Elvidge Andrew David (Orcid ID: 0000-0002-7099-902X)
Sandu Irina (Orcid ID: 0000-0002-1215-3288)
Wedi Nils Peter (Orcid ID: 0000-0001-5324-2749)
Vosper Simon (Orcid ID: 0000-0002-1117-4351)
van Niekerk Annelize (Orcid ID: 0000-0002-4706-1295)
Tolstykh Mikhail (Orcid ID: 0000-0002-6908-1211)

Significant uncertainty in the representation of orography in numerical weather prediction and implications for atmospheric drag and circulation

Andrew D Elvidge[1,2], Irina Sandu[3], Nils Wedi[3], Simon B Vosper[1], Ayrton Zadra[4],

Souhail Boussetta[3], François Bouyssel[5], Annelize van Niekerk[1], Mikhail A Tolstykh[6],

Masashi Ujiie[7]

[1] Met Office, Exeter, UK

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

[3] ECMWF, Reading, UK

[4] RPN, Environment and Climate Change Canada, Dorval, Quebec, Canada

[5] Météo France, Toulouse, France

[6] Institute of Numerical Mathematics Russian Academy of Sciences, and

Hydrometcentre of Russia, Moscow, Russia

[7] Japan Meteorological Agency, Tokyo, Japan

Correspondence:
Andrew D Elvidge
School of Environmental Sciences
University of East Anglia
Norwich Research Park
Norwich, UK
NR4 7TJ
a.elvidge@uea.ac.uk

Key Points

Differences in orography data fields are a principal cause of variation in atmospheric drag and circulation among weather and climate models

Abstract

The representation of orographic drag remains a major source of uncertainty for numerical weather prediction (NWP) and climate models. Its accuracy depends on contributions from both the model grid-scale orography (GSO) and the subgrid-scale orography (SSO). Different models use different source orography datasets and different methodologies to derive these orography fields. This study presents the first comparison of orography fields across several operational global NWP models. It also investigates the sensitivity of an orographic drag parameterisation to the inter-model spread in SSO fields and the resulting implications for representing the northern hemisphere winter circulation in a NWP model.

The inter-model spread in both the GSO and the SSO fields is found to be considerable. This is due to differences in the underlying source dataset employed and in the manner in which this dataset is processed (in particular how it is smoothed and interpolated) to generate the model fields. The sensitivity of parameterised orographic drag to the inter-model variability in SSO fields is shown to be considerable and dominated by the influence of two SSO fields: the standard deviation and the mean gradient of the SSO. NWP model sensitivity experiments demonstrate that the inter-model spread in these fields is of first-order importance to the inter-model spread in parameterised surface stress, and to current known systematic model biases. The revealed importance of the SSO fields supports careful reconsideration of how these fields are generated, guiding future development of orographic drag parameterisations and re-evaluation of the resolved impacts of orography on the flow.

Plain Language Summary

Mountains play a governing role in global atmospheric circulation via the aerodynamic drag they exert on the atmosphere. At smaller scales they influence winds and weather, for example instigating damaging downslope windstorms in their lee, generating winds which power onshore wind farms, and causing clear-air turbulence which affects commercial

aviation. Consequently, it is important that mountains (or 'orography') and their effects are represented accurately in global weather and climate models. Whilst broad mountains are well resolved by these models, smaller mountains and steep slopes are poorly resolved or unresolved. To approximate the drag exerted on the atmosphere by this 'subgrid-scale' orography (SSO), 'missing' hills or mountains are assumed in each grid box, whose height, steepness and shape are defined by data fields derived from the SSO. In this study, it is found that both model grid-scale and subgrid-scale orography fields vary significantly across currently operational models. These differences have a profound effect on the resultant drag, and consequently on the atmospheric circulation. The implication of these results is that changes in how orography is represented in our models have the capacity to bring significant improvements in our ability to model atmospheric circulations across a range of spatial and temporal scales.

1. Introduction

Accounting for the drag on the atmosphere from airflow over and around hills and mountains remains a major source of uncertainty in numerical weather prediction (NWP) and climate modelling. This orographic drag and the consequent momentum exchange plays a governing role in the atmosphere's general circulation on a wide range of spatial and temporal scales (see Smith, 1979 for a review). For example, large-scale mountains have an impact on the location of the mid-latitude jets (Brayshaw et al., 2009); small-scale mountains generate gravity waves that can break and exert influential drag forces on flow as high as the stratosphere and mesosphere (Bacmeister, 1993); and drag over all mountains is associated with local turbulent exchange processes and flow responses which notably affect the weather, climate and inhabitants of mountainous regions (e.g. downslope windstorms: Smith, 1985; foehn: Elvidge and Renfrew, 2016; rotors and boundary layer turbulence: Vosper et al., 2018).

Whilst global models are now of sufficient resolution to resolve large scale mountain waves (Elvidge et al., 2017), the drag exerted on the atmosphere due to subgrid-scale orography (SSO) remains significant and needs to be parameterised in weather and climate models. Several parameterisation schemes are thus used to represent different unresolved orographic drag processes, such as turbulent orographic form drag (TOFD) associated with orographic features with horizontal scales smaller than 5km (Beljaars et al. 2004), low-level flow blocking and gravity waves triggered by orographic features with horizontal scales larger than 5km (i.e. Lott and Miller, 1997; LM97 hereafter). As it is not currently possible to observe orographic drag at global, or even at regional scale, these parameterisations are poorly constrained, and their behaviour varies considerably between models. A model comparison undertaken by Zadra et al. (2013) for the World Meteorological Organization's Working Group for Numerical Experimentation (WGNE) found that over land, and particularly over orography, the parameterised orographic

surface stress can differ by a factor of four (in the zonal average) between NWP models of comparable resolution, while the total parameterised surface stress varied much less. Comparing the stresses between the Met Office Unified Model (MetUM) and the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS), for example, the contributions from different unresolved processes (turbulent drag in the planetary boundary layer, TOFD, low-level flow blocking or gravity wave breaking) differ considerably. For example, the zonally averaged parameterised orographic surface stresses in the MetUM are typically 2.5 to 4 times greater than those in the IFS, whilst the boundary layer surface stresses are typically 1.2 to 2 times smaller (see Figure 1 of Sandu et al., 2016). Elvidge (2019) also revealed significant differences between the MetUM and IFS in the diurnal and spatial variability of surface stresses over high mountain ranges. Sandu et al. (2016) demonstrated that inter-model differences in total parameterised surface stress and its partition between the different subgrid processes (such as those found between the MetUM and IFS) have a significant impact on the NH winter circulation both at daily and seasonal timescales.

Vosper (2015), Vosper et al. (2016) and van Niekerk et al. (2018) used high resolution (km-scale) simulations with the regional version of the MetUM over several mountainous regions (South Georgia, New Zealand, the Himalayas, and Middle East mountains) to evaluate the LM97 scheme used in the operational global MetUM for parameterising the orographic low-level blocking and gravity wave drag. The global MetUM is used for weather predictions and climate projections at horizontal resolutions ranging from tens to hundreds of km; for which these processes require parameterisation. The km-scale simulations, in which orographic drag due to blocking and gravity wave breaking become mostly resolved, demonstrated that while the LM97 scheme can qualitatively represent their effects on the flow, the parameter settings were not optimal and optimal settings varied from one region to another. A common finding

is that the blocking is too strong while the gravity wave breaking effect in the upper troposphere is too weak in the global MetUM. Van Niekerk et al. (2016) also demonstrated that the latitudinal distribution of resolved and parameterised surface stresses in the MetUM varies significantly depending on model resolution, and that the handover between resolved and parameterised surface stress as the model resolution changes is not necessarily handled well.

In summary, current orographic drag parameterisation schemes behave inconsistently across resolutions, across regions and mountain ranges, and across models; and the atmospheric circulation is highly sensitive to these inconsistencies from daily (Zadra et al., 2013, Sandu et al., 2016) to climate timescales (Scinocca et al., 2010, Pithan et al., 2016, van Niekerk et al., 2017). Clearly, orographic drag remains a significant challenge for weather and climate model development. This study aims to shed light on one particular source of uncertainty in the representation of orographic drag in models: the characterization of the grid-scale and subgrid-scale orography.

The first aim of our study is to evaluate model differences in the resolved grid scale orography (GSO) and the unresolved SSO fields. This will be done in Section 3, after describing the sources and processing of orography data in the six NWP models which employ comparable orographic drag parameterisation schemes (Section 2). Our second aim is to investigate to what extent the inter-model spread in SSO fields influences the parameterised orographic drag and ultimately the representation of the NH winter circulation in medium-range weather forecasts. Questions we are aiming to answer here are: (i) how sensitive is the parameterised surface stress to the inter-model spread in each SSO field and which fields are most important?; (ii) to what extent can the inter-model spread in SSO fields explain the inter-model spread in the total parameterised surface stress found by Zadra et al.

(2013)?; and (iii) how sensitive is the NH winter circulation to model differences in SSO fields on daily timescales? To answer the first question (i), we use an offline version of a commonly used orographic drag scheme (LM97 as implemented in the MetUM), forced with meteorological conditions typical of the NH winter. We run this scheme globally using as input various combinations of SSO fields from the six participant models and we examine the changes in the parameterised orographic surface stress (Section 4). To answer (ii) and (iii), we use one of the models (the IFS) to perform several sets of medium-range weather forecasts with different combinations of SSO fields from two of the models (the IFS and MetUM) and we examine the impact on the total parameterised surface stress, on its partition between boundary layer and orographic contributions, and on the large-scale circulation (Section 5).

2. The models and their representation of orography

2.1 The models

The names and key relevant characteristics of the six models used for this study, in their operational weather prediction configurations (as of 2016), are provided in Tables 1 and 2. Note that for the IFS, in addition to the operational version of the model (referred to hereafter as 'IFS Operational'), the GSO and SSO fields corresponding to a coarser TL1279 spectral resolution (global mean grid spacing of 16 km) version of the model (referred to hereafter as 'IFS 16 km') is included in the comparison. This is because the resolution used for the IFS's medium-range operational weather forecasts is significantly higher (at 9 km mean grid spacing) relative to the other centre's operational model resolutions (which vary between 16 and 25 km mean grid spacing).

2.2 The representation of orographic drag

A model's representation of orographic drag depends on the source orography dataset and on how this dataset is processed to generate the GSO and SSO fields; the model's dynamical core, grid type, numerics and resolution; and the orographic drag parameterisation schemes employed, and how these schemes have been tuned to optimise forecast skill or to constrain the model climate (Sandu et al. 2013, Hourdin et al. 2017). As discussed in Section 1, models commonly represent subgrid-scale orographic drag using more than one parameterisation scheme, reflecting the range of spatial scales and processes responsible. In this study, we only focus on the orographic drag exerted on stably stratified flow by mountains with horizontal scales larger than 5 km, through low-level blocking and higher-level gravity wave drag associated with the breaking of vertically propagating mountain gravity waves. Consequently, reference hereafter to orographic drag and SSO parameterisation refers only to blocking drag and gravity wave drag (and not to TOFD).

