and, as has been recently demonstrated by Elvidge et al. (2019), the representation of sub-grid orographic fields. The implications of these differences between the UM and IFS's treatment of parametrized orographic drag for the total drag and its variability over the northern Rocky mountains will be considered.

The regional simulations in this study were conducted with the UM, run across a range of horizontal grid resolutions: 2, 4, 8, 16 and 32 km. These all have 70 vertical levels up to 40 km ASL. The domain size, shown in Figure1, is identical for all regional simulations, which are all driven by the same global simulations. Boundary conditions are provided from a series of 24 hr global UM forecasts at a resolution of 22 km (N512), initialised from Met Office operational analyses at 00UTC. The regional simulations are free-running (not reinitialised every 24 hours) but broken into five, approximately week long, periods: 01-08 December 2015; 07-14 December 2015; 13-20 December 2015; 19-26 December 2015; and 25 December 2015 - 01 January 2016. The first day of each period overlaps with the last day of the previous and can thus be treated as a 'spin-up' and is disregarded. The scientific configuration is identical across the regional simulations, with parametrized drag turned on but parametrized convection turned off (see Vosper et al., 2016 for details of regional configurations).

In this study, the resolved orographic stress and the parametrized boundary layer and orographic stress, both at

the surface and throughout the atmosphere, is diagnosed as a means of validating the orographic drag parametrization schemes. Deriving surface stress and vertical momentum flux from resolved orography over a regional domain is complicated by the fact that the domain contains partial large-scale waves and that there may be mean slopes in the orography across the boundaries of the domain. These large-scale gradients do not contribute to the meso-scale drag that is of interest to the parametrization problem but may dominate the magnitude of the calculated terms. In order to avoid these effects, the area averaged resolved orographic surface stress is computed following Smith et al. (2005) using:

$$\vec{\tau}_{res}(z=0) = A^{-1} \iint_A (p'_* - \overline{p'_*}) \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) h(x,y) dx dy, \tag{1}$$

where $h(x, y)$ is the surface height and A is the domain area. The surface pressure perturbation p'_* is defined as the surface pressure after removal of the standard pressure:

$$p_{std} = p_0 \left(1 - \frac{\gamma h(x,y)}{T_0} \right)^{\frac{g}{R\gamma}}, \tag{2}$$

where $p_0 = 101324\ Pa$ is the standard atmospheric pressure at mean sea level, $T_0 = 288.15\ K$ is the reference temperature, $\gamma = 0.0065\ Km^{-1}$ is an assumed constant temperature lapse rate in the troposphere, $g = 9.81\ ms^{-2}$ is the acceleration due to gravity and $R = 287.05\ m^2 s^{-2} K^{-1}$ is the gas constant of dry air. The overbar in Equation 1 represents the spatial mean. Without the removal of p_{std}, the surface stress magnitudes are much larger. As discussed in Carissimo et al. (1988), this is because the average pressure over the domain (i.e. the mass of the atmosphere) will be a function of $h(x, y)$ and, since the heights at the boundaries are not equal, this average pressure will act on the mean slope. By removing the dependence of p on $h(x, y)$, and its spatial average, the force resulting from the weight of the atmosphere acting on the mean slope is removed. The total surface stress is then given by $\tau_{tot} = |\vec{\tau}_{res} + \vec{\tau}_{bl} + \vec{\tau}_{sso}|$, where $\vec{\tau}_{bl}$ is the parametrized boundary layer surface stress vector and $\vec{\tau}_{sso}$ is the parametrized orographic surface stress vector, which has contributions from both orographic blocking ($\vec{\tau}_{blk}$) and gravity wave momentum flux ($\vec{\tau}_{gwd}$).

The magnitude of the resolved vertical momentum fluxes from gravity waves are computed on constant geometric height levels via:

$$\vec{\tau}_{res}(z) = A^{-1} \iint_A \rho_0(u', v') w' dx dy \tag{3}$$

The perturbation velocities (u', v' and w') are derived using the method described in Kruse and Smith (2015), whereby the fields are detrended (via a linear planar fit) and a high-pass filter is applied in spectral space to remove horizontal scales larger than 800 km. The background density field (ρ_0) is determined by low-pass filtering the density field (removing scales smaller than 800 km), itself computed in a way consistent with the velocity perturbations. The total gravity wave momentum flux in the free atmosphere is given by the sum of the resolved and parametrized gravity wave momentum flux. The total acceleration (or drag) of the large-scale mean zonal wind is computed as:

$$\frac{\partial \overline{u}}{\partial t} = -\frac{1}{\rho_0} \frac{\partial}{\partial z} \left(\tau_{res_x} + \tau_{gwd_x} \right) \tag{4}$$

where τ_{res_x} and τ_{gwd_x} is the resolved and parametrized gravity wave zonal momentum flux, respectively.

3 | RESULTS

3.1 | Assessment of parametrized and resolved orographic drag in two models

In order to establish the accuracy of the orographic drag parametrization schemes employed in the UM and the IFS, we begin by comparing the magnitude of the total surface stress (τ_{tot}) in the global 16 km simulations with that in the 2 km regional simulations. A timeseries of τ_{tot} over the month is shown in Figure 2, demonstrating the ability of the lower resolution global models to capture both the mean and variability of the much higher resolution simulations. This is confirmed by the values of the time-mean and temporal correlation of τ_{tot} between the global and the 2 km regional simulations in Table 1. The total surface stress is very similar in the UM and the IFS, although, perhaps unsurprisingly since they are run with the same model, the global UM has a slightly higher correlation with the 2 km simulations. There appears not to be a significant systematic over or under estimation of the time mean τ_{tot} by either model. There are, however, periods in which the IFS over estimates τ_{tot} and the UM under estimates it and vice versa, relative to the 2 km simulations.

Previous work (Zadra et al., 2013) has shown that, while the total parametrized surface stress is relatively similar in the UM and the IFS, the contributions from the various parametrized components differ. This analysis excluded the resolved surface stress component, which may also differ between models as a result of, for example, the model dynamical formulation, resolution and representation of resolved orography (Elvidge et al., 2019). As a means of quantifying the contribution of the resolved stress to the total surface stress, figure 3 shows the partitioning of τ_{tot} into its parametrized and resolved components for both the UM and IFS. Note that the magnitude of the individual components of the surface stress (τ_{res}, τ_{sso} and τ_{bl}) do not add up to the total surface stress magnitude (τ_{tot}), since τ_{tot} is the magnitude of the vector sum. This implies that the surface stress components are not parallel and that, while τ_{tot} may be similar across resolutions, the direction is not necessarily the same. Figure 3 shows that, despite the fact that the UM and the IFS have significantly different dynamic formulations (e.g. the UM is non-hydrostatic and the IFS is hydrostatic), spatial discretisation (e.g. the UM employs a regular latitude-longitude horizontal grid and a hybrid-height terrain following vertical coordinate, whereas, the IFS employs a spectral dynamical core and a sigma-pressure vertical coordinate system), the resolved surface stress is remarkably similar in both models in terms of its mean and variability. This suggests that most of the differences in τ_{tot} between the two models, although small on average, must be a result of differences in the parametrized stress.

As was found in the studies of Sandu et al. (2016) and Zadra et al. (2013), the breakdown of the individual parametrized components (Figure 3) confirms that the UM has much larger parametrized orographic blocking compared with the IFS, which in turn has much larger boundary layer stress. This difference exemplifies the consequences of having large uncertainties in drag schemes that are left open to tuning and, thus, compensation by other drag processes. The similarity in the total surface stress may not be surprising, since these magnitudes have resulted from many tuning iterations of the parametrized drag to maximise forecast skill. However, Figure 3 does demonstrate that the periods of over and under-estimation of τ_{tot}, relative to the 2 km simulations, may be attributable to the difference in the partitioning of the various subgrid processes. For example, between 17 and 19 December, the ratio of orographic to boundary layer stresses is, in both models, larger than typical for the month. Whilst total stresses are generally well represented by the UM, which has a large orographic component, they are underestimated by the IFS during this period. On the other hand, between 29 and 31 December, the ratio of orographic to boundary layer stresses is smaller than typical and, whilst the total stresses are somewhat overestimated in the IFS, they are underestimated in

the UM, due to underestimated BL stresses. Other periods with a similar difference in partitioning can be identified throughout the timeseries. In fact, there are very few periods where τ_{tot} in both models agree with each other **and** with the 2km simulations. One such period is between 10 to 12 of December, during which the total surface stress is dominated by the resolved component. This dependence of the variability of the total stress on the partitioning of the parametrized drag suggests that changes in the flow regime, and the differences in the flow dependence (such as different stability dependence) of the parametrized orographic drag and the boundary layer drag, may be the cause of these alternating periods of over and under estimation of the total surface stress.

