















## RESEARCH ARTICLE

10.1029/2020MS002160

# CONstraining ORographic Drag Effects (COORDE): A Model Comparison of Resolved and Parametrized Orographic Drag

Annelize van Niekerk<sup>1</sup> , Irina Sandu<sup>2</sup> , Ayrton Zadra<sup>3</sup> , Eric Bazile<sup>4</sup> ,  
Takafumi Kanehama<sup>2,5</sup> , Martin Köhler<sup>6</sup> , Myung-Seo Koo<sup>7</sup> , Hyun-Joo Choi<sup>8</sup> ,  
Yukihiro Kuroki<sup>5</sup> , Michael D. Toy<sup>9</sup> , Simon B. Vosper<sup>1</sup> , and Valery Yudin<sup>9</sup> 

<sup>1</sup>Met Office, Exeter, UK, <sup>2</sup>European Centre for Medium-Range Weather Forecasts, Reading, UK, <sup>3</sup>Atmospheric Numerical Weather Prediction Research Section, Environment and Climate Change Canada, Dorval, Québec, Canada, <sup>4</sup>National Centre for Meteorological Research, Météo-France, Toulouse, France, <sup>5</sup>Japan Meteorological Agency, Tokyo, Japan, <sup>6</sup>Deutscher Wetterdienst, Offenbach, Germany, <sup>7</sup>Korea Institute of Atmospheric Prediction Systems, Seoul, South Korea, <sup>8</sup>Korea Meteorological Administration, Seoul, South Korea, <sup>9</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

## Key Points:

- A multimodel comparison of resolved and parametrized orographic drag is used to evaluate the representation of orographic drag processes
- Models show an underestimation in total orographic gravity wave drag in the stratosphere at climate and seasonal prediction resolutions
- Some models show an overestimation in their parametrized orographic drag within the troposphere at climate resolutions

## Correspondence to:

A. van Niekerk,  
Annelize.vanNiekerk@metoffice.gov.uk

## Citation:

van Niekerk, A., Sandu, I., Zadra, A., Bazile, E., Kanehama, T., Köhler, M., et al. (2020). CONstraining ORographic Drag Effects (COORDE): A model comparison of resolved and parametrized orographic drag. *Journal of Advances in Modeling Earth Systems*, 12, e2020MS002160. <https://doi.org/10.1029/2020MS002160>

Received 29 APR 2020

Accepted 18 SEP 2020

Accepted article online 22 SEP 2020

**Abstract** The parametrization of orographic drag processes is a major source of circulation uncertainty in models. The CONstraining ORographic Drag Effects (COORDE) project makes a coordinated effort to narrow this uncertainty by bringing together the modeling community to: explore the variety of orographic drag parametrizations employed in current operational models; assess the resolution sensitivity of resolved and parametrized orographic drag across models; and to validate the parametrized orographic drag in low-resolution simulations using explicitly resolved orographic drag from high-resolution simulations. Eleven models from eight major modeling centers are used to estimate resolved orographic drag from high-resolution (km-scale) simulations and parametrized orographic drag from low-resolution simulations, typically used for seasonal forecasting (~40 km) and climate projections (~100 km). In most models, at both seasonal and climate resolutions, the total (resolved plus parametrized) orographic gravity wave drag over land is shown to be underestimated by a considerable amount (up to 50%) over the Northern and Southern Hemisphere and by more than 60% over the Middle East region, with respect to the resolved gravity wave drag estimated from km-scale simulations. The km-scale simulations also provide evidence that the parametrized surface stress and the parametrized low-level orographic drag throughout the troposphere are overestimated in most models over the Middle East region, particularly at climate resolutions. Through this process-based evaluation, COORDE provides model developers new valuable information on the current representation of orographic drag at seasonal and climate resolutions and the vertical partitioning of orographic low-level and gravity wave drag.

**Plain Language Summary** Numerical models used for seasonal, climate, or ensemble atmospheric predictions typically cannot resolve mountains with horizontal scales less than a few tens to hundreds of km, which are known to affect the accuracy of these models. As a result, their impacts on the winds, known as drag, are accounted for using simplified theory and approximations, which contain many uncertainties. This work sets out to investigate these uncertainties by comparing these approximations of the drag across several atmospheric models to the drag directly modeled in high-resolution simulations, in which more of the effects of mountains on the atmospheric flow are explicitly resolved. Results show that, in many of the models considered, the approximated mountain drag is underestimated in the upper part of the atmosphere but overestimated in the lower part of the atmosphere.

## 1. Motivation

Currently, short to medium range (up to 10 days) regional and global numerical weather predictions (NWP) are being performed using fine horizontal grid spacings of ~1–10 km. The atmospheric circulation is much better represented in these high-resolution simulations than with coarser grid spacing (tens to hundreds of km). Several studies have shown that the improvement in circulation, particularly over the Northern Hemisphere (NH) during winter, comes predominantly from the fact that the orography is better resolved (e.g., Berckmans et al., 2013; Kanehama et al., 2019). At these resolutions, important orographic effects

©2020. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

become resolved, and there is a lesser need to parametrize them. For example, resolving increasingly smaller orographic scales results in enhanced generation of gravity waves, which can propagate vertically, grow in amplitude, and have a significant nonlocal impact on the large-scale circulation in the stratosphere and middle atmosphere (Fritts & Alexander, 2003). The stratosphere is now well known to be important for medium-range weather prediction, seasonal forecasting and climate projections (Butler et al., 2019), and orographic gravity waves play a significant role in its accurate representation (Alexander et al., 2010). Better resolved orography also leads to a better representation of the low-level flow across mountains and valleys (Vanni re et al., 2019) and, more generally, of the large-scale circulation throughout the atmosphere (Kanehama et al., 2019).