The MetUM, IFS and GDPS SSO drag schemes handle these two drag processes following LM97, who in turn derived their handling of gravity wave drag from Boer et al. (1984) and McFarlane et al. (1987). The ARPEGE and SL-AV schemes (see Catry et al., 2008 for a description) use LM97 for gravity wave drag, and an envelope orography approach (after Wallace et al., 1983) for blocking drag. The GSM's scheme follows Iwasaki et al. (1989), with separate treatments for longwave mountains (wavelengths > 100 km) and shortwave mountains (wavelengths ~ 10 km). The orographic fields derived for the longwave scheme are used for this intercomparison, since the scales accounted for by this scheme are more analogous to those accounted for by the LM97 schemes of the other models.

2.3 The grid-scale and subgrid-scale orography fields

The GSO and SSO fields are generated from the source orography dataset in the manner depicted and described in Figure 1. Note that whilst this methodology is generally consistent across the models, the exact order of the smoothing and interpolation procedures, as well as the strength of the smoothing filters, are model dependent.

The source data used by the models in this study are for the most part from three different digital elevation models (DEMs), all with a horizontal grid spacing of 30 arc seconds (~1 km at mid latitudes). These are the Shuttle Radar Topography Mission dataset (SRTM30) from The National Aeronautics and Space Administration; the Global Land One-km Base Elevation dataset (GLOBE) from The National Oceanic and Atmospheric Administration; and the Global 30 Arc-Second Elevation dataset (GTOPO30) from United States Geological Survey. Of these, GTOPO30 is the oldest, dating back to 1996; GLOBE is newer and incorporates GTOPO30 data amongst other sources; and SRTM30 (Farr et al., 2007) is newer still and was generally considered to be the highest quality freely available global DEM until recently (see Wessel et al., 2018).

The smoothing applied to the GSO (steps 1 & 2 in Figure 1) has been deemed necessary to avoid model grid-point instabilities over steep gradients and to prevent unrealistic gravity waves and aliasing effects (Webster et al., 2003). The filtering applied should entirely eliminate two-grid-length features, which cannot be resolved by models, whilst retaining features of length scales adequately resolved (this limit is model-specific; for the MetUM it is thought to be of order 6-8 grid lengths). The degree of GSO smoothing varies significantly between models, as does the strength of the filter applied to the pre-filtered source orography to generate the GSO (step 2 in Figure 1). This filter defines the orographic scales to be included in the subgrid-scale orography. For the MetUM, it is a scale selective Raymond

filter which eliminates features of scales twice the grid length (Δx) and dampens four-grid-length features by half; as such it is described as a '4-Δx filter'. Filter strength varies between the models in this study from ~2.3-Δx (dampening 2.3-grid-length features by half) for the IFS to ~7-Δx (dampening 7-grid-length features by half) for the SL-AV (Table 2). Moreover, as illustrated in Figure 2, the filters employed also vary significantly in their response functions[1]. The type and strength of orographic filter applied in each model reflect decisions related to the model's stability over steep terrain, forecast optimisation, and technical considerations. However, these decisions were often made many years ago, and may no longer be optimal.

The SSO is generated as the difference between the pre-filtered source orography and the GSO (see step 3 in Figure 1) and consequently describes the subgrid-scale terrain undulations with respect to the GSO. The SSO is then used to compute the "SSO fields" – model grid-scale variables which describe the statistical characteristics of the SSO. In the LM97 scheme, the SSO in each grid box is represented via an array of uniformly distributed, elliptical mountains (hereafter referred to as *parameterized mountains*), as depicted in step 4 of Figure 1. These *parameterized mountains* are defined entirely by the SSO fields.

The various orographic drag schemes used in the participating models use up to four SSO fields (see Table 2), denoted hereafter as *stdev*, *slope*, *anisotropy* and *orientation*. The *stdev* field is the standard deviation of the SSO, and in the LM97 scheme is multiplied by a model-dependent tuning constant to define the assumed height of *parameterized mountains*. The remaining three fields are derived from the grid box average gradients in the SSO (as described in the appendix of LM97). The *slope* field is a measure of grid-box average SSO gradients and defines the slope of *parameterized mountains* in the LM97 scheme. The

---

[1] The response function is the ratio of the output (filtered) orography spectrum to the input (unfiltered) orography spectrum, as a function of distance expressed in grid lengths.

a*nisotropy* field is a measure of the directional dependence of the grid scale orography, defined as the ratio of the minor to major axes of the SSO. *Anisotropy* varies between 0 (indicating anisotropic, "elongated" *parameterized mountains*) to 1 (indicating isotropic, "circular" *parameterized mountains*). Finally, the *orientation* field is the angle of the major axis of the SSO, describing *parameterized mountains* elongated along a North-South axis (0 degrees) to an East-West axis (±90 degrees). A +(-) 45 degree angle indicates mountains elongated along a NE-SW (NW-SE) axis.

For this study, each participating centre provided their GSO and SSO fields on a regular, cell-centred, latitude-longitude grid of spacing 0.25 degrees (~25 km) in both latitude and longitude. Whilst the MetUM, IFS and GDPS employ all four SSO fields in their respective orographic drag schemes, the ARPEGE and SL-AV employ all but the *slope* field, and the GSM employs only the *stdev* field. Note however that for this study the ARPEGE and SL-AV *anisotropy* fields are omitted from analysis for technical reasons.

As an example of global model orographic fields, Figure 3 shows those derived from the MetUM. The greatest *stdev* values are, unsurprisingly, found in regions of precipitous terrain, most notably the Southern edge of the Himalayas. Note that further North over the gentler slopes of the Himalayan plateau, the *stdev* values are generally lower. The *slope* field unsurprisingly exhibits strong correlation with the *stdev* field (global correlation coefficient of 0.96 for the MetUM). As evident in Figure 3, precipitous mountain regions such as the Himalayas and the Alps are generally characterised by greater a*nisotropy* values (i.e. more circular *parameterized mountains*) in the MetUM (global correlation coefficient of 0.21 with *stdev*). This likely reflects that, for flatter terrain (e.g. the dome-like Antarctic ice sheet), *anisotropy* is more strongly influenced by shallow, directional grid-scale slopes than for steep mountainous terrain. The global distribution of the *orientation* field is complex, though some

coherent patterns are discernible. For example, the Andes and Rockies generally exhibit *orientation* values indicative of North-South SSO ridges. This reflects the broadly North-South orientation of the continental plate boundaries along which the orogenic belts giving rise to these mountain ranges formed.

3. Model variability in orographic fields

In this section the inter-model spread in GSO and SSO fields is evaluated, with reference to Figures 4 to 6.

Figure 4 shows the variance of the GSO (scaled by the theoretical $k^{-5/3}$ law – see caption) as a function of the total wavenumber for all models. This spectrum is obtained by transforming the orography into spectral space, i.e. via decomposition into spherical harmonics. The spectrum as a function of total wavenumber is obtained by summing the squared coefficients over all meridional wavenumbers, which is representative of the variance of the field. For further details see for example Malardel and Wedi (2016), and their Figure 1. The variance in global GSO exhibits considerable inter-model divergence towards the smaller scales (high wave numbers). This reflects model differences in resolution and the degree of orographic smoothing applied. The GSO of the two IFS versions exhibit the highest powers at high wavenumbers. This is explained both by their relatively high resolution and by the relatively weak orographic smoothing employed by the IFS relative to the other participant models. At the other end of the scale, the SL-AV's GSO exhibits the weakest power at small scales, reflecting its relatively coarse resolution and acute smoothing.

Probability density functions (PDFs) for each of the orography fields and each model are displayed in Figure 5 for a representative mountain range; the North American Rockies

(domain shown in Figure 3). Focussing here on a mountain range as opposed to the globe helps to illustrate more distinctly the inter-model differences. Qualitatively similar results are found for another mountain range – the Himalayas – and for the globe as a whole (not shown). There is generally good agreement among the models in terms of the PDF of GSO, with no significant systematic differences between the models (Figure 5a). This shows that the models are broadly consistent in their representation of the Rockies. The differences in orographic variance at small-scales shown in Figure 4 are also evident in this figure; models with greater variance exhibiting greater probability densities at both the lowest and highest elevations (for example compare the PDFs for the IFS versions with the SL-AV). More distinct model differences are evident in the *stdev* field; the two IFS versions clearly exhibiting the lowest *stdev* values, followed by the GSM. This is consistent with the differences in small-scale GSO variance, which indicate that these modes resolve more real-world orographic variability than the other models, and consequently require less of this variability to be parameterised. The other models exhibit broadly similar PDFs of *stdev*.

In Figure 6, global mean *stdev* – a measure of the degree to which orography is parameterized or resolved – is plotted against the global mean resolution. As expected, for the higher resolution models, global mean *stdev* is generally lower. However, the relationship is far from linear or even monotonic. For example, the IFS 16 km has significantly lower mean *stdev* values than the MetUM and ARPEGE, despite being of similar resolution. This is likely due to the weaker smoothing filter applied when deriving the GSO, prior to computing the SSO fields, in the IFS (Figure 2). However, the coarsest resolution model – the GDPS – employs a smoothing filter of comparable strength to those used by the MetUM and ARPEGE, yet exhibits similar global mean *stdev*. Such apparent discrepancies are evidently a consequence of other model differences, for example differences in source orography dataset (see Figure 6), model grid type (in particular, the Yin-Yang grid used in the GDPS is notably unique

here), source orography pre-filtering, interpolation methodologies, approaches to avoid erroneously large *stdev* values associated with steep gradients in the source orography, and the algorithms used to derive the orographic fields. The relative importance of each of these differences is the subject of ongoing work. Preliminary results indicate that differences in both the source orography dataset and in the smoothing functions employed have significant impacts on the resultant model orography fields.

In addition to the inter-model spread in the *stdev* field, notable spread is also apparent in the other SSO fields. The *slope* field is only employed in the MetUM, IFS and GDPS orographic drag schemes, and – as previously mentioned – is closely correlated with *stdev*. Consequently, differences in the PDFs of *slope* qualitatively resemble those of *stdev* (compare Figure 5b and 5c). In the *anisotropy* field, there is a positive skew in PDFs for the MetUM and GDPS, whilst the IFS versions have closer to normal distributions. This indicates that the *parameterized mountains* in the MetUM and GDPS are generally more isotropic – dome-like as opposed to ridge-like – than in the IFS (Figure 5d). Finally, the previously mentioned predominantly N-S orientation of subgrid ridges in the Rockies range is clearly evident in the PDFs of *orientation*, with a peak in probability density in all models centred around ~0 ° (Figure 5e). However, there are significant differences in the distributions, with this peak being most pronounced in the ARPEGE and SL-AV, and a greater tendency towards *orientations* > 0° in the MetUM and GDPS. Again, the precise reasons for these differences are not clear and should be the subject of future work.

In summary, the comparison of the GSO and SSO fields used in the six models reveal a considerable inter-model spread in all fields. As the SSO fields are used as input for orographic drag parameterisations, we investigate hereafter to what extent the spread in the SSO fields translates into uncertainties in the representation of parameterised orographic

drag, and consequently of atmospheric circulation in NWP. The sensitivity of orographic drag and atmospheric circulation to differences in the GSO is not discussed here as it has been examined in previous work (Sandu et al., 2017; van Niekerk et al., 2018).