The total surface stress is important for understanding the overall surface-atmosphere exchange but the vertical distribution of the gravity wave drag deposition will provide information on the accuracy of the gravity wave saturation approximations within the parametrization schemes. Vertical profiles of the time-mean momentum fluxes and zonal acceleration (drag) from resolved gravity waves, parametrized gravity waves and the total sum of the two are shown in Figure 4 for the global 16 km and regional 2 km simulations. Between approximately 10-20 km ASL, both the resolved and parametrized gravity momentum flux reduces substantially as the waves saturate and deposit momentum in this layer. This rapid saturation with height is due to the fact that, within the mid-latitudes, there is generally strong negative wind shear at these altitudes. As was the case for the total surface stress, the two models produce very similar total gravity wave momentum fluxes and mean zonal wind forcing, with the vertical distribution of the wave dissipation maximising at roughly the same altitude and reaching similar peak zonal wind deceleration of approximately $-0.26\ ms^{-1}day^{-1}$. They both also compare well, in terms of vertical distribution, with the gravity wave drag in the 2 km simulations. This is reassuring and gives us faith that the saturation hypothesis assumed in these parametrizations is, on average, a good representation of the resolved wave saturation over this region. It is worth noting that parametrizations that assume instantaneous wave propagation, monochromatic wave generation and dissipate according to Lindzen (1981) are expected to perform well in regions of strong negative wind shear towards zero (Kruse and Smith, 2018). This is due to the strong constraint that the rapid decrease in wind speeds, being close to the critical level for orographic gravity waves, apply on the maximum altitude of wave propagation. This means that issues associated with, for example, critical level filtering in directional wind shear are not prominent. However, both models underestimate the total gravity wave drag magnitude by approximately 30%, the reasons for which will be discussed in the subsequent sections.

Unlike the resolved surface stress, the resolved orographic gravity wave drag differs significantly between the UM and the IFS. The UM has less resolved orographic gravity wave drag compared with the IFS and this difference is compensated by the UM having more parametrized orographic gravity wave drag. Elvidge et al. (2019) showed that the filtering of the resolved orography in the UM, which will likely affect the amplitude of resolved orographic gravity waves, is much stronger than that employed in the IFS. This may explain why the UM has less resolved orographic gravity wave drag compared with the IFS. While these differences in partitioning between resolved and parametrized orographic gravity wave drag may seem inconsequential, studies have shown that having different contributions from resolved and parametrized gravity wave drag can be important for the variability of the atmosphere and for the strength of the stratosphere-troposphere coupling (e.g. Polichtchouk et al. (2018)).

3.2 | Assessment of parametrized and resolved orographic drag across resolutions

Weather and climate models are required to perform well across many different resolutions, and are run across different timescales and over particular regions. It is desirable to maintain a single physics configuration across resolutions, since this promotes traceability of model errors and avoids retuning at each resolution. This means that the schemes need to be well behaved and perform 'seamlessly' as the model resolution is varied. The ability of the model to maintain

constant total surface stresses and gravity wave momentum fluxes across resolutions is a good test of the seamlessness of drag parametrization schemes. As the resolved drag likely reduces at coarser resolutions, the parametrized drag should increase to counteract this. This section examines the seamlessness of the representation of drag in the UM across resolutions ranging from 2 km to 32 km, spanning the resolutions used in regional NWP to those used in global ensemble forecasting.

The time-mean total surface stresses and its individual components in the regional simulations across resolutions are shown in Figure 5 (a). The total surface stress is relatively well behaved across resolutions, since it remains somewhat constant. However, a slight increase in total surface stress at the coarsest resolution (32 km UM) is evident. This is due to the increasing parameterized orographic blocking (τ_{blk}) dominating over decreasing resolved orographic and parametrized boundary layer stresses. The boundary layer parametrization accounts for drag processes that are unresolved at the resolutions considered here, i.e. turbulence and orographic form drag. This decrease is, therefore, likely due to the fact that the increased orographic blocking drag with coarser resolution decelerates the low-level winds, which the boundary layer drag depends on, leading to less boundary layer drag. The resolved orographic stress only decreases slightly at coarser resolution due to the fact that it is dominated by the large scales (i.e. O(100km)), which are already well resolved at a grid spacing of 32 km. It is evident that the gravity wave component of the orographic drag parametrization contributes only a small amount to the total surface stress. For the same reason as was given for Figure 3, the individual components of the surface stress vector magnitude do not add up to the total surface stress magnitude in Figure 5(a). Analysis of the zonal and meridional components of the surface stress reveal that the $\vec{\tau}_{res}$ vector is almost perpendicular to the parameterized vectors (not shown), i.e. $\vec{\tau}_{sso}$ and $\vec{\tau}_{bl}$, which would explain why it is in excess of $\vec{\tau}_{tot}$ during particular periods. What is more, the total stress vector rotates by $\sim 20^{\circ}$ clockwise in going from $\Delta x = 2$km to $\Delta x = 32$km.

The degree to which the time-mean and variability of the total orographic stress present in the 2 km UM is captured in the low resolution simulations becomes increasingly dependent on the orographic drag parametrization at coarser resolutions. Figure 6(a) shows a timeseries of the total surface stress, and its individual components, in the 32km UM alongside the total surface stress in the 2 km UM. The time-mean and temporal correlation with the 2 km UM of the total surface stress in the 32 km UM is also stated in Table 1. As is the case for the global 16 km simulations discussed in section 3.1, the variability in the 32 km UM is very similar to the 2 km UM, with a correlation coefficient of 0.966. This implies that, despite the parametrization being a dominant contribution to the model's surface momentum budget, it is able to accurately capture the variability of the explicitly resolved surface stress. This is particularly important towards the end of the month, when the contribution from the resolved and parametrized orographic stress is almost equal. In this case, with the parameter settings used operationally across all configurations of the UM, the parametrization scheme is able to represent the resolved orographic surface stress but it is likely that other orographic regions require different tuning, as was found in Vosper (2015) and Vosper et al. (2016).

Despite its small contribution to the total surface stress, the momentum deposited by orographic gravity waves in the free atmosphere can be substantial and plays an important role in the circulation of the stratosphere, which acts as a source of memory for the atmospheric system (Osprey et al., 2010). An assessment of its contribution across resolutions is, therefore, highly relevant to longer-range (e.g. seasonal) forecasting. The variation with resolution of the total gravity wave momentum flux vector magnitude at 7 km ASL, and the breakdown into resolved and parametrized components, is shown in Figure 5(b). As is to be expected, the resolved flux reduces with coarser resolution. However, the parametrized flux remains almost constant at grid spacing larger than 4 km, leading to an underestimation of the total gravity momentum flux at coarser resolutions. This is consistent with the global 16 km simulations of the UM and the IFS described in section 3.1 and with the findings of van Niekerk et al. (2018), who showed that an underestimation of the orographic gravity wave drag in the lower stratosphere leads to large systematic errors in the lower stratospheric

zonal winds at resolutions ranging from 16 km to 150 km.

The parametrized gravity wave flux dominates over the resolved flux at coarser resolutions, the implication of which is that the model is highly reliant on the gravity wave drag parametrization to capture the variability within the free atmosphere. A timeseries comparing the total, resolved and parametrized gravity wave momentum flux magnitudes at 7 km ASL with the total in the 2 km simulations is shown in Figure 6(b). The parametrization does a remarkable job at capturing the variability of the resolved gravity waves at higher resolution. There are periods in which the parametrization is clearly underestimating the momentum fluxes, particularly during the middle portion of the month. The orographic gravity wave drag parametrization does, however, capture the highly intermittent behaviour that may be important for the long-term atmospheric variability and mean state of the stratosphere, as has been shown to be the case for non-orographic gravity wave drag (de La Cámara et al., 2016).

The lack of resolution sensitivity in the parametrized orographic gravity wave momentum fluxes can be explained by examining the formulation of the orographic drag parametrization and how it partitions the drag into contributions from low-level blocking and gravity wave generation. In particular, the formulation of the scheme is such that the parametrized momentum flux is insensitive to resolution when the flow is in the partially blocked flow regime (see Vosper (2015) for further details of the formulation). This is due to the fact that, in this regime, the gravity wave amplitude is limited to the effective mountain height (h_{eff}), i.e. the proportion of the mountain that is able to generate waves. In the UM, this is given by $h_{eff} = min(h, \frac{U}{NF_c})$, where h is the height of the sub-grid mountains, U and N are the wind speed and stability averaged over the height of the sub-grid mountains and F_c is the critical Froude number, a parameter currently set to 4 in the UM. The depth of the blocked layer is given by $Z_b = h - h_{eff}$. In the IFS, the formulation is similar but Z_b is computed through iteratively solving the inequality $\int_{Z_{blk}}^{h} \frac{N(z)}{U(z)} dz \geq H_{crit}$ for Z_{blk}, where H_{crit} is some critical value and N and U vary with height rather than being depth averaged. The expression for the gravity wave surface stress is proportional to h_{eff}^2 and in the direction of the surface wind it is of the form:

$$\tau_{gwd} = \frac{1}{4} \rho_s U_s N_s \frac{\alpha}{\sigma} h_{eff}^2 GB \tag{5}$$

where ρ_s, U_s and N_s are near-surface averaged density, winds and stability, respectively. α is the mean slope of the sub-grid orography, σ is the standard deviation of the sub-grid orography, G is a constant tuning coefficient and B is a function of the subgrid mountain anisotropy. On inspection of the expression for h_{eff} it becomes evident that, in the case where the flow is partially blocked ($h > \frac{U}{NF_c}$), $h_{eff} = \frac{U}{NF_c}$ and the expression for the gravity wave momentum flux becomes independent of the sub-grid mountain height. Consequently, if the flow is frequently in a blocked flow regime over most of the domain and the depth-averaged near-surface winds and stability do not exhibit significant resolution dependence, then the gravity wave momentum flux can only vary with resolution through α/σ. On the other hand, the blocked layer depth (Z_b) is still able to scale with the sub-grid mountain height. The sensitivity of the surface stress and gravity wave momentum flux to σ and α fields used in global NWP models are demonstrated in Elvidge et al. (2019), and the transition of gravity wave stress between sub-critical ($h < \frac{U}{NF_c}$) and super-critical ($h \geq \frac{U}{NF_c}$) flow regimes is discussed and illustrated in their Figure 8c.