Global circulation models also continue to be used across a range of resolutions for many applications, spanning multiple time scales. Given the elevated computing costs associated with long integration times and the requirement for large ensemble sizes, simulations for subseasonal to seasonal forecasting and climate projections are performed at relatively coarse horizontal grid spacing ( $\sim 20$ – $150$  km). At these resolutions, the model dynamics cannot accurately represent effects from orography at horizontal scales smaller than 3 to 8 times the grid spacing. Orographic drag parametrizations must therefore be used to represent the effects of unresolved orographic features on the atmospheric flow. Orographic drag parametrizations for processes such as gravity wave drag, low-level flow blocking, or turbulent orographic form drag have been shown to be key for model accuracy in the troposphere and stratosphere across time scales (see Sandu et al., 2019 and references therein). However, the representation of unresolved orographic drag processes within models remains uncertain (Sandu et al., 2019) due to the lack of observational or theoretical constraints on orographic drag within the troposphere and in the lower stratosphere at global or even at local scale.

Due to the lack of constraints and the importance of orographic processes for circulation, many of the choices made in the representation of unresolved orographic drag processes are the result of repeated tuning exercises aimed at improving forecast skill or model climate fidelity (Hourdin et al., 2017; Mauritsen et al., 2012; Williams et al., 2020). These choices include, but are not limited to: the processes which are parametrized: the setting of uncertain parameters within the parametrization schemes; and the degree of filtering of the mean orography, or how to derive subgrid-scale orographic fields. In recent years, community efforts have focused on narrowing the uncertainty in the representation of orographic drag processes in models through dedicated activities within the Working Group for Numerical Experimentation (WGNE) of the World Meteorological Organization and through a series of studies targeted at answering open questions related to these processes. A catalyst to this was the WGNE Drag project (Zadra et al., 2013), which demonstrated that, at global NWP resolutions (10–20 km), operational models differ in their total parametrized subgrid surface stress, particularly over land and especially over orography, but also in their partitioning between different drag processes. Consequently, processes such as turbulent drag over vegetative or urban canopies, turbulent orographic form drag over orographic features with scales smaller than 5 km, and low-level flow blocking and gravity wave drag associated with orographic features with scales larger than 5 km are being used interchangeably. Another WGNE effort later demonstrated that a considerable part of the intermodel differences in parametrized orographic surface stress is due to choices made in the filtering of resolved orography and derivation of subgrid orographic fields (Elvidge et al., 2019). In parallel, a series of studies revealed that intermodel differences in drag partitioning between different processes (partly the result of repeated tuning exercises) significantly affects the large-scale circulation over the NH (Pithan et al., 2016; Sandu et al., 2016) and the circulation response to climate change (van Niekerk et al., 2017). Further work also showed that, for at least a couple of models, the change in resolved orographic drag is not accurately balanced by the change in parametrized orographic drag when horizontal resolution is varied (Sandu et al., 2019; van Niekerk et al., 2016; Vosper et al., 2016, 2019). This is a key requirement for ensuring that circulation remains robust across resolutions, so that a model can be used seamlessly for a range of applications. Given the aforementioned findings, uncertainty in orographic drag processes is likely to be a major contributor to the intermodel spread in circulation (Shepherd, 2014).

A number of questions related to the representation of orographic drag processes still remain open. In particular, how do parametrized surface stresses differ between models at lower resolutions used for seasonal forecasting and climate projections? Do intermodel differences in drag partitioning result in differences in the vertical and regional distribution of drag? Can we get a better handle on parametrized orographic drag and on its partitioning between different processes from high-resolution simulations, rather than rely on

repeated tuning exercises? Is the poor handover between resolved and parametrized orographic drag when the resolution is varied a common feature for several models?

These questions have motivated the COORDE project, which takes place under the auspices of WGNE and Global Atmospheric System Studies (GASS). COORDE aims at further narrowing the uncertainties in the representation of orographic drag processes by coordinating the modeling community efforts to address some of the related open questions listed above. This is done by applying the approaches and techniques of van Niekerk et al. (2018) and Vosper et al. (2019) to a wider number of models from several operational modeling centers. In these previous studies, high-resolution (km-scale) simulations that explicitly resolve orographic low-level blocking and gravity wave effects to a large extent were used to evaluate the representation of orographic drag and its impacts on circulation in two operational models, that is, the Met Office Unified Model and the Integrated Forecasting System of the European Centre for Medium-Range Weather Forecasts (ECMWF). The accuracy of such km-scale simulations has been demonstrated using satellite-derived global observations of gravity waves in the stratosphere (Holt et al., 2017; Stephan et al., 2019), and employing them to constrain orographic drag processes constitutes one of the avenues outlined by Sandu et al. (2019) for reducing uncertainties in the representation of these processes.

COORDE uses coordinated simulations performed with 11 models spanning eight major modeling centers. Simulations are performed both at low resolutions, in which drag processes need to be parametrized (i.e., resolutions typically used for climate projections and seasonal forecasting) and at high (km-scale) resolutions, in which low-level flow blocking and gravity waves are largely resolved so that any remaining parameterized effect is significantly smaller than at low resolution.