4. Sensitivity of parameterised orographic surface stress to subgrid orographic fields

In this section, we investigate the extent to which the inter-model spread in SSO fields affects the representation of parameterised orographic drag. This is done via a set of experiments with an 'offline' version of the commonly used LM97 scheme, in the form employed by the MetUM and using operational MetUM parameter settings (see Appendix, or Webster et al. (2013), for a description of the LM97 implementation in the MetUM).

The 'offline' LM97 scheme uses as input global SSO fields and key meteorological fields describing the flow incident on SSO (flow velocities, Brunt Vaisala frequencies and air densities), and yields as output momentum fluxes from both low-level flow blocking and gravity wave sources, and the corresponding surface stresses (the surface stress representing the vertical integral of the momentum flux derivative). To ensure the relevance of this analysis to simulations of real atmospheric conditions, the input meteorological fields have been gathered from instantaneous MetUM output at a lead time of 24 hr from global short-range forecasts initiated daily at 00UTC for the month of December 2016, at n768 resolution (grid spacing of ~16 km).

The 'offline' LM97 scheme is first run using the SSO fields of each model for which all four SSO fields are available (MetUM, IFS, GDPS). It is then run using all but one of the SSO fields from the MetUM and the remaining field from one of the other models. The scheme is run successively for each SSO field and each model. These experiments reveal the sensitivity

of the parameterised orographic surface stress to the inter-model variability in each of the SSO fields. All spatial averages discussed in this section include land model grid points only (omitting grid points over ocean).

The IFS Operational SSO fields lead to the lowest zonally averaged SSO surface stress magnitudes across all latitudes, with values typically 6 to 8 times smaller than those obtained for the MetUM and GDPS SSO fields (Figure 7a). The MetUM and GDPS SSO fields yield quite similar stresses across the latitude bands. Figures 7c-f illustrate the impact on the parameterised SSO surface stress due to varying the model source of the *stdev, slope, anisotropy* and *orientation* fields, respectively (note that missing lines reflect unavailable SSO fields in certain models – see Table 2). SSO surface stresses appear to be virtually insensitive to the model spread in *anisotropy* (Figure 7e) and *orientation* (Figure 7f). The inter-model spread in SSO surface stress (Figure 7a, for example IFS Operational *vs* MetUM) is almost entirely due to the differences in the *stdev* (Figure 7c) and *slope* fields (Figure 7d).

The stress sensitivities to each SSO field are explored further by analysing the response of the global mean parameterised SSO surface stress to the (normalised) inter-model variability in the global mean of each SSO field (Figure 8). Each point in Figure 8 corresponds to one of the SSO field substitution experiments. The x-axis shows the global mean of the substituted SSO field normalised such that an inter-model range for each field is linearly mapped to a consistent range on the axis (between "lowest" and "*control*"), while the y-axis shows the difference in global mean surface stress relative to that incurred by the *control* experiment. To facilitate a fair comparison of the impact of each SSO field, the inter-model range refers to the range only across those four models which incorporate all four SSO fields (namely the IFS Operational, IFS 16 km, MetUM and GDPS). The *control* experiment refers to that associated with the highest global mean value of the relevant SSO field across these four

models. Note that due to *orientation* being a circular quantity the global mean of its modulus (i.e. |*orientation*|) is used, which ranges from North-South orientated SSO (small values) to East-West orientated SSO (large values).

Figure 8 corroborates the results for the zonally averaged SSO surface stresses (Figure 7); i.e. there is considerable variability in the globally averaged SSO surface stress across the experiments, and the *stdev* and *slope* exert far greater influence than the *anisotropy* and *orientation*. This finding also holds in terms of the global standard deviation of the impact on parameterised SSO surface stresses obtained for the various experiments with respect to the *control* experiment (see 'whiskers' in Figure 8). There is a clear positive correlation between the globally averaged parameterised SSO surface stresses and both *stdev* and *slope*, with the generally lower, weaker sloped sub-grid orography of the IFS versions yielding the lowest stresses. The IFS's more anisotropic, ridge-like SSO (lower *anisotropy* values) results in *marginally* lower global mean stresses than the more isotropic, dome-like SSO used in the MetUM and GDPS. No correlation is discernible between global mean |*orientation*| and global mean SSO surface stress.

The latitude-weighted global mean parameterised SSO surface stress varies considerably; from 0.015 W m^{-2} when using SSO fields exclusively sourced from the IFS Operational to 0.115 W m^{-2} when using exclusively MetUM-sourced fields. 74 % of this difference can be accounted for by substituting only *slope* between the two models, whilst 73 % can be accounted for by substituting only *stdev*. Meanwhile, only 6 % and 1 % of this difference can be accounted for by substituting, respectively, only *anisotropy* and only *orientation*. Note these percentages do not add up to 100, indicating that the combined influence of all the SSO fields do not equal the sum of the individual influences of each SSO field. This is due to nonlinearities in the LM97 scheme equations – most notably the nonlinear influence of *stdev*

(discussed below and in the Appendix) – which act to modulate the influence of the other SSO fields.

The surface stresses from these experiments are predominantly due to low-level blocking (Figure 7b). Likewise, as evident from Figure 8, the sensitivity of surface stress to the SSO fields is dominated by that caused by blocking. Furthermore, the degree to which the blocking drag dominates over the gravity wave drag is also sensitive to the SSO fields. Figure 7b shows that the MetUM and GDPS SSO fields yield a greater dominance (typically 85-95 % due to blocking), whilst IFS Operational fields yield the weakest dominance (typically 60-80 % due to blocking). The reason for this sensitivity becomes apparent on inspection of Figure 8 and the LM97 equations (Appendix). As global means, both the low-level blocking drag (D_b) and the gravity wave drag (D_{gw}) exhibit a roughly linear, positive correlation with *slope* across the SSO substitution experiments (Figure 8b). Indeed, as evident from Appendix Equations 1, 2 and 4, $D_b \propto slope$ and $D_{gw} \propto slope$. Meanwhile, whilst D_b is positively correlated with *stdev* (albeit nonlinearly), D_{gw} varies non-monotonically with *stdev* (note the blue fitted curve having a parabolic shape in Figure 8c). Consequently, the generally lower *stdev* values in the IFS results in, relative to the MetUM and GDPS, a greater decrease in D_b than in D_{gw}. The nonlinear sensitivity of D_{gw} to *stdev* is briefly explored in the Appendix.

5. Sensitivity of parameterised surface stress and large-scale circulation to subgrid orographic fields

In this section, we explore the relative importance of the SSO fields for the representation of the total parameterised surface stress and the NH winter large-scale circulation on daily timescales. Several global weather forecast experiments have been conducted with the IFS 16

km (Tl1279 resolution). Each of these experiments consists of a set of 31 ten-day forecasts, initialized daily at 00 UTC throughout December 2016.

In the IFS$_{\text{IFS-SSO}}$ experiment, the IFS was run in its operational configuration (with the exception of the coarser resolution), with the corresponding IFS SSO fields. In a second IFS experiment (IFS$_{\text{MetUM-SSO}}$) the SSO fields were replaced by those from the MetUM (at n768; 17.0 km mean global grid spacing). Finally, five further IFS experiments used various combinations of SSO fields from both the IFS and MetUM. Note that the MetUM and IFS 16 km SSO fields are roughly representative of the extremes of the inter-model spread in key fields, so their respective influence (as revealed in these experiments) provide a valuable counterpoint for comparison. The IFS 16 km exhibits, on average, the smallest *stdev* and smallest *slope* values among the models of comparable resolution, whilst the MetUM exhibits on average relatively large *stdev* and the largest *slope* values.

5.1 Impacts on total parameterised surface stress and its partition between SSO and boundary layer components

We first examine the impact of using the various combinations of SSO fields on the zonally averaged total parameterised surface stress (τ_{total}), and on its constituent components: SSO stress (τ_{SSO}) and turbulent boundary layer stress (τ_{BL}) (Figure 9 a-c). One of the questions we are interested in is to what extent the inter-model spread in SSO fields can explain differences in the total surface stress and its components between the IFS and MetUM found by Zadra et al. (2013) and Sandu et al. (2016). For that purpose, we also use the zonally averaged values of τ_{total}, τ_{SSO} and τ_{BL} from the MetUM short range forecasts performed to derive the meteorological variables used to force the 'offline' LM97 scheme in section 4 (dark blue line in Figures 9 a-c). Recall these consisted of 31, daily, 24-hour forecasts at N768 resolution for December 2016.

In the IFS$_{\text{IFS-SSO}}$ experiment, the zonally averaged τ_{BL} is greater than τ_{SSO} across all latitudes, whilst in the MetUM experiment, the partition of stress between the two components is generally more evenly balanced, with the dominant component varying with latitude (Figure 9). In terms of τ_{total}, there is generally good agreement between the two models, except across the NH mid latitudes – the most mountainous latitudes – where MetUM stresses exceed those of the IFS$_{\text{IFS-SSO}}$ by approximately 15-20%, reflecting the findings of Zadra et al. (2013) and Sandu et al. (2016).

In the IFS$_{\text{MetUM-SSO}}$ experiment (red line in Figure 9a-c) τ_{SSO} is more than doubled across all latitude bands, relative to the IFS$_{\text{IFS-SSO}}$ experiment. Much of this increase in τ_{SSO} is offset by a compensating decrease (by typically 5-15 %) in τ_{BL}. This is a result of a general deceleration in low level winds (τ_{BL} being proportional to wind speed) resulting from the increase in τ_{SSO}. Despite this offset, the net effect on the total parameterised surface stress remains significant and positive, with increases in τ_{total} between 5 and 25 % across all latitudes. These changes bring the IFS into much closer agreement with the MetUM experiment in zonally averaged τ_{SSO}, τ_{BL} and τ_{total}. Clearly the SSO fields have a significant impact on the total parameterised surface stress and on its partition between the SSO and boundary layer components. And in this case these impacts explain a large part of the differences in parameterised surface stress – from both τ_{SSO} and τ_{BL} – between the IFS and MetUM. In terms of globally averaged, latitude-weighted stresses, 55 % of τ_{SSO} , 35 % of τ_{BL} and 80 % of τ_{total} differences between the IFS$_{\text{IFS-SSO}}$ (black line) and the MetUM (blue line) experiments can be explained by differences in SSO fields between the two models (i.e. are replicated in the differences between the IFS$_{\text{IFS-SSO}}$ and IFS$_{\text{MetUM-SSO}}$ (red line) experiments).

Remaining differences in parametrized surface stress between the IFS$_{\text{IFS-SSO}}$ and MetUM experiments are likely to be largely due to differences in orographic drag scheme parameter

settings, differences in the implementation of the LM97 scheme and indirectly via differences in GSO. For instance, as discussed in Elvidge (2019), in the MetUM there is a considerably larger LM97 scheme blocking coefficient, C_d, a different method for deriving the local Froude number used by the LM97 scheme (see Appendix), and a different parameterisation for turbulent orographic drag in the boundary layer (which employs a drag coefficient smaller by half than that recommended from a physical perspective by Mason, 1987). This results in elevated τ_{SSO} and lower τ_{BL} relative to the IFS, which uses the TOFD scheme for boundary layer drag, following Beljaars et al. (2004). Moreover, the MetUM's n768 resolution is marginally coarser than the IFS 16 km's TL1279, and so, irrespective of the SSO fields, slightly higher parameterised stresses should be expected in the MetUM. Differences in GSO will indirectly bring about differences in surface stresses due to the modulation of background conditions by the GSO and nonlinear interactions between resolved and parameterized drag (van Niekerk et al., 2018).