Figure 7(a) shows the ratio of the domain that is in a blocked flow regime on average over the period of integration, i.e. the percentage of the domain over which $h > \frac{U}{NF_c}$, where $\frac{U}{NF_c}$ is a time averaged quantity. Evidently, the flow is in a blocked regime over the majority of the domain considered. Figures 7(a) also shows that U/N does not change substantially with model resolution. This implies that the parametrized momentum fluxes can only scale with resolution via α/σ, which is plotted as a function of grid spacing in Figure 7(b). Comparing this with the change in gravity wave momentum flux across resolutions, shown in figure 5(b), it is clear that the variation of the parametrized

gravity wave momentum fluxes with resolution is reflected in α/σ. Note that the parametrized orographic blocking stress is still able to scale with the sub-grid mountain height, which explains why we see a monotonic increase in the parametrized blocking stress at coarser resolutions. It is likely that a value of $F_c = 4$ leads to a too frequent occurrence of flow blocking, which generally increases with coarser resolution (see Figure5(b)). This would be consistent with the fact that the blocking surface stress increases too rapidly with coarser resolution. However, in order to fully understand this behaviour it is necessary to consider the resolution sensitivity and the physical meaning of α/σ, delineated below.

The expression for the gravity wave surface stress vector in a zonal flow that is constant with height is given in spectral space by (Phillips, 1984; Teixeira, 2014):

$$\vec{\tau} = A^{-1} 4\pi^2 \rho N U \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{k\vec{k}}{|\vec{k}|} |\hat{h}(k,l)|^2 dk\, dl \tag{6}$$

where A is the area of the domain, $\vec{k} = (k, l)$ is the wavenumber vector and $\hat{h}(k, l)$ is the two-dimensional Fourier transform of the orography. The parametrization assumes that the subgrid orography is made up of several elliptical mountains each given by $h(x, y) = \frac{h}{(1+(x/a)^2+(y/b)^2)^{3/2}}$, where h is the maximum height of the mountain and a and b are the half-widths in the x and y direction, respectively. In a flow aligned with the x-axis, the stress in the x direction for one single elliptical mountain becomes:

$$\tau_x = A^{-1} \rho N U b h^2 G B \tag{7}$$

To determine the total stress from several elliptical mountains, the expression is multiplied by $\frac{A}{2a2b}$, i.e. the number of mountains in the grid box. This gives the expression:

$$\tau_x = \frac{1}{4a} \rho N U h^2 G B \tag{8}$$

The half-width a is then estimated as σ/α, i.e. the standard deviation of the subgrid orography divided by the mean slope. Upon substituting $a = \sigma/\alpha$, this becomes the expression for the parametrized gravity wave stress as given by Equation 5. σ/α is, therefore, an indication of the mean half-width of the subgrid mountains, calculated so as to determine the number of mountains within a grid-box. From Figure 7(b), we see that this mean half-width increases at lower resolutions, as may be expected from the fact that the sub-grid mountains are getting larger. The jump from 4km to 2km is due to the fact that a pre-filtering of the sub-grid orographic fields removes scales smaller than 6km. This is done so as to remove sub-grid variability that is not able to generate vertically propagating gravity waves. Smoothing of the small-scale slopes thus leads to large sub-grid length-scales relative to the 2km resolved orography. This is confirmed in Figure 7(b), in which the α/σ fields are also plotted with this pre-filtering removed. The removal of this filtering generally reduces the sub-grid wavelengths but also removes the jump between resolutions of 4km and 2km.

The factor $1/a = \alpha/\sigma$ is analogous to the zonal wavenumber ($k = 1/(2a)$) in Equation 6 if one considers a surface corrugation with only one single length scale of the mountain (Nappo, 2002). This monochromatic wave assumption is commonly made in gravity wave drag parametrizations, including those used in the UM and IFS. As the length scale of the orography (σ/α) increases at coarser resolutions (Figure 7(b)), the single wavenumber accounted for by the

parametrization scheme decreases. Combined with the fact that the resolved gravity wave momentum fluxes reduce at coarser resolutions, this results in a large decrease of the total gravity wave stress at coarse resolutions. Clearly, this formulation of the gravity wave drag does not allow the parametrized orographic gravity wave momentum flux to vary accurately with resolution and, by design, it cannot compensate for the loss off explicitly resolved momentum flux at coarser resolution. Other parametrizations such as those described by Kim and Doyle (2005) and Pierrehumbert (1986), which calculate a characteristic length-scale of the unresolved orography and use this to compute the number of sub-grid mountains within the grid-box, are likely to experience similar resolution sensitivities. Further justification for why this representation is inappropriate is discussed in the following section.

3.3 | Orographic gravity wave drag momentum flux scale contributions

So far, we have shown that the high resolution simulations can provide some constraint on the time-mean total and the variability of the parametrized orographic stress. Further insight into the behaviour of orographic gravity waves across different scales can be gained by performing a spectral analysis of the resolved orographic gravity wave momentum fluxes. This not only provides knowledge of which scales are being resolved but, consequently, also which scales need to be accounted for by the parametrization.

The co-spectra of the total momentum flux at 5km ASL are shown in Figure 8 for the various different horizontal resolutions. These were produced using the high pass filtered velocity perturbations, using the method described in Kruse and Smith (2015). The contributions are then computed in spectral space using:

$$\tau_{res} = |A^{-1} 4\pi^2 \rho_0 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\hat{u}^*, \hat{v}^*) \hat{w}\, dk\, dl|, \tag{9}$$

where $(\hat{.})$ is the two-dimensional Fourier transform and $(^*)$ denotes the complex conjugate. The two-dimensional momentum flux contributions are then binned into total wavenumber bands. Note that Figure 8(a) has a linear wavenumber axis, while 8(b) has a logarithmic axis and the flux is multiplied by the total wavenumber, $|\vec{k}|$, in order that the area under the curve is representative of the total momentum flux contribution from each wavenumber band (see Pope et al. (2000), their equation 2.262). The peak of the momentum flux occurs at a wavelength of about 400 km, which should be a large enough scale to be accurately resolved by the lower resolution models. While the larger scale waves tend to produce the largest amplitude momentum fluxes, they are only a narrow contribution towards the total wave spectrum (most notable in Figure 8(a)). In fact, the contribution from the high wavenumbers makes up a large part of the total resolved momentum flux at the higher resolutions. This is better illustrated in Figure 8(b), in which the area under the curve for scales smaller than ~80 km is approximately equal to that for scale larger than ~80 km at a resolution of 2 km. The increase in momentum flux at higher resolution is, therefore, as a result of increased momentum flux from smaller scales. This is also evident from Figure 5(b), in which the total resolved gravity wave momentum flux more than doubles between resolutions of 16km and 2km and increases sharply between 8km and 2km.

There is also no clear dominant wavenumber band at scales smaller than 100km in the high resolution simulations, with the tail of the spectrum plateauing in Figure 8(a). The momentum fluxes produced by wavelengths shorter than 100 km, but large enough to be hydrostatically balanced, therefore, carry somewhat equal momentum. This has interesting implications for the monochromatic wave assumption made in gravity wave drag parametrizations. These schemes implicitly assume that the small-scale waves produce a negligible amount of momentum flux compared with the dominant scale (given by α/σ). However, the broad spectrum of small-scale waves produced by the Rocky moun-

tains are demonstrably of significant importance to the total gravity wave momentum flux. The spectrum, along with Equation 6, suggest that, in order for the gravity wave drag scheme to behave 'seamlessly', it should account for the integral of the gravity wave momentum flux over all of the sub-grid scales rather than just the dominant scale.