The distinct aims of COORDE and this study are, thus, the following:

- (i) *Take stock of the variety of different orographic drag parametrizations employed in current operational models:* Implications of the choices made in representing orographic drag at low resolutions can be understood by comparing the vertical and horizontal distributions and diurnal cycles of the parametrized drag and surface stresses across the models.
- (ii) *Explore the resolution sensitivity of the resolved and parametrized orographic drag:* The handover between resolved and parametrized orographic drag as the resolution is varied is ascertained for the different models by comparing them at climate and seasonal forecasting resolutions.
- (iii) *Compare the parametrized orographic drag with explicitly resolved orographic drag from high-resolution simulations:* High-resolution simulations are used to evaluate the parametrized orographic drag and its partition between low-level drag effects and vertically propagating gravity wave drag in low-resolution simulations.

COORDE's experimental design, the participating models, and their orographic drag parametrizations are described in section 2. Section 3 then discusses the global distribution of resolved and parametrized orographic drag in the participating models, while section 4 focuses on a particular region of complex orography, namely, the Middle East (ME) region. For this region, the high-resolution simulations are used to perform a process-based evaluation of the parametrized orographic drag in the low-resolution experiments and to infer information about its partitioning between low-level drag and gravity wave drag. Conclusions are drawn in section 5.

## 2. Methodology

Results from simulations performed following the experimental protocol in section 2.2 were submitted by eight major operational modeling centers using 11 different models. The models are summarized in Table 1, along with their nominal grid spacing for each experiment, choice of initializing analysis, and relevant references. To address the first aim of this study, we begin by describing the different orographic drag parametrizations used in the models.

### 2.1. Orographic Drag Parametrizations

The last column of Table 1 lists the subgrid orographic drag processes, along with references to the particular parametrization schemes that are employed in the models. Within this set of models, there are a number of ways in which the subgrid orographic drag is represented. The orographic drag can generally be separated

**Table 1**  
Summary of the COORDE Models, Their Initializing Analysis, Nominal Grid Spacing, and References to Orographic Drag Parametrizations and Model

Center	Model	Grid spacing	Parametrizations and model reference
<b>ECMWF</b> Analysis: ECMWF	IFS	LR: 80 km MR: 40 km HR: 9 km (Global)	Orographic flow blocking and hydrostatic gravity wave drag: Lott and Miller (1997) Turbulent orographic form drag (explicit vertical distribution): Beljaars et al. (2004) Implementation: Part 4, Chapter 5 of ECMWF (2019) Model: ECMWF (2019)
		LR: 100 km MR: 40 km	Hydrostatic orographic gravity wave drag: McFarlane (1987) Orographic flow blocking: Lott and Miller (1997)
<b>Environment Canada</b> Analysis: Canadian Meteorological Centre	GDPS	LR: 100 km MR: 40 km	Turbulent orographic form drag (effective roughness): McTaggart-Cowan et al. (2019) Implementation: Zadra et al. (2015, 2003)
	RDPS	HR: 3 km (Regional)	
<b>Met Office</b> Analysis: ECMWF	UM	LR: 120 km MR: 36 km HR: 1.8 km (Regional)	Model: McTaggart-Cowan et al. (2019) Orographic flow blocking and hydrostatic gravity wave drag: Lott and Miller (1997) Turbulent orographic form drag (effective roughness): Wood and Mason (1993) Implementation: Webster et al. (2020); Vosper (2015)
		GSM1705	9 km (Global) LR: 100 km MR: 40 km
<b>JMA</b> Analysis: JMA	GSM2003	LR: 100 km	Orographic flow blocking and hydrostatic gravity wave drag: Lott and Miller (1997), Vosper (2015)
		MR: 40 km HR: 10 km (Global) LR: 100 km MR: 40 km HR: 6 km (Global)	Turbulent orographic form drag (explicit vertical distribution): Beljaars et al. (2004), Wood et al. (2001) Model: Yonehara et al. (2020) Orographic flow blocking, hydrostatic, and nonhydrostatic gravity wave drag: Kim and Doyle (2005) Turbulent orographic form drag (effective roughness): Richter et al. (2010), Koo et al. (2018) Implementation: Choi and Hong (2015) Model: Hong et al. (2018)
<b>DWD</b> Analysis: ECMWF	ICON	LR: 80 km MR: 40 km HR: 2.5 km (Regional)	Orographic flow blocking and hydrostatic gravity wave drag: Lott and Miller (1997) Turbulent orographic form drag (explicit vertical distribution): Beljaars et al. (2004) Implementation: Prill et al. (2019)
		FV3GFS	LR: 70 km
<b>NOAA</b> Analysis: ECMWF	WRF-ARW	HR: 3 km (Regional)	Model: Skamarock et al. (2008)
		ARPEGE	LR: 90 km MR: 40 km
<b>Météo-France</b> Analysis: ARPEGE	AROME	HR: 2.5 km (Regional)	AROME differs from ARPEGE only in using a nonhydrostatic switch in the dynamical core and different microphysics and shallow convection parametrization schemes.