The IFS experiments with a combination of IFS and MetUM SSO fields reveal the relative importance of each SSO field for τ_{SSO}, τ_{BL} and τ_{total} (Figure 9 d-f). Corroborating the results from the 'offline' LM97 experiments discussed in Section 4, the *stdev* and *slope* impart considerably greater impacts than *anisotropy* and *orientation* on surface stress (Fig 9d-f). Their combined impact is responsible for the vast majority of the impact on τ_{total}, τ_{SSO} and (indirectly) τ_{BL} obtained when all four IFS SSO fields are replaced with those of the MetUM (87 % in terms of latitude-weighted global average for τ_{total}). The combined impact of *stdev* and *slope* explains 69 % of the total difference in latitude-weighted global average τ_{total} between the IFS$_{IFS-SSO}$ and MetUM experiments.

5.2 Impact on the large-scale circulation

The sensitivity of the large-scale circulation to the SSO fields is now explored. The SSO fields used in orographic drag parameterisations, such as the LM97 scheme, can affect both the lower-troposphere through low-level blocking as well as upper levels via gravity wave breaking. Figure 10 shows the impact on NH surface pressure of all the IFS experiments in which one or more MetUM SSO fields are used relative to the IFS$_{IFS-SSO}$ experiment, at a forecast lead time of 24 hours. These changes in surface pressure are associated with changes in the orographic drag strength (illustrated for example by Figure 9f) resulting from the use of MetUM-derived SSO fields. For the IFS$_{MetUM-SSO}$ experiment there is a coherent region of significantly increased mean surface pressure (up to 0.3 hPa higher) over the high latitudes (>60 °N) and extending to mid latitudes (30-60 °N), especially so between approximately 20°E and 100°E across east Europe and much of the north of Asia (Figure 10a). This pattern roughly resembles a known longstanding positive polar pressure bias in the MetUM, which has been shown to be highly sensitive to surface drag; both τ_{SSO} (Elvidge, 2019) and τ_{BL} due to sea ice roughness (Renfrew et al., 2019). There is a broad decrease in pressure elsewhere, at lower latitudes. The responses in surface pressure generally resemble those obtained by Sandu et al. (2016) in IFS sensitivity experiments where the strength of either low-level blocking or TOFD were enhanced.

In Section 5.1, the impact on parameterised drag in the IFS$_{MetUM-SSO}$ experiment was found to be nearly entirely explained by the combined effects of both the *stdev* and *slope* fields. The same is found here for the impact on surface pressure (Figure 10b). The *combined* effect of *stdev* and *slope* is required to explain the surface pressure response; replacing each of these two fields individually produces comparable patterns, though of much smaller amplitude (Figure 10 c,d). As expected, given their only small influence on surface stresses, the

anisotropy and *orientation* fields have only minor influence on the surface pressure (10 e,f). These findings also hold for the impact on NWP skill in the ten-day weather forecasts performed with the IFS. The net effect, for example, on the root mean square error of geopotential height at different levels in the atmosphere (from near-surface to the stratosphere) is broadly explained by the combined effect of *stdev* and *slope*, cannot be explained by either of these fields independently and is relatively insensitive to *anisotropy* and *orientation* (not shown). It is worth noting that the use of MetUM SSO fields in the IFS degrades the representation of the circulation significantly up to day 3. This reflects the fact that the IFS is optimally tuned to its own SSO fields.

In summary, it has been demonstrated that the global variability in parameterised surface stress due to the inter-model spread in SSO fields is of comparable magnitude and can explain a significant portion of variability in parameterised surface stress across models. Both the 'offline' LM97 experiments discussed in Section 4 and the global SSO field sensitivity experiments discussed here demonstrate that the *stdev* and *slope* SSO fields exert considerably more influence on the parameterised surface stress and on global circulation than do *anisotropy* and *orientation*. The combined influence of both fields is crucial to the local and large scale atmospheric response.

6. Conclusions

In the first major model comparison of the grid-scale orography (GSO) and subgrid-scale orography (SSO) fields used in global NWP models, notable inter-model variability is found across six operational models. This variability stems from several different sources. The global mean of the standard deviation of SSO *(the *stdev* SSO field) – a measure of the degree

to which orography is parameterized or resolved – is evidently correlated with model resolution, though this relationship is shown to be far from linear, revealing the importance of other factors such as model grid type, fidelity of the source orography dataset and the manner in which this dataset is processed to derive the GSO and SSO fields. In particular, there is considerable variation in the degree and type of smoothing applied. The influence of these technical and methodological differences on the resultant depiction of model orography is difficult to quantify, and further work could be done to isolate their impact.

Sensitivity experiments using an 'offline' version of the popular LM97 orographic drag scheme over global orography during typical atmospheric conditions were used to evaluate the impact of the SSO fields on parameterised orographic surface stress and its partition between low-level flow blocking and higher level wave breaking components. By using the inter-model spread in each SSO field as a benchmark for their plausible range, we were able to ascertain the relative importance of each field. It has been found that parameterised orographic surface stresses are most sensitive to the inter-model variability in the *stdev* and *slope* SSO fields. The influences of the *anisotropy* and *orientation* fields are considerably smaller. The partition of parameterized orographic drag into its constituent components – low-level blocking drag and gravity wave drag – is also highly sensitive to the SSO fields. In the LM97 scheme, a greater proportion of blocking drag is found for those models with typically greater *stdev* values, due to the gravity wave drag being on the whole less sensitive than the blocking drag to *stdev*.

Global weather forecasts performed with the IFS, with combinations of SSO fields derived from the IFS and MetUM, were used to examine the impact of the SSO fields on the total parameterised surface stress and its partition between orographic and boundary layer components, and on large-scale circulation on daily timescales. These experiments have

demonstrated that 55 % of global-average variability in SSO stress and 80 % of the global-average variability in total surface stress between the MetUM and IFS can be accounted for by differences in SSO fields. The current uncertainty in SSO fields is evidently of first-order importance to current uncertainty in total parameterised surface stress. Corroborating the results of the 'offline' LM97 sensitivity experiments, the combined differences in the *stdev* and *slope* fields between the MetUM and IFS were found to account for 86 % of the impact on total parameterised surface stress of the combined differences in all SSO fields. The combined influence of these fields on surface pressure is similarly dominant.

Via model comparison, this study has established the importance of orographic fields in NWP models, and consequently the significance of the current uncertainty in how they are sourced, generated and implemented in orographic drag parameterisation schemes. Careful re-consideration of each of these steps is recommended. For example, of the models in this study, only the IFS currently employs the global source dataset generally considered to be the highest quality (SRTM30). Yet recent as-yet unpublished work has revealed that the choice of source dataset has a significant effect on orographic drag but cannot explain the total spread seen among models. In other recent work, the degree of orographic smoothing employed has been found to have a considerable effect on the resultant orographic fields and consequently on modelled drag, yet the appropriate amount of smoothing, and precisely how and at what stage in the orography data processing this smoothing is applied, remains unclear. Addressing this uncertainty requires consideration of what scales are being resolved, and consequently which scales require parameterisation. In all the models assessed in this study, the current motivation for filtering the GSO is to limit the scales to be resolved so as to avoid model instability over steep orography. However, it is known that models do not fully resolve the effects of orography at the grid-scale, but at a coarser resolution (evidence from a MetUM experiment over New Zealand suggests wavelengths shorter than 8-10 grid lengths are poorly

resolved; Vosper et al., 2016). Given the strength of the filters currently used, some scales finer than this "effective resolution" remain after smoothing, and consequently are neither parameterised nor resolved. So there is a case for increasing the degree of orographic smoothing applied in the generation of SSO fields so that they represent scales up to the effective resolution of the model, whilst retaining as much high resolution resolved orography as possible, under the restraints of model stability. However, this argument assumes that the parameterisation schemes are designed (and tuned) to hand over between resolved and parameterised scales in a well behaved way. Vosper et al. (2016) show that this is not universally the case in schemes which are tuned to optimise global forecasts. They point towards the development of a regionally variable (scale-aware) tuning framework as a potential solution.

Note that this study has not addressed the impacts of model variability in GSO on drag and circulation. This matter has been touched upon in previous work by Sandu et al. (2017) – who showed significant large-scale circulation forecast sensitivities to differences in GSO – and will be the subject of further work.

Future development in orographic drag parameterisation should consider the relative importance of each SSO field in the LM97 scheme. The component of the LM97 scheme representing the influence of the horizontal shape of SSO via *anisotropy* and *orientation* has been shown to contribute trivially to resultant parameterised stresses, at least in the mean sense and in terms of the impact on the large-scale circulation, and so is arguably not required (however, the extent to which these fields matter for the prediction of local flow patterns and near surface weather conditions in mountainous regions requires further investigation). Meanwhile, consideration could be given to developing parameterisation complexity in the more important aspect of the influence of the vertical cross-sectional characteristics of SSO.

For example, only half of the study's participant models employ an independently varying *slope* field in their drag schemes. When *stdev* and *slope* are incorporated in the manner of LM97, their combined description of the SSO's vertical cross sectional characteristics is limited. For instance, they are able to define SSO characterised by undulating terrain in precisely the same manner as SSO characterised by a single high mountain peak surrounding otherwise flat land. A more complete description of grid box variability in *stdev* and *slope* is arguably warranted, the nature and influence of which could be explored using high resolution modelling of real and idealised topography. Indeed, due to the 4-dimensional complexity of orographic drag, future developments in its parameterisation are likely to rely on high resolution (km-scale) simulations (e.g. Van Niekerk et al., 2018). Nevertheless, observations targeting the effect of orography on the atmosphere are essential in providing real-world grounding and a means to validate the high resolution simulations, and remain in high demand.