Two further interesting observations from the spectra in Figure 8 can be made. Firstly, the reduction in the momentum flux at higher wavenumbers that is seen in the coarser resolution simulations occurs at much larger scales than the grid-scale. For example, in Figure 8(b) the momentum flux in the 32 km simulation begins to reduce relative to the 2 km simulations' at scales smaller than about 150 km. Similarly, the momentum fluxes in the 16 km simulations drop off rapidly at scales smaller than ~100 km, the 8km simulations drop off at scales smaller than ~60 km and the 4 km simulations drop off at scales smaller than ~50 km. That is roughly between 5 and 8 times the grid scale, which is consistent with the effective resolution estimated from other studies (e.g. Davies and Brown, 2001; Vosper et al., 2016). For the gravity wave drag parametrization to scale accurately with resolution it, therefore, needs to not only account for the correct sub-grid scales but also for those that are not fully resolved by the model dynamics, i.e. those 5 to 8 times the grid-scale.

The other point to note from Figure 8 is that the momentum flux contributions from the large wavelengths reduce as the grid size reduces. For example, comparing the momentum fluxes in the wavelength band 800 km to 200 km in the 32 km simulations with those in the 2 km simulations, there is a clear decrease in the amplitude with smaller grid spacing. This behaviour is unexpected, since one expects the large scales to be well resolved at all the resolutions considered and, thus, only changes in the small scales should be seen. Three reasons for why this may occur are:

(i) The small scale gravity waves lead to a change in the large-scale mean flow, causing a feedback onto the large-scale waves

(ii) At coarser horizontal resolutions, the model is not able to accurately represent some of the grid-scale waves and, as a result, they are aliased onto the smallest resolveable scale

(iii) The low-level drag is larger at coarser resolutions (as is evident from Figure 5(b)) which reduces the low level winds. This causes the flow to become more 'envelope following' as opposed to 'terrain following', as described below.

The phenomenon in (iii) was observed in the DEEPWAVE field campaign (Smith et al., 2016), during which the wave spectrum was observed to shift to larger scales during weak wind events and to smaller scales during strong wind events. Smith et al. (2016) attributed this to the flow being more able to penetrate into the valleys of the mountains during the strong wind events, thus being described as 'terrain following' and generating waves at smaller scales. On the other hand, the weak wind events led to pooling of air in the valleys, the flow skimming over the tops of the valleys and being described as 'envelope following'. The waves were, consequently, generated on larger scales.

While it is difficult to discount the first two hypotheses using our current model setup, the final hypothesis can be tested by switching off the parametrized orographic drag at the coarser resolutions. We expect that, as the surface winds increase with reduced drag, the flow should become more 'terrain following' and we should observe a decrease in momentum fluxes at the large wavelengths, accompanied by an increase at small wavelengths. Figure 8 shows the spectrum of the momentum fluxes in the 32km regional UM simulation with the parametrized orographic drag turned off. There is a clear decrease in the power at large wavelengths but no clear increase at small wavelengths. This may be as a result of the horizontal resolution limiting the generation of waves at small scales. Although these results are inconclusive, they do suggest that hypothesis (iii) is plausible and illustrate a strong interaction between the resolved gravity waves and the parametrized orographic drag, as was also found in van Niekerk et al. (2018).

4 | CONCLUSIONS

The work presented here demonstrates the significant value of high resolution simulations over complex terrain for understanding the accuracy of and deficiencies in orographic drag parametrizations. A comparison of drag components in low resolution simulations, in which a large proportion of the drag is parametrized, with those in higher resolution simulations, in which the orography and atmospheric processes are better resolved, has provided a constraint on the time-mean and the variability of orographic surface stress and gravity wave momentum flux. In the first part of this study, the drag components in two global models, the UM and the IFS, at approximately 16 km horizontal resolution were evaluated in this way. Both the mean and variability of the total orographic stress at the surface was found to compare well with the highest resolution simulations. Neither model exhibited a systematic over or under estimation of the total surface stress. The partitioning of the parametrized stress into boundary layer and orographic stress was found to vary greatly between the two models, with the UM having more parametrized orographic stress compared with boundary layer stress and the IFS having the opposite, corroborating previous studies (Sandu et al., 2016; Zadra et al., 2013). Although difficult to quantify definitively, the accuracy of the total surface stress during particular periods had an apparent dependence on the partitioning of the parametrized surface stress. This suggests that the different flow dependence between the boundary layer and orographic drag parametrization may have an impact on the forecast skill during specific atmospheric regimes.

An important aspect of the orographic gravity wave drag parametrization scheme is its ability to accurately deposit momentum in the lower stratosphere. The vertical distribution of the total orographic gravity wave stress and drag in the global UM and IFS was shown to compare well with the total in the 2 km simulations. However, an underestimation of the orographic gravity wave drag was identified in both the UM and the IFS. Moreover, the partitioning of the resolved and parametrized orographic gravity wave drag differed in the two models, with the UM having more parametrized orographic gravity wave drag in place of resolved gravity wave drag. The high level of accuracy of the parametrized orographic drag in the two models at a resolution of 16 km can be attributed to many iterations of tuning to minimise the forecast error in the winds, temperatures and surface pressure. This study has shown how high resolution simulations may be a means of avoiding the many iterations of tuning, in favour of a more direct approach to constraining the magnitude of the parametrized orographic drag.

The seamlessness of the orographic drag parametrization was then examined by comparing the various different components of the stress across resolutions. While the total time-mean surface stress was shown to be almost independent of resolution (apart from at the lowest resolution considered), the total gravity wave momentum flux was not and was found to decrease at lower resolutions. This was attributed to a lack of resolution sensitivity in the parametrized orographic gravity wave momentum flux. As the resolved flux decreased with coarser resolution, this was not compensated by an increase in parametrized flux. The formulation of the gravity wave drag parametrization is such that the parametrized flux is not able to increase at coarser resolution when the flow is in a blocked regime, which it frequently is during the period considered here. It is possible that a re-tuning of the parametrization may lead to improved scaling with resolution but we suggest that it is the monochromatic wave assumption made by the parametrization scheme that prohibits it from being appropriately sensitive to resolution. This assertion was reinforced by a spectral decomposition of the resolved momentum fluxes. The complex orography of the Rocky mountains was shown to produce a broad spectrum of waves carrying almost equal momentum across the smaller wavelengths, implying that the monochromatic wave assumption is too crude for this region. These findings are consistent with observed statistics from the DEEPWAVE campaign over New Zealand, where scales larger and smaller than 60 km were found to contribute the same amount to the total momentum flux on average (Smith and Kruse (2017), their Table 2).

A possible solution to this problem would be to formulate a scheme which accounts for all of the sub-grid scales. The natural choice would be to transform the sub-grid orography to spectral space, evaluate the momentum fluxes across all the wavenumbers and deduce the total stress via Equation 6. Such a scheme may be prohibitively expensive and, as a result, a simpler approach may need to be devised. Another way would be to assume some dependence of mountain height on mountain scale and to use this function in the integral of (6). A similar approach is taken in Garner (2005), in which an empirically derived relationship between the sub-grid mountain height and length-scale is used when computing the grid-box mean of the ensemble of sub-grid mountains. One could also imagine a scheme which uses the information about the dominant length-scale within the grid box that is currently used (namely α/σ) and assumes some drop-off in the momentum flux with horizontal scale, empirically derived from a high resolution momentum flux spectrum (e.g. Figure 8) or from the orographic spectrum. For example, one could assume a spectrum of the subgrid mountains that scales as $|\hat{h}|^2 \propto h|\vec{k}|^{-5/3}$, which is quite close to reality (Uhrner, 2001; Beljaars et al., 2004), and use equation (6) to derive an expression for the full spectrum of momentum fluxes. With appropriate integral limits, this integral can be expressed in terms of the highest wavenumber accounted for by the parametrization (i.e. $k_{max} = \frac{2\pi}{6km}$, thus removing wave scales close to the buoyancy frequency) and the lowest wavenumber that is appropriate for each grid-box at a given resolution (i.e. $k_{min} = \frac{2\pi\alpha}{\sigma}$). The evaluation of this integral is non-trivial in two-dimensions but in one-dimension it is:

$$\tau \propto A^{-1} 2\pi\rho NUh \int_{-\infty}^{\infty} k^{-2/3} dk = A^{-1} 2\pi\rho NUh[3k^{1/3}]_{k_{min}}^{k_{max}} \tag{10}$$

As the resolution becomes coarser, k_{min} decreases but k_{max} remains the same and, as a result, the expression for the stress becomes larger. A formulation of this kind would then, at least, provide some resolution sensitivity of the orographic gravity wave parametrization that accounts for the spectrum of waves generated by the topography. In the case of a blocked flow regime, the amplitude of the stress could then also be limited to the effective mountain height, i.e. $h = h_{eff} = \left(\frac{U}{NF_c}\right)^2$. Other aspects of the monochromatic wave assumption, such as the variable horizontal direction of the wave vectors and vertical group speed, could also be relaxed but, for simplicity, one could assume that they were equal for all wavenumbers.