*Note.* In the final column, “Implementation” refers to where details of the orographic drag parametrization implementation can be found, where available. “Model” refers to where the model is described.  
Abbreviations: AROME, Application of Research to Operations at Mesoscale; ARPEGE, Action de Recherche Petite Echelle Grande Echelle; DWD, Deutscher Wetterdienst; ECMWF, European Centre for Medium-range Weather Forecasts; FV3GFS, Finite-Volume Cubed-Sphere Global Forecast System; GDPS, Global Deterministic Prediction System; GSM1705, Global Spectral Model 1705; GSM2003, Global Spectral Model 2003; ICON, Icosahedral Nonhydrostatic; IFS, Integrated Forecasting System; JMA, Japan Meteorological Agency; KIAPS, Korea Institute of Atmospheric Prediction Systems; KIM, Korean Integrated Model; NOAA, National Oceanic and Atmospheric Administration; RDPS, Regional Deterministic Prediction System; UM, Unified Model; WRF-ARW, Weather Research and Forecasting Model; LR, low resolution; MR, middle resolution; HR, high resolution.

into that from vertically propagating hydrostatic gravity waves and low-level orographic drag processes. The vertically propagating gravity waves deposit momentum predominantly within the stratosphere and middle atmosphere, while the low-level orographic drag processes act to decelerate the flow within the troposphere in the vicinity of topography. It is worth noting that other nonorographic elements of the Earth's surface, such as urban or vegetative canopies, also apply drag at low levels over land. These are often treated using Monin-Obukhov similarity theory with roughness lengths derived from land cover data (Garratt, 1994). The uncertainties in these parametrizations, as well as the land cover data itself, are also large, and turbulent drag can, therefore, be used as a surrogate for low-level orographic drag and vice versa (Williams et al., 2020).

The drag from vertically propagating hydrostatic orographic gravity waves is the only component that is employed in all of the participating models. The models use similar expressions for the gravity wave surface stress, based on linear theory of a monochromatic wave (at horizontal scales of  $> \sim 5$  km) being generated over idealized topography, and most, excluding GSM1705 and GDPS, include some aspect of anisotropy of the subgrid orography. The treatment of wave breaking and saturation of the vertically propagating gravity waves is somewhat different between the models. For example, IFS, ICON, KIM, and FV3GFS employ a Richardson number wave saturation criterion (described in, e.g., Palmer et al., 1986), whereas, the UM, GDPS, ARPEGE, and GSM1705 employ an amplitude-based saturation criterion (as described in McFarlane, 1987). The ARPEGE model also accounts for nonlinear gravity wave effects, such as resonance and reflection (Roehrig et al., 2020).

The models differ more in their approach to low-level drag, likely due to the fact that these processes are nonlinear, can involve several co-occurring processes, and are difficult to observe and, therefore, to constrain. Low-level drag encompasses processes such as orographic low-level flow blocking, vertically trapped orographic waves, and turbulent drag, both from small-scale orographic features (turbulent orographic form drag) and from land-cover. All but one of the models, namely, JMA's GSM1705, employ an orographic low-level flow blocking parametrization. One of the first orographic low-level flow blocking drag parametrizations was introduced by Lott and Miller (1997) and is the basis for most of the schemes listed in Table 1. The formulation of this component of the orographic drag scheme is such that the depth of orographic flow blocking and the mountain wave amplitude is determined by a low-level Froude number. If the flow is too weak (or the flow is very stable and vertical displacement is suppressed) and is not able to go over the mountain to generate gravity waves, it is instead deflected around it. Within the vertical layer over which the flow is blocked, a bluff body drag is applied. With the exception of ARPEGE and GDPS, parametrizations that include a low-level flow blocking component are generally coupled to the vertically propagating component by limiting the amplitude of hydrostatic gravity waves when the flow is blocked.

Two of the models (GSM1705 and KIM) additionally employ a parametrization which accounts for drag due to nonhydrostatic gravity waves. This accounts for drag from vertically trapped waves generated by flow over topography that typically have horizontal scales  $\sim 10$  km, although in theory, trapping of waves is dependent on the flow conditions (i.e., through the Scorer parameter Crook, 1988). These schemes use the same expression for the nonhydrostatic as that of the hydrostatic wave surface stress but include some enhancement or scaling factor for the nonhydrostatic gravity wave stress. These waves are assumed to dissipate and deposit momentum at low levels.

Turbulent form drag from small-scale orography (typically horizontal scales of  $< \sim 5$  km) is also parametrized in several of the models, with some of them employing an effective roughness parametrization in which the vegetative roughness length is increased to represent the effects of subgrid orography (Wood & Mason, 1993) and others using an explicit vertical profile of the orographic stress (Beljaars et al., 2004; Wood et al., 2001). Throughout the manuscript, we use the terms parametrized orographic drag or SSO to refer to drag from subgrid-scale orographic low-level flow-blocking, low-level gravity wave breaking, and hydrostatic gravity wave drag, which are largely resolved in the km-scale simulations. We use BL to refer to turbulent orographic form drag and momentum fluxes over various land types within the boundary layer, which are parametrized in the low- and high-resolution simulations. Envelope orography is also employed in the ARPEGE model and acts to increase the volume of the resolved orography as a means of representing the blocking drag from unresolved orography (see Wallace et al., 1983) and thus is represented through the model dynamics.

**Table 2**  
*Summary of Experiments*

Experiment name	Description
LR CTL	Low-resolution (~100 km) control experiment with subgrid-scale orographic drag on
LR NOSSO	Low-resolution (~100 km) control experiment with subgrid-scale orographic drag off
MR CTL	Middle-resolution (~40 km) control experiment with subgrid-scale orographic drag on
HR CTL	High-resolution (~1.8 – 10 km) control experiment with subgrid-scale orographic drag off
HR LROR	High-resolution (~1.8 – 10 km) control experiment with subgrid-scale orographic drag off and low-resolution mean orography

*Note.* Refer to section 2.2 for further details.