Appendix. The LM97 implementation in the MetUM

The LM97 scheme is the predominant parameterisation for orographic drag used in NWP and climate models. Here we describe its implementation in the MetUM. The scheme is based on two separate conceptual models to represent drag due to both low-level orographic flow blocking and mountain wave breaking. The blocking drag component is derived following bluff body dynamics and takes the form:

$$\boldsymbol{D_b} \sim - \rho \; C_d' \; l \; \frac{u|u_{av}|}{2} , \qquad (1)$$

where $\boldsymbol{D_b}(z)$ is the drag exerted on the flow at levels within the blocked flow layer, C_d' is a variable drag coefficient which incorporates the tuning coefficient C_d but is allowed to vary

with the SSO's aspect ratio as seen by the incident flow (being a function of wind direction, *anisotropy* and *orientation*), $\rho(z)$ is air density, $\boldsymbol{u}(z)$ is wind speed, $\boldsymbol{u_{av}}$ is a depth-averaged wind speed and $l(z)$ represents the cumulative horizontal width of subgrid orography as seen by the flow across the grid box at height z. For flow across a single grid box,

$$l \sim L^2 \frac{\alpha}{2\sigma} \sqrt{\frac{z_b - z}{z + \sigma}} \max(\cos \varphi, \gamma \sin \varphi), \qquad (2)$$

where L is the grid box length scale, z_b is the blocked layer depth, φ is the angle between the incident flow direction and SSO *orientation*, and σ, α γ are the *stdev*, *slope* and *anisotropy* SSO fields. Note that $l = 0$ where $z = z_b$ as the flow at this height is assumed to ascend to the parameterized mountain tops, and the level $z = 0$ (i.e. the GSO) is assumed to be at height σ above the *parameterized mountain* valleys. The dependency of z_b on the mountain flow regime is given by

$$z_b = \max\left[0, n\sigma\left(1 - \frac{F_{av}}{F_c}\right)\right], \qquad (3)$$

where n and the critical Froude number, F_c are tunable constants and $F_{av} = \boldsymbol{u_{av}}/(N_{av}\, n\sigma)$ is the Froude number – the ratio of the flow's vertical buoyancy oscillation length scale to mountain height and a measure of the degree to which the incident flow is blocked by the mountain. Here, N_{av} is a depth-averaged Brunt Vaisala frequency and F_{av} – itself a function of $\boldsymbol{u_{av}}$ and N_{av} – is derived iteratively, as described in Vosper (2015).

The wave drag component of the LM97 scheme is based on linear gravity-wave theory and given by

$$\boldsymbol{D_{gw}} = \rho_l \, \boldsymbol{u_l} \, N_l \, G \, (\alpha/\sigma) \left(\frac{h_{gw}^2}{4}\right) A, \qquad (4)$$

where the subscript l denotes a low-level depth-averaged quantity, A is a term representing the effects of anisotropy (being a function of wind direction, *anisotropy* and *orientation*), G is a tuning constant, and h_{gw} is the depth of the uppermost portion of the parameterized mountains that are responsible for the forcing of parameterised gravity waves, provided by

$$h_{gw} = n\sigma - z_b. \qquad (5)$$

In the current global operational MetUM, tuning parameters C_d, n, F_c and G have the values 4, 2.5, 4 and 0.5, respectively.

The nonlinear, non-monotonic relationship sensitivity of D_{gw} to *stdev*, σ is now briefly explored. It can be shown from Equations 4 and 5 that, for vertically uniform background flow (a reasonable assumption for the purpose of this exploration), $D_{gw} \propto \sigma^{-1}$ in sub-critical Froude number flows (i.e. where $F_{av} < F_c$). In this regime, the depth of the uppermost portion of the *parameterized mountains* responsible for forcing gravity waves (h_{gw}) is independent of σ (with changes in σ reflected in equivalent changes in the flow blocking depth, z_b). However, in the LM97 scheme framework, higher subgrid-scale mountains (with no change in *slope*) effectively means fewer *parameterized mountains* in the gridbox and consequently weaker D_{gw}. Conversely, in super-critical Froude number flows (i.e. where $F_{av} > F_c$ and therefore, from Equation 3, $z_b = 0$), gravity waves are forced by the full depth of SSO, i.e. $h_{gw} \propto \sigma$. In this regime – noting that the linear hydrostatic expression for D_{gw} has $D_{gw} \propto h_{gw}^2 / \sigma$ (see Equation 4) – it can be shown that $D_{gw} \propto \sigma$. So, the sensitivity of D_{gw} to σ is Froude number dependent, with an inflection point in D_{gw} (as a function of Froude number) at the critical Froude number (F_c).

Acknowledgements:

References

ARPEGE-Climate Version 5.2 (2011). Documentation – Chapter 13: Gravity wave drag. http://www.umr-cnrm.fr/gmgec/arpege-climat/ARPCLI-V5.2/doca/arp52ca.pdf. Accessed 31 Jan 2019

Bacmeister, J. T. (1993). Mountain-wave drag in the stratosphere and mesosphere inferred from observed winds and a simple mountain-wave parameterization scheme. *Journal of the atmospheric sciences*, *50*(3), 377-399.

Beljaars, A., Brown, A. R., & Wood, N. (2004). A new parameterisation of turbulent orographic form drag. *Quarterly Journal of the Royal Meteorological Society*, *130*(599), 1327-1347.

Boer, G. J., McFarlane, N. A., Laprise, R., Henderson, J. D., & Blanchet, J. P. (1984). The Canadian Climate Centre spectral atmospheric general circulation model. *Atmosphere-Ocean*, *22*(4), 397-429.

Bouteloup, Y. (1995). Improvement of the spectral representation of the earth topography with a variational method. *Mon. Wea. Rev.*, 123, 1560–1573.

Brayshaw, D. J., Hoskins, B., & Blackburn, M. (2009). The basic ingredients of the North Atlantic storm track. Part I: Land–sea contrast and orography. *Journal of the Atmospheric Sciences*, *66*(9), 2539-2558.

Buehner, M., McTaggart-Cowan, R., Beaulne, A., Charette, C., Garand, L., Heilliette, S., Lapalme, E., Laroche, S., Macpherson, S.R., Morneau, J., & Zadra, A. (2015). Implementation of deterministic weather forecasting systems based on ensemble–variational

data assimilation at Environment Canada. Part I: The global system. *Monthly Weather Review*, *143*(7), 2532-2559.

Catry, B., Geleyn, J. F., Bouyssel, F., Cedilnik, J., Brožková, R., & Derková, M. (2008). A new sub-grid scale lift formulation in a mountain drag parameterisation scheme. *Meteorologische Zeitschrift*, *17*(2), 193-208.

ECMWF IV, I. F. S. (2016). Documentation—Cy41r2 Part IV: Physical Processes. European Centre for Medium-Range Weather Forecasts: Reading, UK.

Elvidge, A. D. (2019). Orographic drag in the Met Office Unified Model: sensitivity to parameterisation and insights from inter-model variability in drag partition. *Met Office Forecasting Research Tech. Rep. 635.*

Elvidge, A. D., & Renfrew, I. A. (2016). The causes of foehn warming in the lee of mountains. Bulletin of the American Meteorological Society, 97(3), 455-466.

Elvidge, A. D., Vosper, S. B., Wells, H., Cheung, J. C., Derbyshire, S. H., & Turp, D. (2017). Moving towards a wave-resolved approach to forecasting mountain wave induced clear air turbulence. Meteorological Applications, 24(3), 540-550.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of geophysics*, *45*(2).

Hourdin, F., Gueye, M., Diallo, B., Dufresne, J. L., Escribano, J., Menut, L., Marticoréna, B., Siour, G., & Guichard, F. (2015). Parameterization of convective transport in the boundary

layer and its impact on the representation of the diurnal cycle of wind and dust emissions. *Atmospheric Chemistry and Physics*, *15*(12), 6775-6788.

Hourdin, F., T. Mauritsen, A. Gettelman, J. Golaz, V. Balaji, Q. Duan, D. Folini, D. Ji, D. Klocke, Y. Qian, F. Rauser, C. Rio, L. Tomassini, M. Watanabe, and D. Williamson (2017), The Art and Science of Climate Model Tuning. Bull. Amer. Meteor. Soc., 98, 589–602, https://doi.org/10.1175/BAMS-D-15-00135.1

Iwasaki, T., Yamada, S., & Tada, K. (1989). A parameterisation scheme of orographic gravity wave drag with two different vertical partitionings. *Journal of the Meteorological Society of Japan. Ser. II*, *67*(1), 11-27.

JMA (2019) Outline of the operational numerical weather prediction at the Japan Meteorological Agency (March 2019), Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System (GDPFS) and Numerical Weather Prediction (NWP) Research. 229pp. http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2019-nwp/index.htm.

Lawrence, B. N., Bennett, V. L., Churchill, J., Juckes, M., Kershaw, P., Pascoe, S., Pepler, S., Pritchard, M. & Stephens, A. (2013). Storing and manipulating environmental big data with JASMIN. In *2013 IEEE international conference on big data* (pp. 68-75). IEEE.

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

Malardel, S., & Wedi, N. P. (2016). How does subgrid-scale parametrization influence nonlinear spectral energy fluxes in global NWP models?. Journal of Geophysical Research: Atmospheres, 121(10), 5395-5410.

Mason, P. J. (1987). Diurnal variations in flow over a succession of ridges and valleys. *Quarterly Journal of the Royal Meteorological Society*, *113*(478), 1117-1140.

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

van Niekerk, A., Shepherd, T. G., Vosper, S. B., & Webster, S. (2016). Sensitivity of resolved and parameterised surface drag to changes in resolution and parameterisation. *Quarterly Journal of the Royal Meteorological Society*, *142*(699), 2300-2313.

van Niekerk, A., Scinocca, J. F., & 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*(8), 2557-2574.

van Niekerk, A., Sandu, I., & 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.

Pithan, F., Shepherd, T. G., Zappa, G., & Sandu, I. (2016). Climate model biases in jet streams, blocking and storm tracks resulting from missing orographic drag. *Geophysical Research Letters*, *43*(13), 7231-7240.

Renfrew, I. A., Elvidge, A. D., & Edwards, J. M. (2019). Atmospheric sensitivity to marginal-ice-zone drag: Local and global responses. *Quarterly Journal of the Royal Meteorological Society*.

Sandu, I., Bechtold, P., Beljaars, A., Bozzo, A., Pithan, F., Shepherd, T. G., & Zadra, A. (2016). Impacts of parameterised orographic drag on the Northern Hemisphere winter circulation. *Journal of advances in modeling earth systems*, *8*(1), 196-211.

Sandu, I., Zadra, A., & Wedi, N. (2017). Impact of orographic drag on forecast skill. *ECMWF Newsletter, 150*, 18–24.

Scinocca, J. F., & Sutherland, B. R. (2010). Self-acceleration in the parameterization of orographic gravity wave drag. *Journal of the Atmospheric Sciences*, *67*(8), 2537-2546.

Smith, R. B. (1979), The Influence of Mountains on the Atmosphere, Adv. Geophys., 21, 87-230, doi: 10.1016/S0065-2687(08)60262-9.

Smith, R. B. (1985). On severe downslope winds. Journal of the atmospheric sciences, 42(23), 2597-2603.

Tolstykh, M. A., Fadeev, R. Yu., Shashkin, V. V., Goyman, G. S., Zaripov, R. B., Kiktev, D. B., Makhnorylova, S. V., Mizyak, V. G., & Rogutov, V. S. (2018). Multiscale Global Atmosphere Model SL-AV: the Results of Medium-range Weather Forecasts. *Russian Meteorology and Hydrology*, 43(11), 773-779.

Vosper, S. B. (2015). Mountain waves and wakes generated by South Georgia: implications for drag parameterisation. *Quarterly Journal of the Royal Meteorological Society*, *141*(692), 2813-2827.

Vosper, S. B., Brown, A. R., & Webster, S. (2016). Orographic drag on islands in the NWP mountain grey zone. Quarterly Journal of the Royal Meteorological Society, 142(701), 3128-3137.