Other interesting features of the gravity wave momentum flux spectrum were also identified: the effective resolution of the model is roughly 5 to 8 times the grid-length; and at increasing resolutions the momentum fluxes from large-scale gravity waves reduced. It is unclear exactly why this reduction at the large scales occurs but tests suggest that this may be due to the interaction with parametrized drag, perhaps as a result of the additional low-level drag at coarser resolutions. Current parametrizations assume some scale separation between what is resolved and parametrized. The change in the momentum flux contributions from the large scales as the resolution increases exemplifies the difficulties associated with this assumption, i.e. the 'greyzone' problem. While the scales of the momentum fluxes at the different resolutions may help to inform the parametrization schemes, they are also of interest in themselves. This is because the direction and speed of propagation of the waves can depend not only on the background flow but also on the scale of the waves. For example, shorter horizontal wavelength waves propagate faster in the vertical (Kruse and Smith, 2018) and are more readily refracted (Amemiya and Sato, 2016). This means that, although not investigated here, the scales of the resolved waves could have an impact on the large-scale atmospheric circulation.

acknowledgements

The authors would like to thank the developers of the Unified Model Nesting Suite, Stuart Webster and Chris Short. We also thank Steve Garner and one anonymous reviewer for their careful readings and suggestions that helped to improve the manuscript.

references

Alexander, M. J. and Barnet, C. (2007) Using satellite observations to constrain parameterizations of gravity wave effects for global models. *Journal of the Atmospheric Sciences*, **64**, 1652–1665. URL: https://doi.org/10.1175/JAS3897.1.

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**, 237–256.

Beljaars, A. C. M., Brown, A. R. and Wood, N. (2004) A new parametrization of turbulent orographic form drag. *Q. J. R. Meteorol. Soc.*, **130**, 1327–1347. URL: http://doi.wiley.com/10.1256/qj.03.73.

Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J. and Shelly, A. (2012) Unified modeling and prediction of weather and climate: A 25-year journey. *Bulletin of the American Meteorological Society*, **93**, 1865–1877. URL: https://doi.org/10.1175/BAMS-D-12-00018.1.

Carissimo, B. C., Pierrehumbert, R. T. and Pham, H. L. (1988) An estimate of mountain drag during alpex for comparison with numerical models. *Journal of the Atmospheric Sciences*, **45**, 1949–1960. URL: https://doi.org/10.1175/1520-0469(1988)045<1949:AEOMDD>2.0.CO;2.

Davies, L. A. and Brown, A. R. (2001) Assessment of which scales of orography can be credibly resolved in a numerical model. *Quarterly Journal of the Royal Meteorological Society*, **127**, 1225–1237. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712757405.

ECMWF (2015) Part iv: Physical processes. *IFS Documentation CYXXX*. URL: https://www.ecmwf.int/node/9204.

Elvidge, A. D., Sandu, I., Wedi, N., Vosper, S. B., Zadra, A., Boussetta, S., Bouyssel, F., van Niekerk, A., Tolstykh, M. A. and Ujiie, M. (2019) Significant uncertainty in the representation of orography in numerical weather prediction and implications for atmospheric drag and circulation. *Journal of Advances in Modeling Earth Systems*, **0**. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001661.

Fritts, D. C., Smith, R. B., Taylor, M. J., Doyle, J. D., Eckermann, S. D., Dörnbrack, A., Rapp, M., Williams, B. P., Pautet, P.-D., Bossert, K., Criddle, N. R., Reynolds, C. A., Reinecke, P. A., Uddstrom, M., Revell, M. J., Turner, R., Kaifler, B., Wagner, J. S., Mixa, T., Kruse, C. G., Nugent, A. D., Watson, C. D., Gisinger, S., Smith, S. M., Lieberman, R. S., Laughman, B., Moore, J. J., Brown, W. O., Haggerty, J. A., Rockwell, A., Stossmeister, G. J., Williams, S. F., Hernandez, G., Murphy, D. J., Klekociuk, A. R., Reid, I. M. and Ma, J. (2016) The deep propagating gravity wave experiment (deepwave): An airborne and ground-based exploration of gravity wave propagation and effects from their sources throughout the lower and middle atmosphere. *Bulletin of the American Meteorological Society*, **97**, 425–453. URL: https://doi.org/10.1175/BAMS-D-14-00269.1.

Garner, S. T. (2005) A topographic drag closure built on an analytical base flux. *Journal of the Atmospheric Sciences*, **62**, 2302–2315. URL: https://doi.org/10.1175/JAS3496.1.

— (2018) Ground-truth model evaluation of subgrid orographic base-flux parameterization. *Journal of the Atmospheric Sciences*, **75**, 3653–3670. URL: https://doi.org/10.1175/JAS-D-17-0368.1.

Kim, Y.-j. and Doyle, J. D. (2005) Extension of an orographic-drag parametrization scheme to incorporate orographic anisotropy and flow blocking. *Quarterly Journal of the Royal Meteorological Society*, **131**, 1893–1921. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.04.160.

Kruse, C. G. and Smith, R. B. (2015) Gravity wave diagnostics and characteristics in mesoscale fields. *Journal of the Atmospheric Sciences*, **72**, 4372–4392. URL: https://doi.org/10.1175/JAS-D-15-0079.1.

— (2018) Nondissipative and dissipative momentum deposition by mountain wave events in sheared environments. *Journal of the Atmospheric Sciences*, **75**, 2721–2740. URL: https://doi.org/10.1175/JAS-D-17-0350.1.

de La Cámara, A., Lott, F., Jewtoukoff, V., Plougonven, R. and Hertzog, A. (2016) On the gravity wave forcing during the southern stratospheric final warming in lmdz. *Journal of the Atmospheric Sciences*, **73**, 3213–3226. URL: https://doi.org/10.1175/JAS-D-15-0377.1.

Lindzen, R. S. (1981) Turbulence and stress owing to gravity wave and tidal breakdown. *Journal of Geophysical Research: Oceans*, **86**, 9707–9714. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JC086iC10p09707.

Lott, F. and Miller, M. J. (1997) A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Q. J. R. Meteorol. Soc.*, **123**, 101–127. URL: http://adsabs.harvard.edu/cgi-bin/nph-data{_}query?bibcode=1997QJRMS.123..101L{\&}link{_}type=EJOURNAL{\%}5Cnpapers3://publication/doi/10.1002/qj.49712353704.

Nappo, C. (2002) *An Introduction to Atmospheric Gravity Waves*. International Geophysics. Elsevier Science. URL: https://books.google.co.uk/books?id=I6KDVUU3myoC.

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**, 2527–2547. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001417.

van Niekerk, A., Scinocca, J. F. and Shepherd, T. G. (2017) The modulation of stationary waves, and their response to climate change, by parameterized orographic drag. *Journal of the Atmospheric Sciences*, **74**, 2557–2574. URL: https://doi.org/10.1175/JAS-D-17-0085.1.

Osprey, S. M., Gray, L. J., Hardiman, S. C., Butchart, N., Bushell, A. C. and Hinton, T. J. (2010) The climatology of the middle atmosphere in a vertically extended version of the met office's climate model. part ii: Variability. *Journal of the Atmospheric Sciences*, **67**, 3637–3651. URL: https://doi.org/10.1175/2010JAS3338.1.

Phillips, D. S. (1984) Analytical surface pressure and drag for linear hydrostatic flow over three-dimensional elliptical mountains. *J. Atmos. Sci.*, **41**, 1073–1084. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1984)041{\%}3C1073:ASPADF{\%}3E2.0.CO{\%}3B2.

Pierrehumbert, R. (1986) An essay on the parameterization of orographic gravity wave drag. In *Seminar/Workshop on Observation, Theory and Modelling of Orographic effects. Seminar: 15-19 September 1986, Workshop: 19-20 September 1986*, vol. 1, 251–282. ECMWF, Shinfield Park, Reading: ECMWF. URL: https://www.ecmwf.int/node/11673.

Pithan, F., Shepherd, T. G., Zappa, G. and Sandu, I. (2016) Climate model biases in jet streams, blocking and storm tracks resulting from missing orographic drag. *Geophys. Res. Lett.*, **43**, 7231–7240.

Plougonven, R., Hertzog, A. and Guez, L. (2013) Gravity waves over antarctica and the southern ocean: consistent momentum fluxes in mesoscale simulations and stratospheric balloon observations. *Quarterly Journal of the Royal Meteorological Society*, **139**, 101–118. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.1965.

Polichtchouk, I., Shepherd, T. G., Hogan, R. J. and Bechtold, P. (2018) Sensitivity of the brewer–dobson circulation and polar vortex variability to parameterized nonorographic gravity wave drag in a high-resolution atmospheric model. *Journal of the Atmospheric Sciences*, **75**, 1525–1543. URL: https://doi.org/10.1175/JAS-D-17-0304.1.