The orographic processes accounted for in the models overlap to some extent, and many models use the Lott and Miller (1997) scheme as a basis for their low-level flow blocking formulation. However, the implementation of the orographic drag schemes (including Lott & Miller, 1997) also depends on many subjective choices, such as the tuning of uncertain parameters, the partitioning between different drag processes or flow regimes, and the vertical averaging of input variables. Moreover, the representation of parametrized orographic drag depends not only on the parametrizations themselves but also on the description of the sub-grid orography fields and the grid-box mean orography (Elvidge et al., 2019). To aid interpretation of results presented here, a reference to detailed technical descriptions of the implementation of the schemes in each model, where available, is also included in Table 1.

## 2.2. Experimental Design

COORDE's experimental design is similar to that described in van Niekerk et al. (2018) but is reiterated here for clarity. A series of experiments was designed to evaluate the parametrized orographic drag at climate (~100 km) and seasonal forecasting (~40 km) resolutions using high-resolution (1.8–10 km) global and regional simulations. Since we are interested in relatively fast processes and in attributing errors that are local to the orography, short-range forecasts initialized from analyses are performed. A set of fourteen 24-hr forecasts were produced for each of the experiments summarized in Table 2 from the models listed in Table 1. The 14 forecasts were initialized from the analyses listed in Column 1 of Table 1 at 00Z each day from 1–14 January 2015. A number of 14 forecast start dates was chosen to keep computational costs at a minimum for the participating modeling groups while also providing a large enough set to gauge systematic effects.

The abbreviations and a description of each experiment, given in Table 2, is described here. The nominal horizontal grid spacing used for the experiments are given in the third column of Table 1: LR indicates “low resolution” (~100 km); MR indicates “middle resolution” (~40 km); and HR indicates “high resolution” (1.8–10 km). These grid spacings are values taken within the midlatitudes and, vary globally depending on the choice of model grid. Experiments LR CTL and MR CTL are global simulations performed with the standard configuration of each model. Comparing the LR CTL and MR CTL simulations allows us to assess the resolution sensitivity of the resolved and parametrized orographic drag.

The LR NOSSO experiments are the same as LR CTL but with the subgrid-scale orographic (SSO) drag from orographic processes such as flow blocking, low-level gravity wave breaking, and hydrostatic gravity wave drag turned off. Note that turbulent boundary layer (BL) drag is left on in all the experiments. The impact of parametrized orographic drag on the circulation (on zonal wind) at low resolution can be deduced by differencing the sets of experiments LR CTL and LR NOSSO.

In order to determine the implications of using different model configurations for resolved orographic drag, we use a number of high-resolution simulations. This multimodel ensemble of high-resolution simulations allows us to assess the robustness of the resolved orographic drag to resolution, dynamical core, and physics. If there is little difference between the resolved drag obtained from the high-resolution simulations performed with the different models, this would support the use of such simulations in validating parametrized orographic drag at lower resolutions. Since it is computationally expensive to perform global simulations at horizontal grid spacings of 1.8–3 km with several models, some HR simulations were performed only over

the ME region, extending from 20.0–50.0°N and from 28.0–68.0°E. The ME region was chosen due to its complex orography and its position relative to the midlatitude jet, which makes it dynamically important.

The model setup for the HR CTL experiments is the standard configuration of each model but with the SSO turned off while BL effects still remain active. Since all the high-resolution models employ parametrized orographic drag in their operational NWP configuration, this setup is likely to degrade forecast scores. As discussed in van Niekerk et al. (2018), this choice was made to minimize interactions between the resolved drag and the parametrized orographic drag. The experiments HR LROR are identical to HR CTL except that the high-resolution model orography is replaced with a low-resolution orography. The low-resolution orography was generated by linearly regridding the mean orography from the global LR CTL onto the HR CTL grid. As proposed by van Niekerk et al. (2018), the impact on the flow from resolved orographic low-level drag and gravity wave drag (which should be parametrized in the LR CTL experiment) can be deduced by taking the difference between the HR CTL and HR LROR experiments. For some of the modeling centers, namely, Environment Canada, NOAA, and Météo-France, a different model is used for the low-resolution and high-resolution simulations. Except for a few resolution-dependent parameters, the atmospheric model sets of GDPS/RDPS and ARPEGE/AROME share the same dynamical-core configuration and (mostly) the same set of physical parametrizations. While the model configurations differ between FV3GFS/WRF-ARW, these high-resolution simulations are intended to provide further indication of the intermodel spread in resolved orographic drag.

By comparing the impact of parametrized orographic drag across the models, we can determine how the variety of parametrizations listed in Table 1, their implementation, and the intermodel differences in the subgrid orography fields affect the modeled flow. Then, by comparing the impacts of the resolved orography with those of the parametrized orographic drag, we can determine how credibly orographic effects on the flow are represented at climate and seasonal resolutions.