Vosper, S., Ross, A., Renfrew, I., Sheridan, P., Elvidge, A., & Grubišić, V. (2018). Current challenges in orographic flow dynamics: turbulent exchange due to low-level gravity-wave processes. Atmosphere, 9(9), 361.

Wallace, J. M., Tibaldi, S., & Simmons, A. J. (1983). Reduction of systematic forecast errors in the ECMWF model through the introduction of an envelope orography. *Quarterly Journal of the Royal Meteorological Society*, *109*(462), 683-717.

Walters, D., Brooks, M., Boutle, I., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., & Bushell, A. (2017). The Met Office unified model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations. *Geoscientific Model Development*, *10*(4), 1487-1520.

Webster, S. (2003). Orography Ancillary File Creation. *Unified Model Documentation Paper No. 74.*

Webster, S., Wells, H., & Vosper, S. (2013). Gravity wave drag. Unified Model Documentation Paper, (22). http://cms.ncas.ac.uk/wiki/Docs/MetOfficeDocs

Webster, S., Brown, A. R., Cameron, D. R., & Jones, C. P. (2003). Improvements to the representation of orography in the Met Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, *129*(591), 1989-2010.

Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., & Roth, A. (2018). Accuracy assessment of the global TanDEM-X Digital Elevation Model with GPS data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *139*, 171-182.

Zadra, A. (2013). *WGNE drag project: An inter-model comparison of surface stresses*. Technical report.

http://collaboration.cmc.ec.gc.ca/science/rpn/drag_project/documents/wgne_drag_project_report01.pdf

Zadra, A. (2018). Notes on the new low-pass filter for the orography field. Internal Report, RPN-A, Meteorological Research Division, Environment and Climate Change Canada. http://collaboration.cmc.ec.gc.ca/science/rpn/drag_project/documents/topo_lowpass_filter.pdf

Model acronym	Full model name	Institute	Grid type	Native horizontal resolution		Orographic Drag parameterisation scheme framework
				Global mean grid spacing (km)	Details	
MetUM	Unified Model	Met Office	Regular lat/lon	17.0	N768: From 12 km over poles to 22 km over equator	Following Lott and Miller (1997), Boer et al. (1984) and McFarlane et al. (1987).
IFS Operational	Integrated Forecast System	European Centre for Medium-Range Weather Forecasts	Gaussian reduced octahedral	8.8	Tco1279 spectral resolution: From 6 km over poles to 11 km over equator	
IFS 16 km			Gaussian reduced	15.5	Tl1279 spectral resolution: From 11 km over poles to 20 km over equator	
GDPS	Global Deterministic Prediction System	Canadian Meteorological Centre	Rotated Yin-Yang grid	25.0	The Yin-Yang grid allows the resolution to be approximately uniform across all latitudes	
ARPEGE	Action de Recherche Petite Echelle Grande Echelle	Meteo-France	Schmidt projection	16.3	T1198 spectral resolution: From 7 km over France to 33 km over the South Pacific	Flow blocking scheme following envelope orography approach of Wallace et al. (1983). Gravity wave scheme following Lott and Miller (1997), Boer et al. (1984) and McFarlane et al. (1987).
SL-AV	Semi-Lagrangian – Absolute Vorticity Model	Hydrometeorological Research Centre of Russia	Regular lat/lon	19.2	from 12 km in NH mid latitudes to 26 km degrees in SH low latitudes	
GSM	Global Spectral Model	Japan Meteorological Agency	Gaussian reduced	19.7	Tl959 spectral resolution: 20 km in mid-latitudes	Longwave scheme (wavelengths > 100 km) and shortwave scheme (wavelengths approx. 10 km) based on Iwasaki et al. (1989).

Table 1. The names of the models, and details of the native grids and orographic drag schemes employed operationally.

Model and appropriate reference(s)	Source orography dataset	Orographic smoothing		SSO fields employed
		Approximate filter strength (Δx = grid length)	Details	
MetUM (Walters et al., 2017; Webster, 2003)	GLOBE	4.0Δx	Scale selective Raymond filter (see Webster et al., 2003)	$\sigma, \gamma, \theta, \alpha$
IFS Operational IFS 16 km (ECMWF, 2016)	SRTM30 from 60˚N to 60˚S; GLOBE north of 60˚N; Antarctic RAMP2 south of 60˚S; BPRC for Greenland; Iceland DEM	2.3Δx	A general smoothing operator is used in grid point space. The smoothing operator is a top hat function with smooth edges (see ECMWF, 2016)	$\sigma, \gamma, \theta, \alpha$
GDPS (Buehner et al., 2015)	GTOPO30	4.0Δx	Smoothing operator in grid point space (9-point filter with weight parameter f = 1/2). See Zadra (2018) for more details.	$\sigma, \gamma, \theta, \alpha$
ARPEGE (ARPEGE, 2011)	GTOPO30	3.0Δx	The spectral orography is computed with a "quadratic" truncation minimizing a cost function which is a measure of a distance between the mean orography and the spectrally fitted orography (following Bouteloup, 1995).	σ, γ, θ
SL-AV (Tolstykh et al., 2018)	GTOPO30	7.2Δx, before additional 9-point filtering	1-d Raymond filter, followed by a 9-point filter applied to sharp mountains only with the coefficient of the central point equal to 0.25	σ, γ, θ
GSM (JMA, 2019)	GTOPO30	2.7Δx	The grid-averaged elevation is spectrally smoothed by multiplying the spectral coefficients by a smoothing function (section 3.2.11 in JMA, 2019).	Σ

Table 2. Details of the sourcing and processing of orography for each model. The SSO fields are *stdev*, σ; *slope*, α; *anisotropy*, γ; and *orientation*, θ.
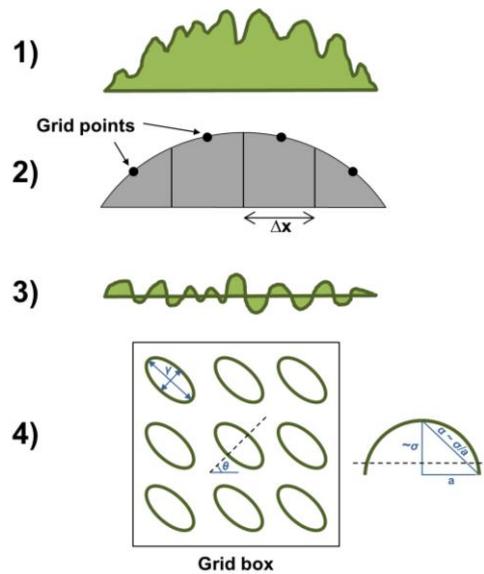
Figure 1. Schematic summarising the method for generating the GSO and SSO fields. Note the exact order of the smoothing and interpolation procedures are model dependent. 1) The high resolution source dataset is 'pre-filtered' (and also sometimes interpolated to a coarser grid) to remove the finest scales (these scales, typically < 5 km, are dealt with by TOFD and effective roughness schemes); 2) The resulting dataset is then smoothed and interpolated to the model grid to provide the GSO; 3) The SSO orography is then generated as the difference between the pre-filtered source orography and the GSO interpolated to the pre-filtered source data grid (or, in the IFS, smoothed versions of the pre-filtered source data); 4) The SSO fields (*stdev*, σ; *slope* along the SSO major axis, α; *anisotropy*, γ; and *orientation* of the SSO major axis, θ) are computed to define, for each grid box, an array of uniformly distributed, elliptical *parameterized mountains*.