Pope, S., Pope, S., Eccles, P. and Press, C. U. (2000) *Turbulent Flows*. Cambridge University Press. URL: https://books.google.co.uk/books?id=HZsTw9SMx-0C.

Preusse, P., Eckermann, S. D. and Ern, M. (2008) Transparency of the atmosphere to short horizontal wavelength gravity waves. *Journal of Geophysical Research: Atmospheres*, **113**. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD009682.

Sandu, I., Bechtold, P., Beljaars, A., Bozzo, A., Pithan, F., Shepherd, T. G. and Zadra, A. (2016) Impacts of parameterized orographic drag on the northern hemisphere winter circulation. *Journal of Advances in Modeling Earth Systems*, **8**, 196–211. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015MS000564.

Sandu, I., van Niekerk, A., Shepherd, T. G., Vosper, S. B., Zadra, A., Bacmeister, J., Beljaars, A., Brown, A. R., Dörnbrack, A., McFarlane, N., Pithan, F. and Svensson, G. (2019) Impacts of orography on large-scale atmospheric circulation. *Nature Partners Journal Climate and Atmospheric Science*, **2**. URL: https://doi.org/10.1038/s41612-019-0065-9.

Shutts, G. and Broad, A. (1993) A case study of lee waves over the lake district in northern england. *Quarterly Journal of the Royal Meteorological Society*, **119**, 377–408. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711951102.

Smith, R. B. (1978) A measurement of mountain drag. *Journal of the Atmospheric Sciences*, **35**, 1644–1654. URL: https://doi.org/10.1175/1520-0469(1978)035<1644:AMOMD>2.0.CO;2.

Smith, R. B. and Kruse, C. G. (2017) Broad-spectrum mountain waves. *Journal of the Atmospheric Sciences*, **74**, 1381–1402.

Smith, R. B., Nugent, A. D., Kruse, C. G., Fritts, D. C., Doyle, J. D., Eckermann, S. D., Taylor, M. J., Dörnbrack, A., Uddstrom, M., Cooper, W., Romashkin, P., Jensen, J. and Beaton, S. (2016) Stratospheric gravity wave fluxes and scales during deepwave. *Journal of the Atmospheric Sciences*, **73**, 2851–2869. URL: https://doi.org/10.1175/JAS-D-15-0324.1.

Smith, S. A., Doyle, J. D., Brown, A. and Webster, S. (2005) Sensitivity of resolved mountain drag to model resolution for map case studies. *Hrvat. Meteoroloski Cas.*, **132**, 264–267.

Teixeira, M. A. C. (2014) The physics of orographic gravity wave drag. *Frontiers in Physics*, **2**, 43. URL: https://www.frontiersin.org/article/10.3389/fphy.2014.00043.

Uhrner, U. (2001) The impact of new sub-grid scale orography fields on the ecmwf model. *ECMWF*, 49. URL: https://www.ecmwf.int/node/12862.

Vosper, S. B. (2015) Mountain waves and wakes generated by south georgia: Implications for drag parametrization. *Q. J. R. Meteorol. Soc.*, **141**, 2813–2827.

Vosper, S. B., Brown, A. R. and Webster, S. (2016) Orographic drag on islands in the nwp mountain grey zone. *Q. J. R. Meteorol. Soc.*

Vosper, S. B. and Mobbs, S. D. (1997) Measurement of the pressure field on a mountain. *Quarterly Journal of the Royal Meteorological Society*, **123**, 129–144. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712353705.

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K. and Zerroukat, M. (2019) The met office unified model global atmosphere 7.0/7.1 and jules global land 7.0 configurations. *Geoscientific Model Development*, **12**, 1909–1963. URL: https://www.geosci-model-dev.net/12/1909/2019/.

Wood, N. and Mason, P. (1993) The pressure force induced by neutral, turbulent flow over hills. *Q. J. R. Meteorol. Soc.*, **119**, 1233–1267.

Wright, C. J., H. N. P. H. L. A. M. J. and Mitchell, N. J. (2017) Exploring gravity wave characteristics in 3-d using a novel s-transform technique: Airs/aqua measurements over the southern andes and drake passag. *Atmospheric Chemistry and Physics*, 8553–8575. URL: https://doi.org/10.5194/acp-17-8553-2017.

Zadra, A., Bacmeister, J., Bouyssel, F., Brown, A., Lock, A., Figueroa, S. N., Innocentini, V., Nakagawa, M., Roff, G. and Tolstykh, M. (2013) WGNE Drag Project, http://collaboration.cmc.ec.gc.ca/science/rpn/drag_project/. URL: http://collaboration.cmc.ec.gc.ca/science/rpn/drag{_}project/.

TABLE 1 The time-mean (middle row) and temporal correlation with the regional 2 km UM (bottom row) of the total surface stress in the global 16 km IFS, global 16 km UM, regional 32 km UM and regional 2 km UM.

	Global IFS	Global UM	32 km UM	2 km UM
τ_{tot} (Pa)	0.268	0.240	0.279	0.265
$corr[\tau_{tot}, \tau_{tot}(2km)]$	0.822	0.912	0.966	1.00

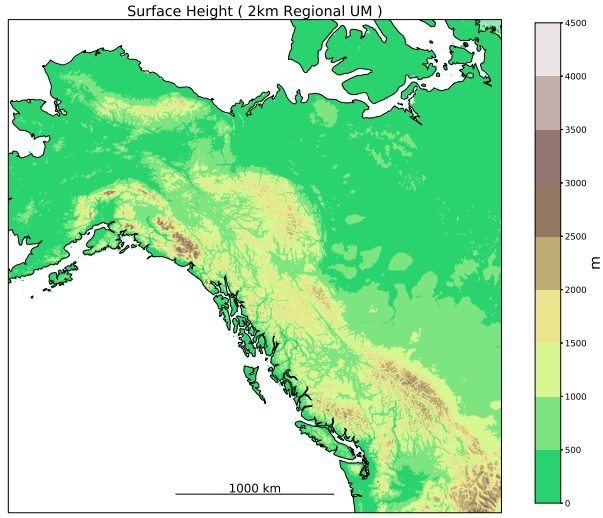

FIGURE 1 Domain used for global and regional model evaluation. Colour shading shows model orographic height in the 2 km regional UM simulation. To illustrate the scale, a distance of 1000 km is shown.

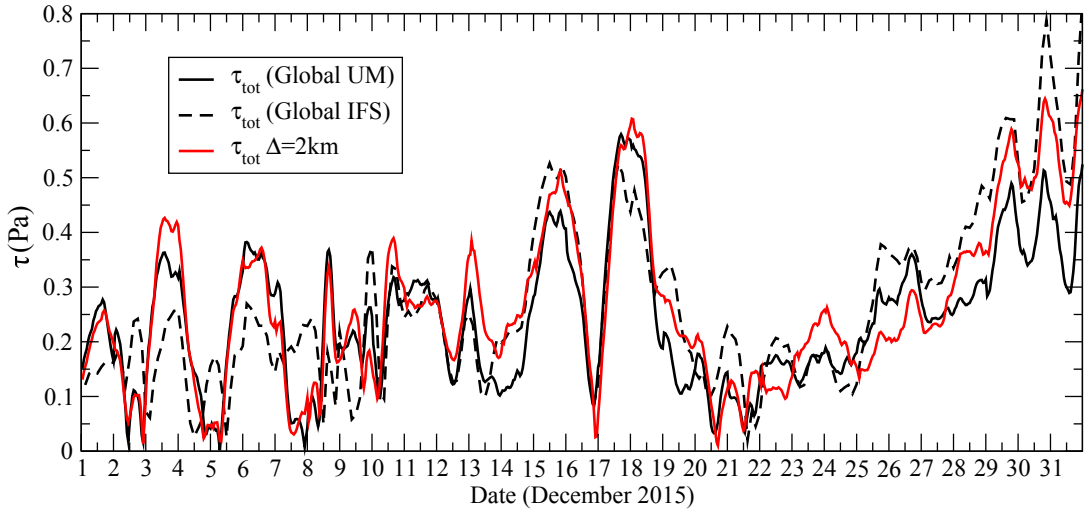

FIGURE 2 Time series of the total surface stress τ_{tot} in the global 16km UM (solid black), global 16 km IFS (dashed black) and regional 2 km UM (solid red).