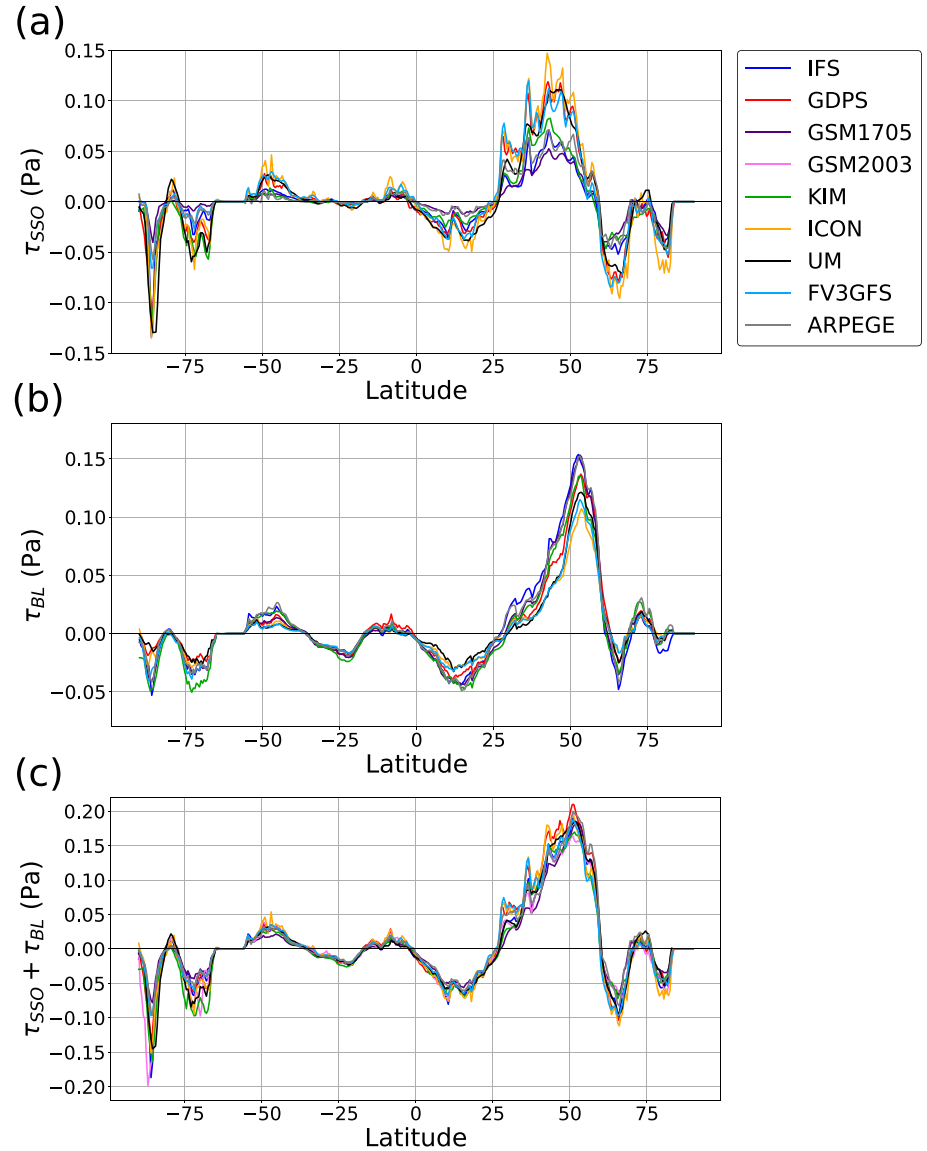
### 2.3. Diagnostics

As well as the circulation impacts (on the zonal wind) from resolved and parametrized orographic drag, relevant components of the zonal momentum budget are used to validate and compare output from the parametrizations. Two commonly used diagnostics that are produced by drag parametrization schemes and that can be computed from high-resolution simulations are the zonal wind surface stresses (vertical flux of zonal momentum at the surface) and vertical profiles of the deceleration of the zonal winds (the drag). The relationship between vertical flux of zonal momentum and deceleration is given by

$$\frac{du(z)}{dt} = -\frac{1}{\rho} \frac{\partial \tau(z)}{\partial z}, \quad (1)$$

where  $u$  is the mean zonal wind averaged over a particular area,  $\rho$  is density,  $\tau(z)$  is the vertical flux of zonal momentum, and  $z$  is altitude above sea level. The surface stress from parametrized boundary layer turbulence is denoted by  $\tau_{BL}$  and includes contributions from the turbulent mixing and orographic form drag within the boundary layer. The surface stress from parametrized subgrid-scale orographic drag is denoted by  $\tau_{SSO}$  and includes the stress from the orographic low-level flow blocking, vertically propagating gravity waves, and low-level wave breaking (depending on the model parametrizations). The vertical profiles of the parametrized zonal wind drag from the boundary layer turbulence and subgrid-scale orographic drag within the troposphere are denoted by  $\frac{du}{dt}_{BL}$  and  $\frac{du}{dt}_{SSO}$ , respectively. Away from the boundary layer, the parametrized drag is dominated by the vertically propagating gravity wave drag (GWD), which is denoted by  $\frac{du}{dt}_{GWD}$ . Note that  $\frac{du}{dt}_{GWD}$  is part of  $\frac{du}{dt}_{SSO}$ . The zonal wind drag was output on 16 pressure levels ranging from 1,000 to 10 hPa. Both the parametrized surface stresses and zonal wind drag were output as 6-hourly means.

To compare the total (resolved plus parametrized) orographic drag from the low-resolution simulations with the resolved orographic drag from high-resolution simulations, we also compute the resolved surface stresses and gravity wave drag in the free atmosphere in the UM and IFS. The resolved surface stress calculation follows the method described in Smith et al. (2005) and Vosper et al. (2019) using



**Figure 1.** Zonal mean parametrized zonal surface stress (land only) in LR CTL experiments from (a) low-level flow blocking and orographic gravity wave drag ( $\tau_{SSO}$ ), (b) turbulent friction and turbulent orographic form drag ( $\tau_{BL}$ ), and (c) their sum. Note that only the sum of  $\tau_{BL} + \tau_{SSO}$  is shown for GSM2003.