Figure 2. Response functions of the orographic filters applied in the MetUM, IFS, GDPS, SL-AV and GSM to the pre-filtered source orography prior to the derivation of the SSO. Note the curve for ARPEGE is missing due to the filter it employs not lending itself to illustration in this form.
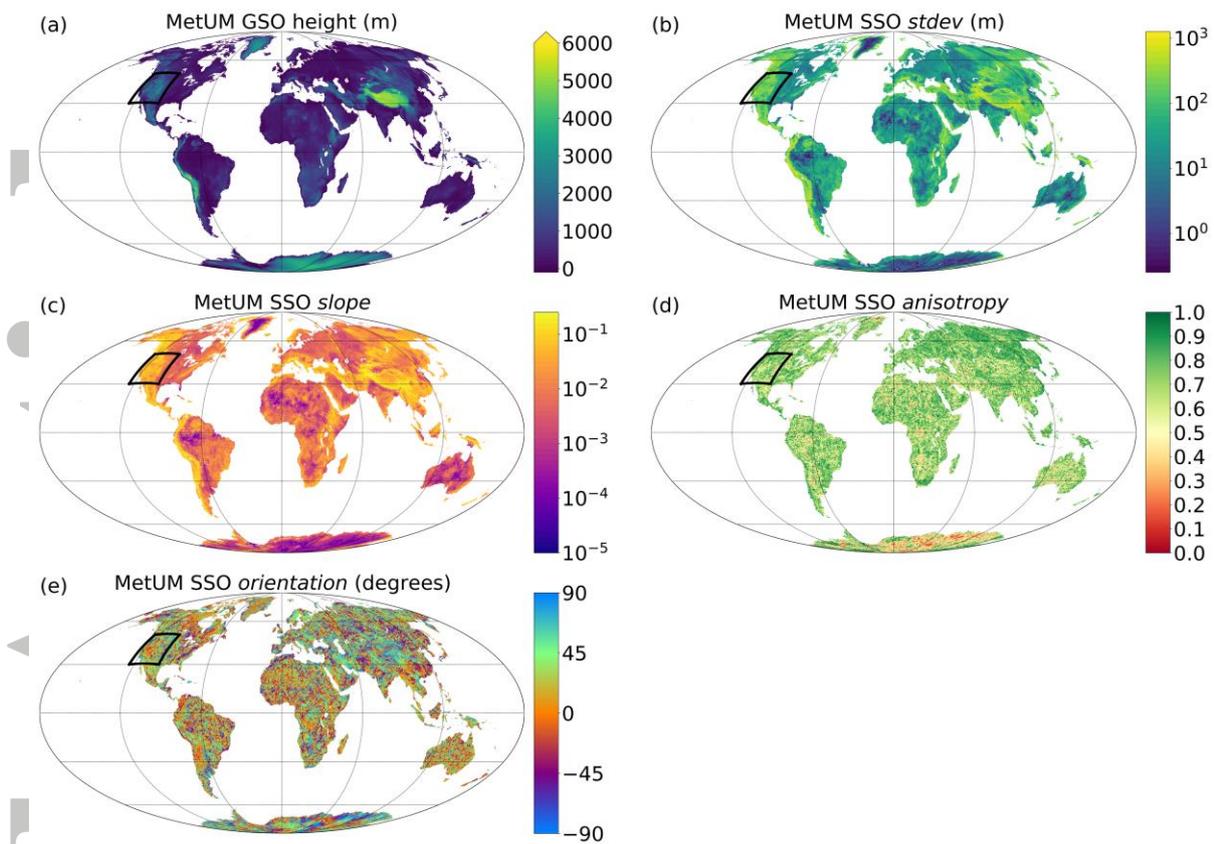
Figure 3. Global GSO height and SSO fields in the MetUM. In each figure panel, the black box encloses the Rockies region used for the probability density function plots (Figure 5).
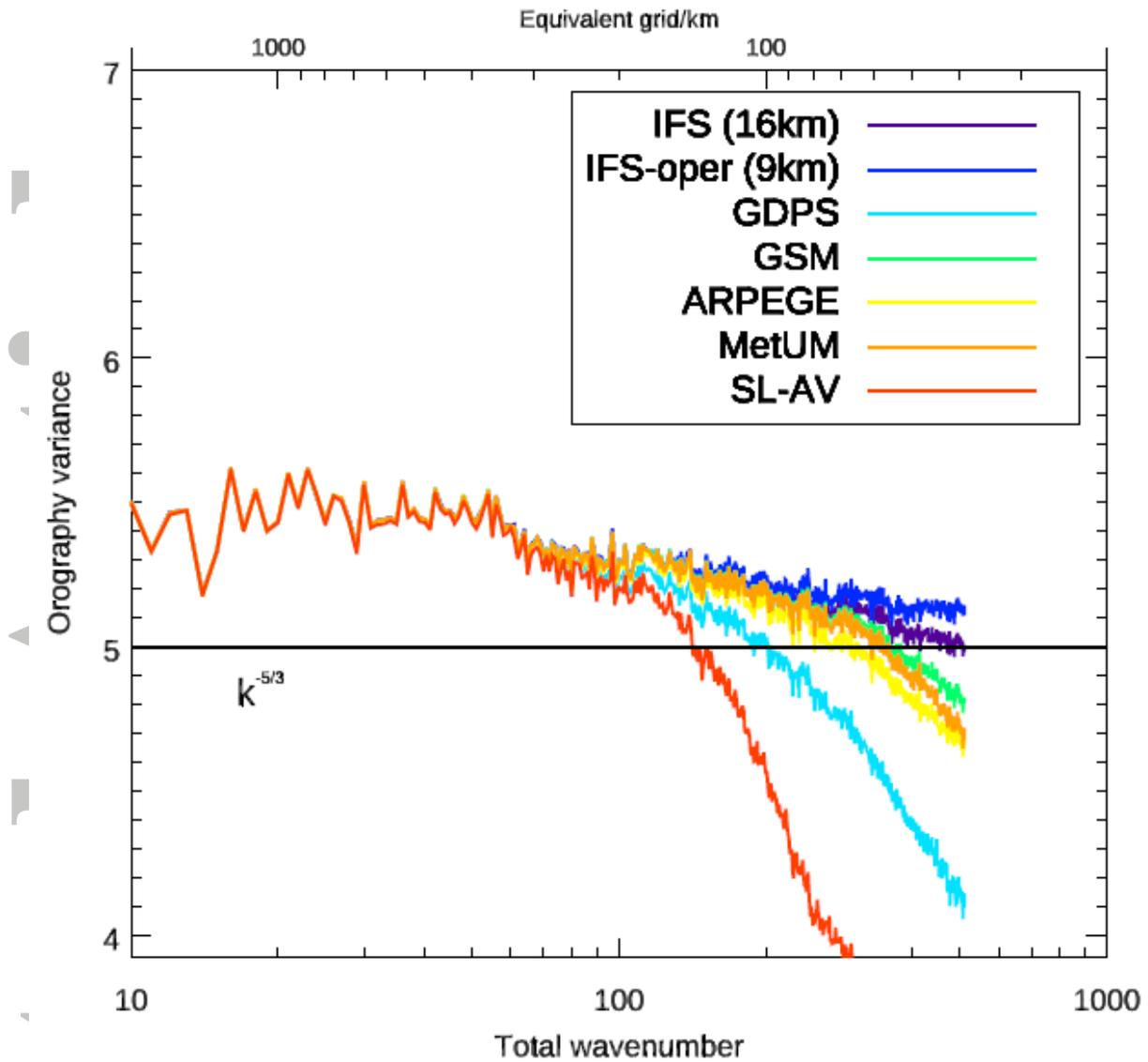
Figure 4. Variance of the resolved global orography as a function of the total wavenumber, k, for all models. Note that, for clarity, all spectra are multiplied by $k^{5/3}$. Then the horizontal line identifies $k^{-5/3}$. For this plot the 0.25 x 0.25 degree gridded data have been spectrally truncated to 511 wavelengths (39 km at the equator).

Figure 5. Probability density functions for a) GSO height; b) SSO *stdev*; c) SSO *slope*; d) SSO *anisotropy*; and e) SSO *orientation* over all land points within a region covering the Rocky Mountains (between 100 and 124 degrees West; and between 30 and 50 degrees North), for each of the models (for which the respective field is available).
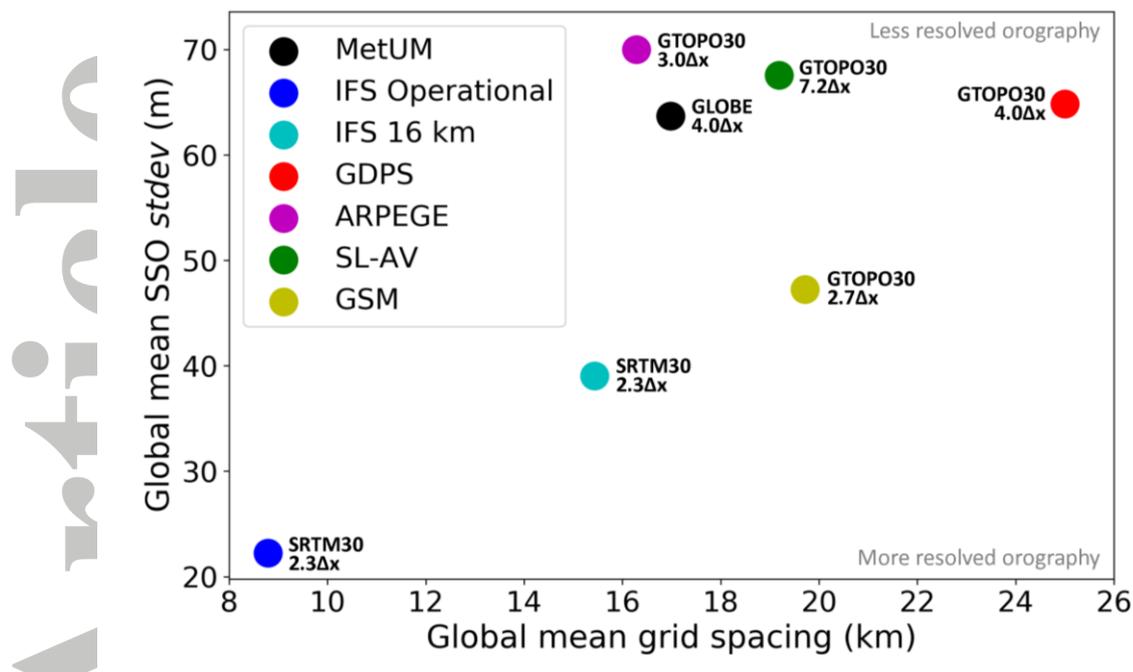
Figure 6. Global mean SSO *stdev* as a function of global mean model resolution. Data points for each model are annotated by the main source orography datasets employed and the filter strengths used to smooth the GSO before deriving the SSO (where Δx refers to the model grid spacing).
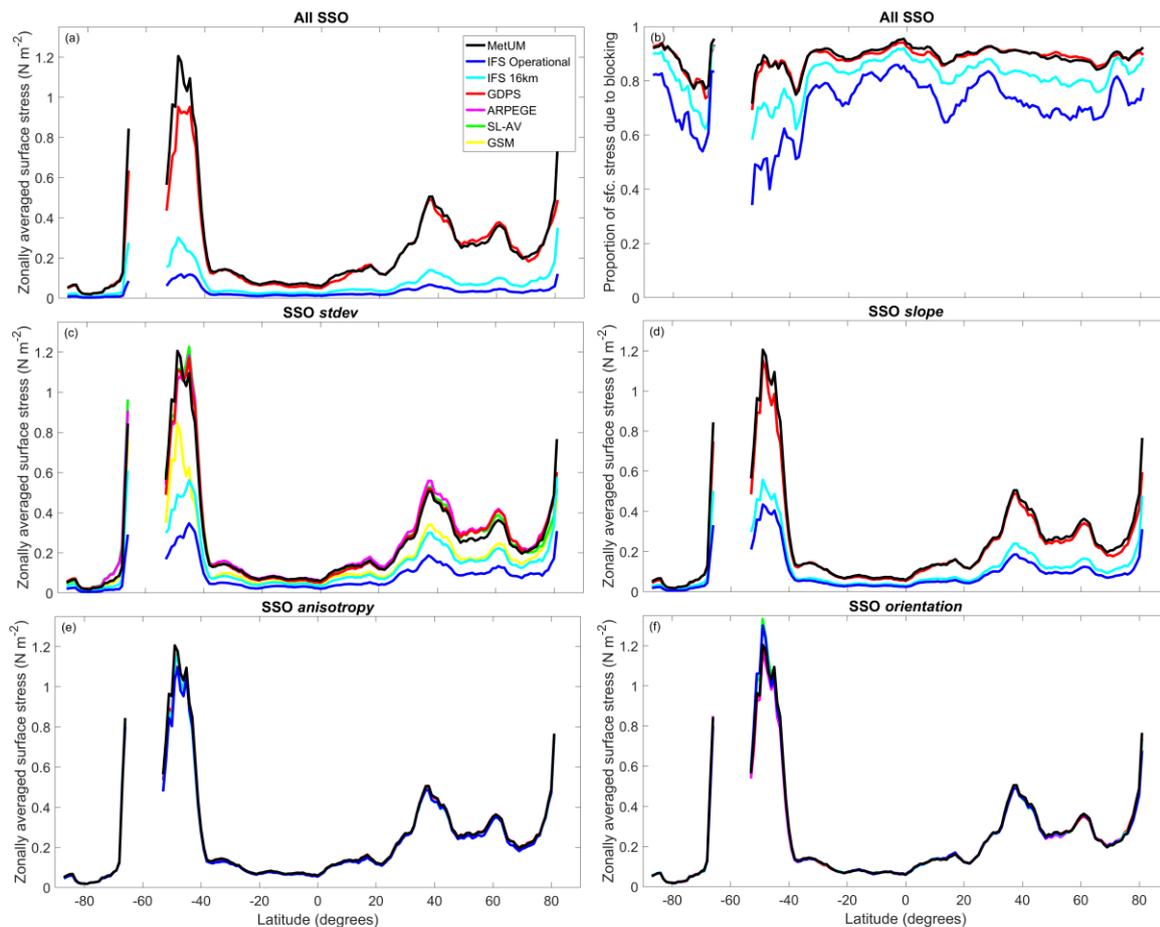
Figure 7. The sensitivity of zonally-averaged orographic surface stress magnitude to the inter-model variability in SSO fields, from the 'offline' LM97 scheme experiments. In panels (a) and (b) each line corresponds to an experiment where all SSO fields are derived from a single model (colour coded; see legend). In the experiments illustrated in panels (c-f), all but one of the SSO fields are sourced from the MetUM – the exception being c) *stdev*, d) *slope*, e) *anisotropy* and f) *orientation*, which for each coloured line is sourced from the model indicated in the legend. Panel (b) shows the fraction of the parameterised orographic surface stress that is due to low-level blocking drag (the rest being due to gravity wave drag) in the experiments in which the four SSO fields are sourced from the same model. The data are taken from all land points on the MetUM n768 grid, then interpolated to a 1-degree by 1-degree grid and subjected to a 5-degree running mean.
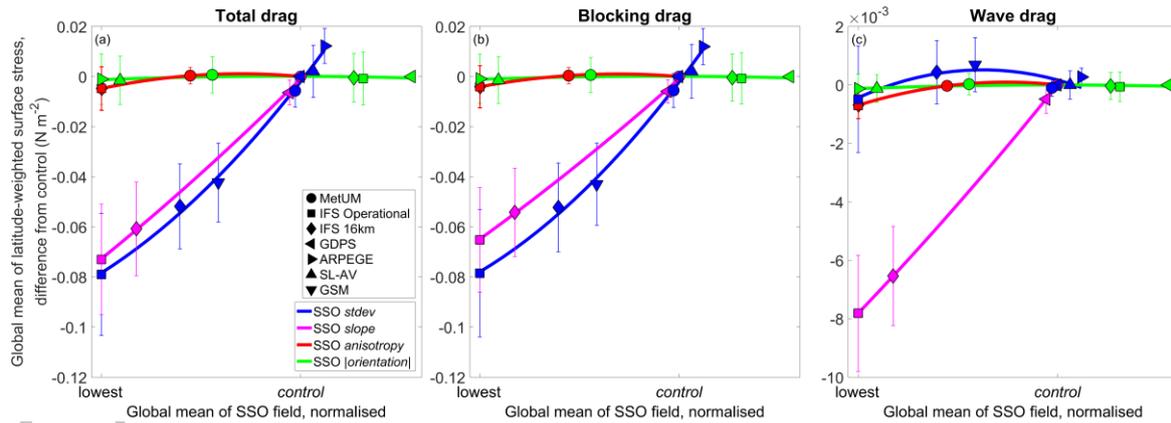
Figure 8. The sensitivity of a) total, b) low-level blocking and c) gravity wave surface stress magnitude (expressed as difference relative to a control) to the inter-model variability in each SSO field, from offline LM97 drag scheme experiments where all but one of the SSO fields are sourced from the MetUM. The exception (i.e. the SSO field that is varied) and the source model for this field are indicated in the legends. The *control* experiment is that associated with the highest global mean value of the relevant SSO field amongst those models for which all the SSO fields are available (the MetUM, both IFS versions and GDPS). For each data point, whiskers indicate one tenth of the standard deviation in stress difference relative to the *control* experiment across all grid points. Note that the stresses used to create the global means have been latitude-weighted (multiplied by the cosine of the latitude) to account for the fact that grid box areas reduce towards the poles, and only land points were considered for these averages.
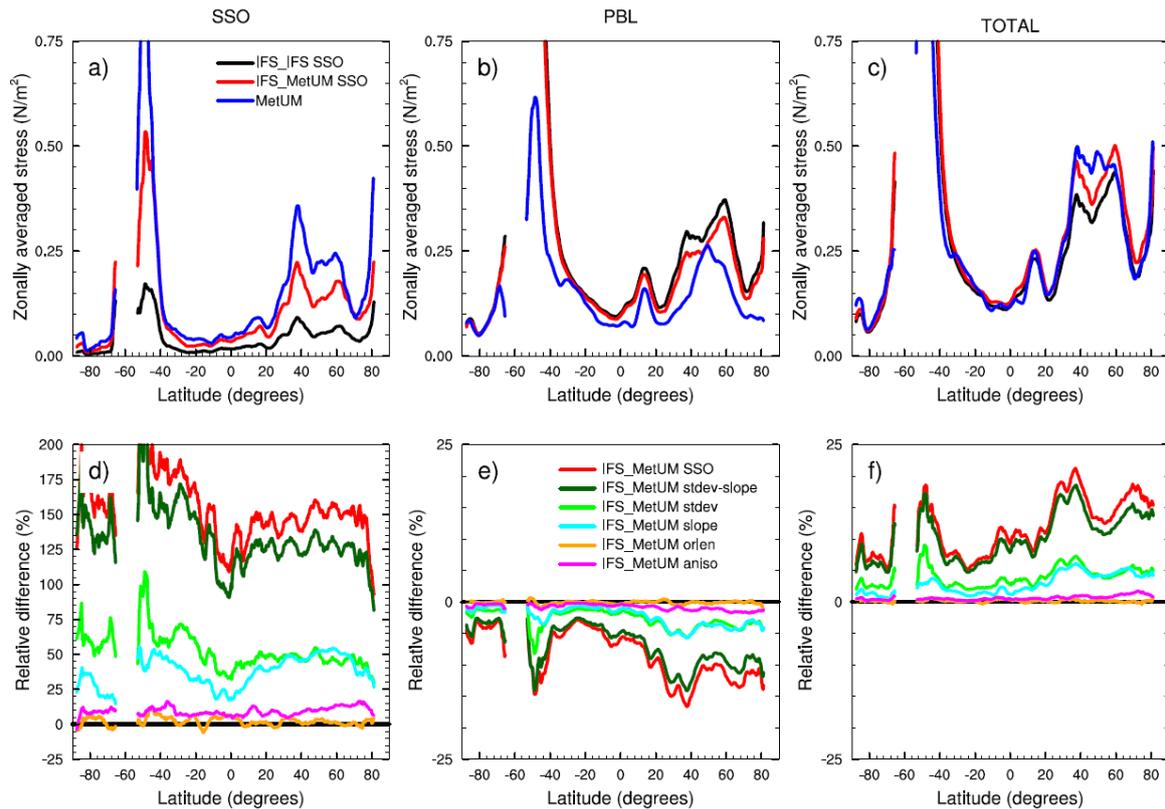
Figure 9. (a) to (c) Zonally averaged (over land points only) contributions to parameterised surface stresses from SSO and planetary boundary layer (PBL) schemes, and total surface stress, from the IFS$_{IFS-SSO}$ model experiment (black), the IFS$_{MetUM-SSO}$ experiment (red) and the MetUM experiment (blue). (d) to (f) Relative change in SSO, PBL and total zonally averaged parameterised stress in each of the IFS SSO-sensitivity experiments with respect to the IFS$_{IFS-SSO}$ experiment (i.e. which of the SSO fields have been replaced by MetUM SSO fields). The legend indicates which SSO-sensitivity experiment is represented by each line. The stresses represent the mean over the first 24 hours of simulation, and have been averaged over the 31 daily forecasts performed in each experiment for December 2016. The data have been interpolated to a 1-degree by 1-degree grid and subjected to a 5-degree running mean.