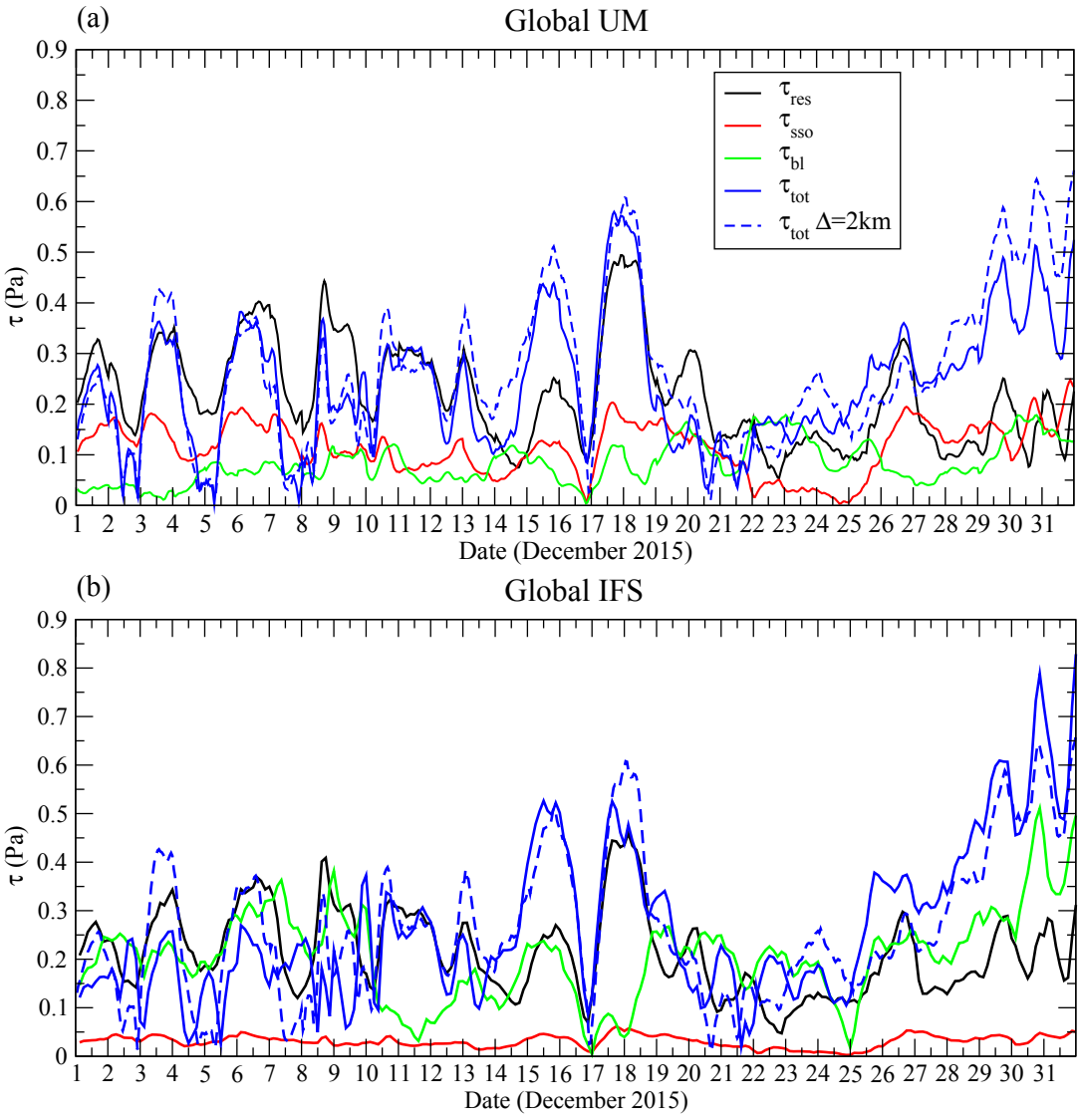

FIGURE 3 Time series of the resolved surface stress τ_{res} (solid black), parametrized orographic surface stress τ_{sso} (solid red), parametrized boundary layer surface stress τ_{bl} (solid green) and total surface stress τ_{tot} (solid blue) for (a) global 16 km UM and (b) global 16 km IFS. The total surface stress for the regional 2 km UM, $\tau_{tot}\Delta = 2km$, is also shown in both panels (dashed blue).

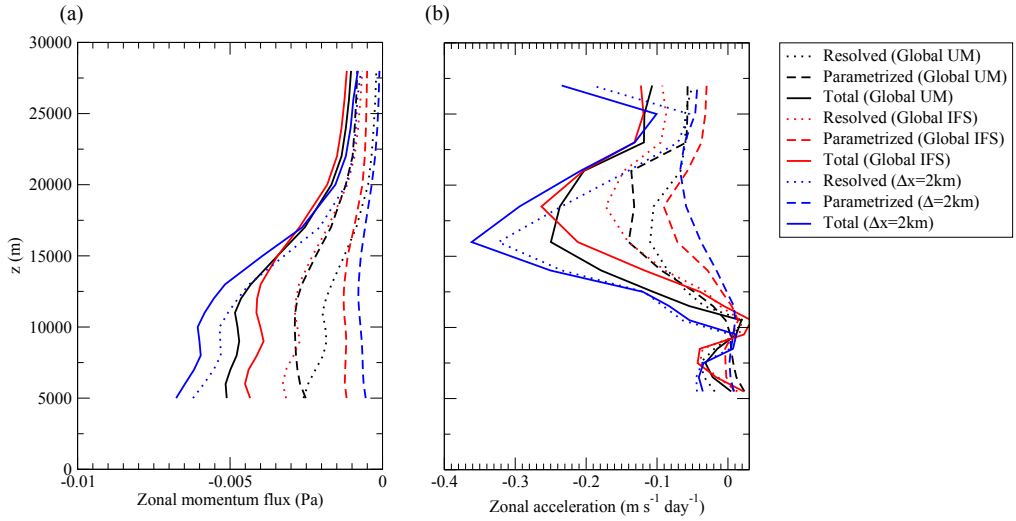

FIGURE 4 Time mean vertical profile of resolved (dotted), parametrized (dashed) and total (solid) (a) vertical flux of zonal momentum (see Equation 3) and (b) vertical divergence of zonal momentum flux (see Equation 4) in the global 16km UM (black), global 16km IFS (red) and regional 2km UM (blue).

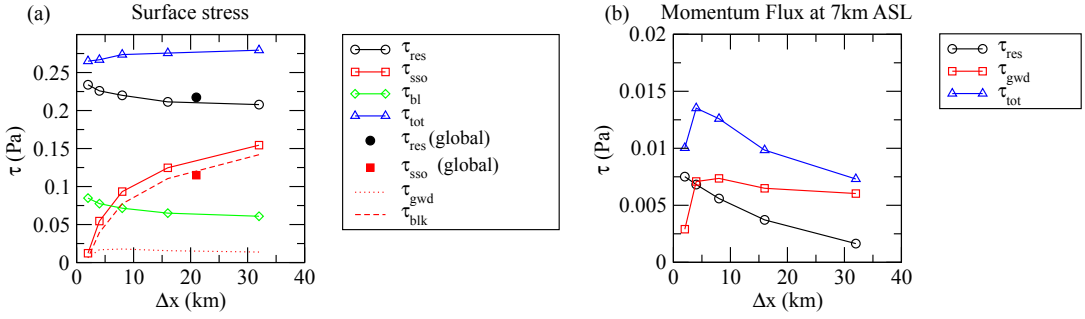

FIGURE 5 (a) Variation with horizontal resolution of τ_{res} (black open circles), τ_{sso} (red open squares), τ_{bl} (green open diamonds), τ_{tot} (blue open triangles), τ_{gwd} and τ_{blk} at the surface for the regional UM simulations. Resolved surface stress (black filled circle) and parametrized orographic drag (red filled square) for the 22km global driving model is also shown. (b) Variation with horizontal resolution of τ_{res} (black open circles), τ_{gwd} (red open squares) and τ_{tot} (blue open triangles) at 7 km ASL the regional UM simulations.

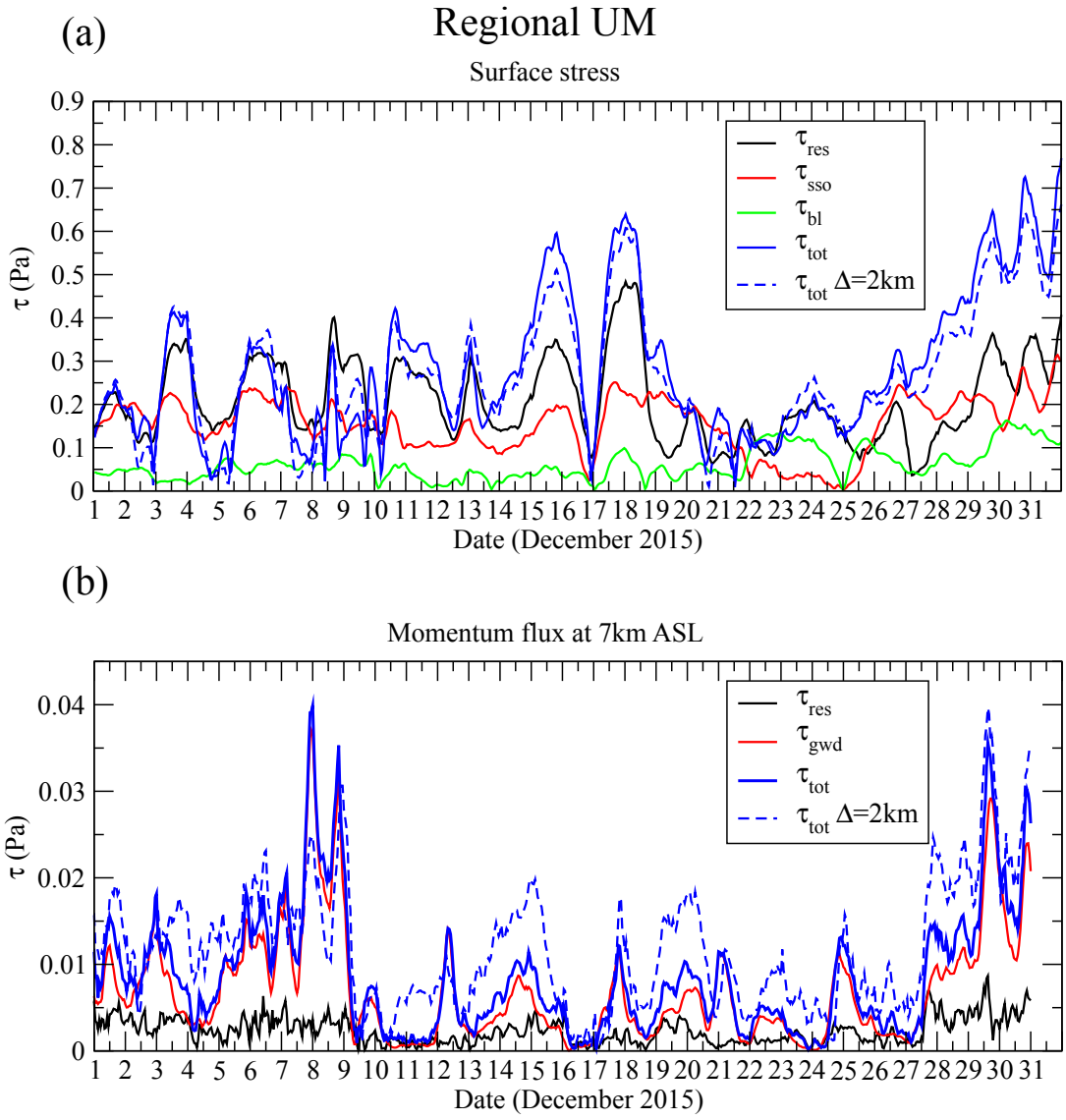

FIGURE 6 (a) Time series of τ_{res} (solid black), τ_{sso} (solid red), τ_{bl} (solid green) and τ_{tot} (solid blue) at the surface for the 32km regional UM simulations. τ_{tot} at the surface is shown for the 2km regional UM (dashed blue). (b) Time series of τ_{res} (solid black), τ_{gwd} (solid red) and τ_{tot} (solid blue) at 7 km ASL for the 32km regional UM simulations. τ_{tot} at 7 km ASL is shown for the 2km regional UM (dashed blue).

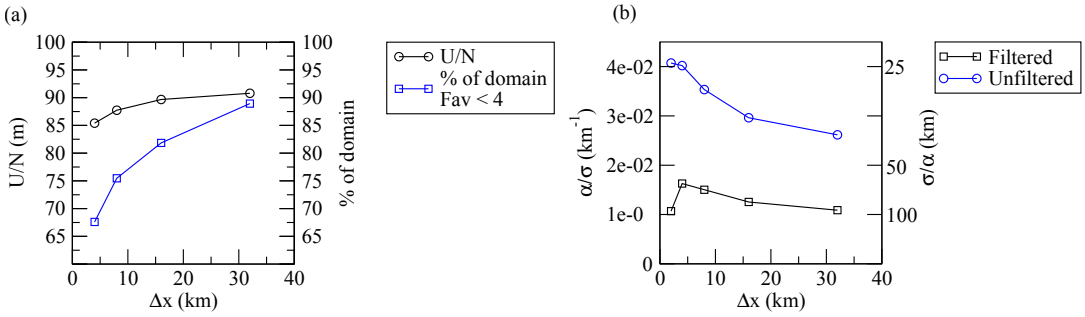

FIGURE 7 (a) Variation with horizontal resolution of the depth averaged U/N over the Rockies domain (black open circles), with values shown on left hand axis, and the percentage of the domain that is in a blocked flow regime on average during the period of simulation (blue open squares), with values shown on right hand axis. i.e. % of domain that is in blocked flow regime is the % of domain in which $h \geq \frac{U}{NF_c}$, see section 3.2 for details. (b) Variation with horizontal resolution of α/σ with standard pre-filtering of the source orography (black open squares) and without the pre-filtering (blue open circles) averaged over the Rockies domain. The inverse, σ/α, is shown on the right-hand axis to aid interpretation of the half-width a of the sub-grid orography.

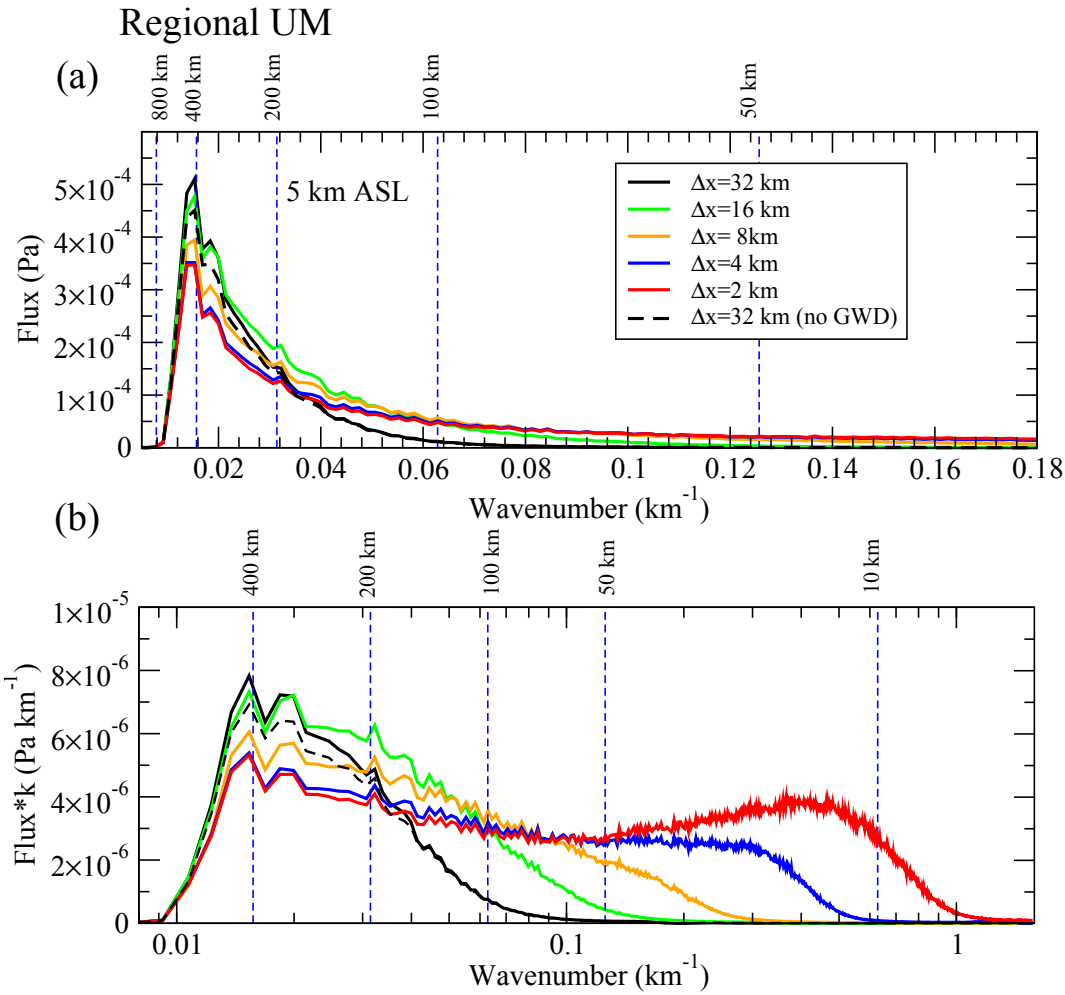

FIGURE 8 Spectral contributions to the resolved momentum flux as a function of total horizontal wavenumber ($|\vec{k}|$) at 5km ASL. (a) Flux against $|\vec{k}|$ using a linear x-axis. (b) Flux multiplied by $|\vec{k}|$ using a log x-axis. In both (a) and (b) the area under the curves represent the contribution towards the momentum fluxes from a particular wavenumber band. The spectrum is shown for all the regional simulations at different horizontal resolutions, including the 32km simulation with parametrized orographic drag turned off, as given by the legend. A number of wavelength values are marked with dashed lines.