IFS_MetUM all

IFS_MetUM stdev-slope

IFS_MetUM stdev

IFS_MetUM slope

IFS_MetUM orlen
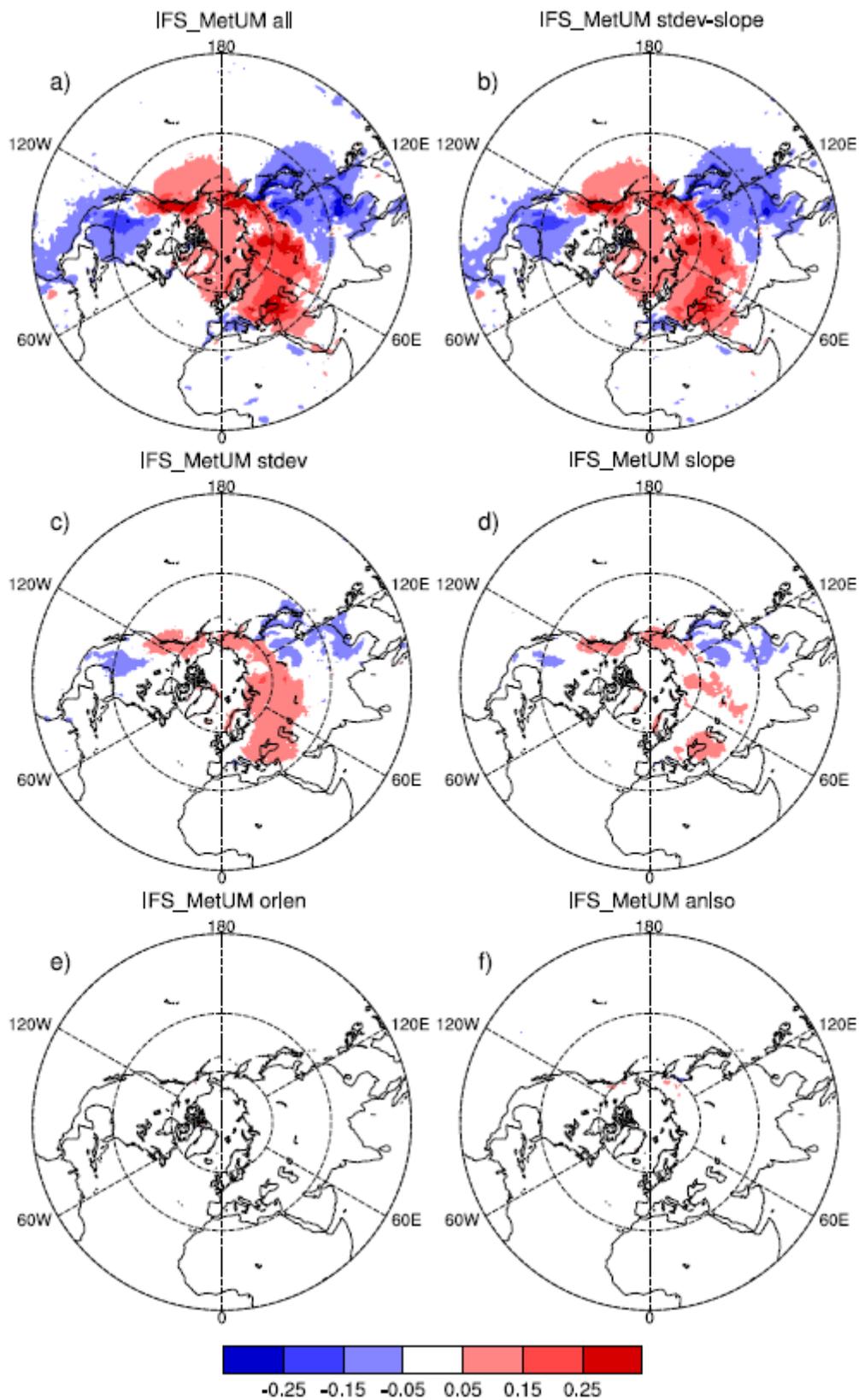
IFS_MetUM anlso

-0.25 -0.15 -0.05 0.05 0.15 0.25

Figure 10. The mean impact on IFS-simulated December 2016 Northern Hemisphere surface pressure (in hPa) of replacing (a) all, (b) *stdev* and *slope* (c) *stdev*, (d) *slope*, (e) *orientation* and (f) *anisotropy* IFS SSO field(s) with the equivalent MetUM SSO field(s). The plots represent the surface pressure at a lead time of 24 hours, averaged over the 31 daily forecasts performed for December 2016.