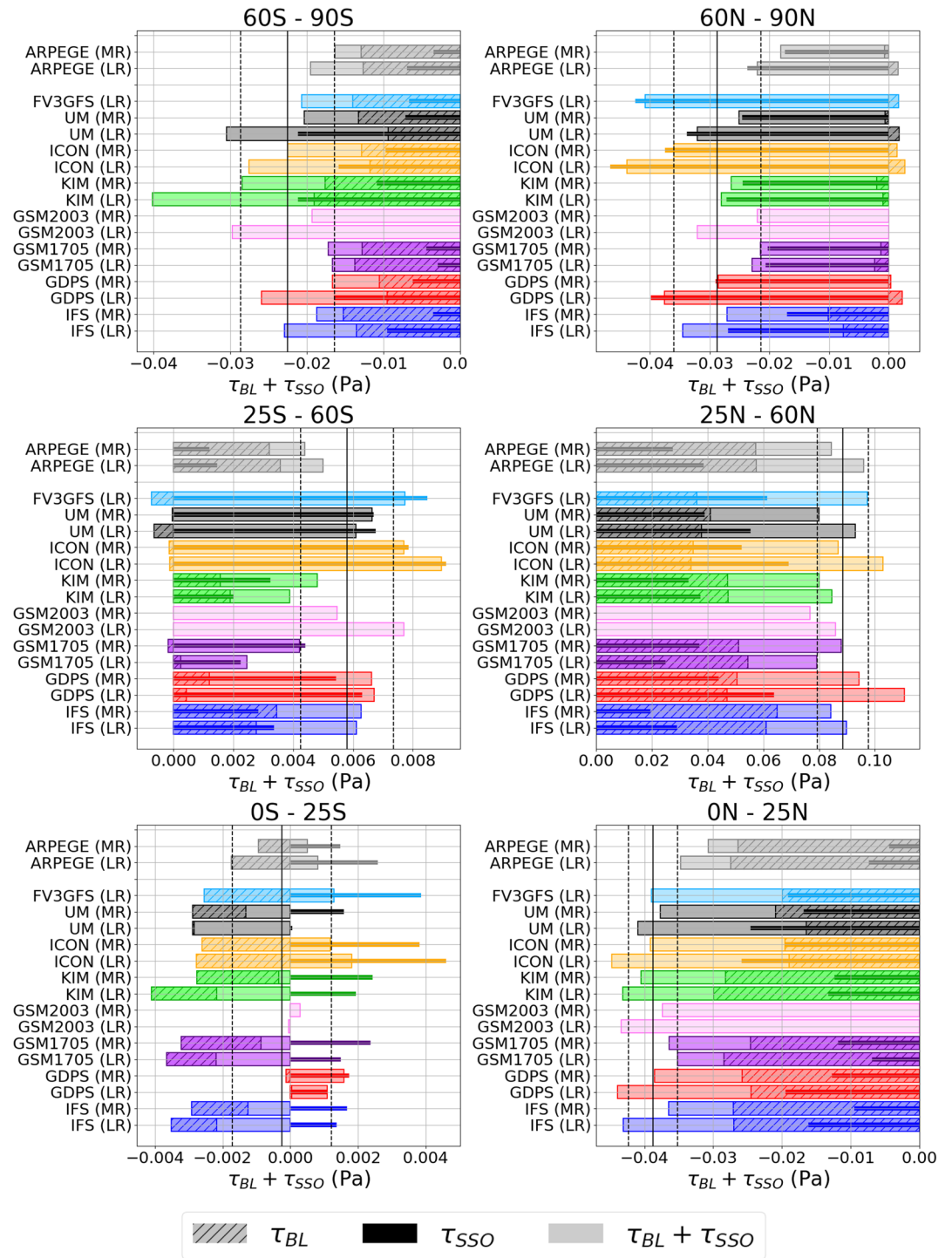
$$\tau_{RES}(z = 0) = A^{-1} \iint_A (p'_* - \bar{p}'_*) \frac{\partial}{\partial x} h(x, y) dx dy, \quad (2)$$

where  $z = 0$  denotes that the quantity is computed at the Earth's surface, the height of which varies longitudinally and latitudinally and is given by  $h(x, y)$ .  $A$  is the domain area, such as a hemisphere or geographical region, over which the quantity is integrated. The surface pressure perturbation  $p^*'$  is defined as the surface pressure ( $p_*$ ) after the removal of the standard atmospheric pressure:

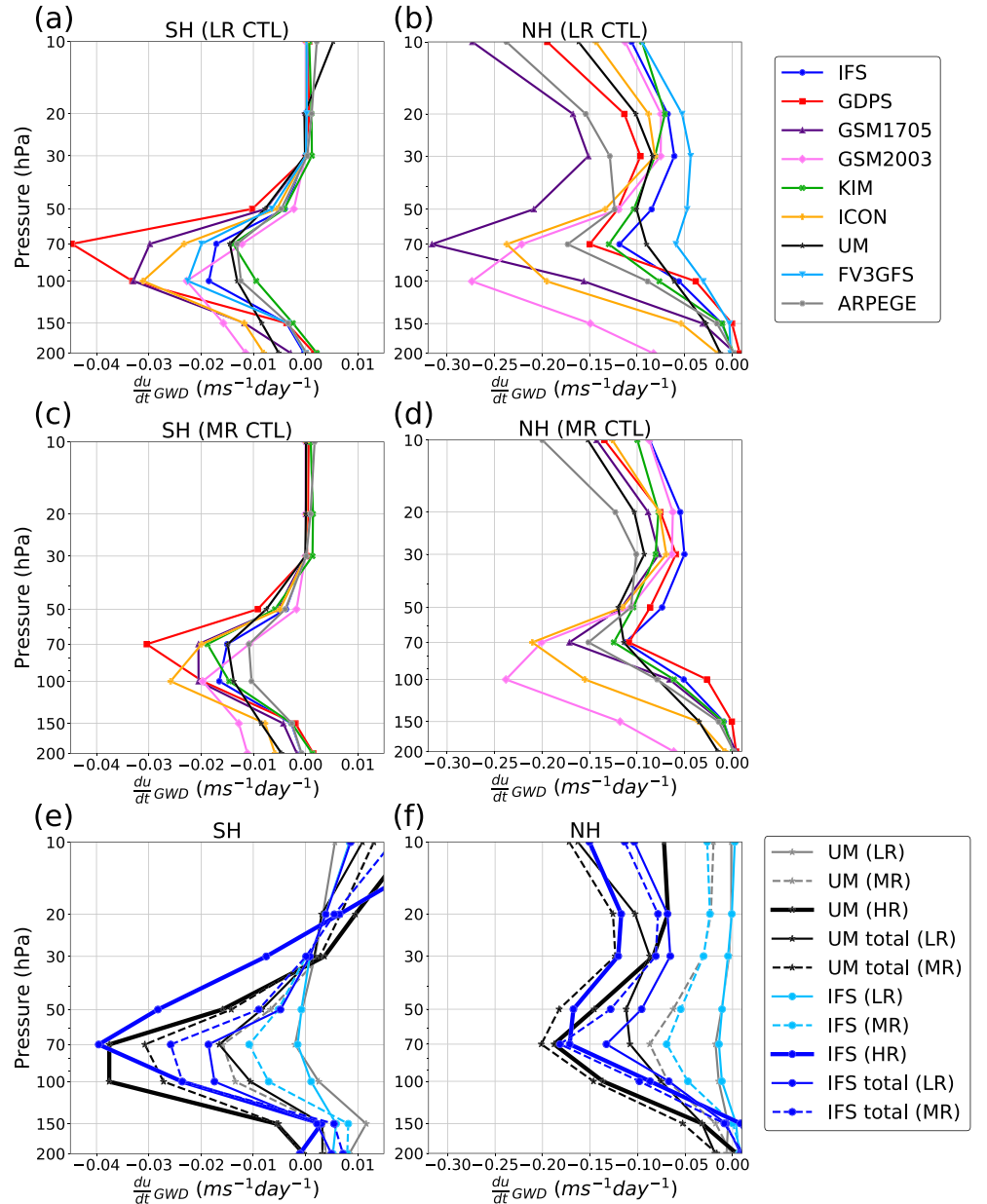
$$p_{std} = p_0 \left( 1 - \frac{\gamma h(x, y)}{T_0} \right)^{\frac{p_0}{\gamma T_0}}, \quad (3)$$

where  $p_0 = 101324 \text{ Pa}$  is the standard atmospheric pressure at mean sea level,  $T_0 = 288.15 \text{ K}$  is the reference temperature,  $\gamma = 0.0065 \text{ K m}^{-1}$  is an assumed constant temperature lapse rate in the troposphere,





**Figure 2.** Area-averaged parametrized zonal surface stress (land only) over different latitude bands, as indicated by titles.  $\tau_{BL}$ ,  $\tau_{SSO}$ , and their sum are shown, as indicated by the legend, for LR CTL and MR CTL. Solid and dashed vertical lines are the mean and standard deviation, respectively, of the total parametrized stress ( $\tau_{BL} + \tau_{SSO}$ ) across the models in LR CTL and MR CTL. Note that only the sum of  $\tau_{BL} + \tau_{SSO}$  is shown for GSM2003.



**Figure 3.** Area-averaged gravity wave zonal wind drag over (a, c, e) the Southern Hemisphere and (b, d, f) the Northern Hemisphere stratosphere from (a–d) parametrized gravity wave drag in all the models and (e, f) resolved/total gravity wave drag in the UM and IFS. Panels (a) and (b) show LR CTL models, and panels (c) and (d) show MR CTL models. In panels (e) and (f), model resolutions are indicated by the legend, where total is the sum of the resolved and parametrized gravity wave drag.

$g = 9.81 \text{ m s}^{-2}$  is the acceleration due to gravity, and  $R = 287.05 \text{ m}^2 \text{ s}^{-2} \text{ K}^{-1}$  is the gas constant of dry air. The overbar in Equation 2 represents the spatial mean over the domain  $A$ .

The resolved gravity wave momentum fluxes are computed using

$$\tau_{RES}(z) = A^{-1} \iint_A \rho_0 u' w' dx dy. \quad (4)$$

The perturbation velocities ( $u'$ ,  $v'$ , and  $w'$ ) are derived using the method described in Kruse and Smith (2015), whereby a high-pass filter is applied in spectral space to remove horizontal scales larger than  $\sim 1,600 \text{ km}$ . For regional momentum fluxes, the fields are detrended (via a linear planar fit) prior to high-pass

filtering. The background density field ( $\rho_0$ ) is determined by low-pass filtering the density field (removing scales smaller than 1,600 km). The resolved surface stress and momentum fluxes were computed from 6-hourly instantaneous values of  $p_*$ ,  $u$ ,  $w$ , and  $\rho$ , and the gravity wave drag is then computed via (1).

### 3. Global Drag Distribution

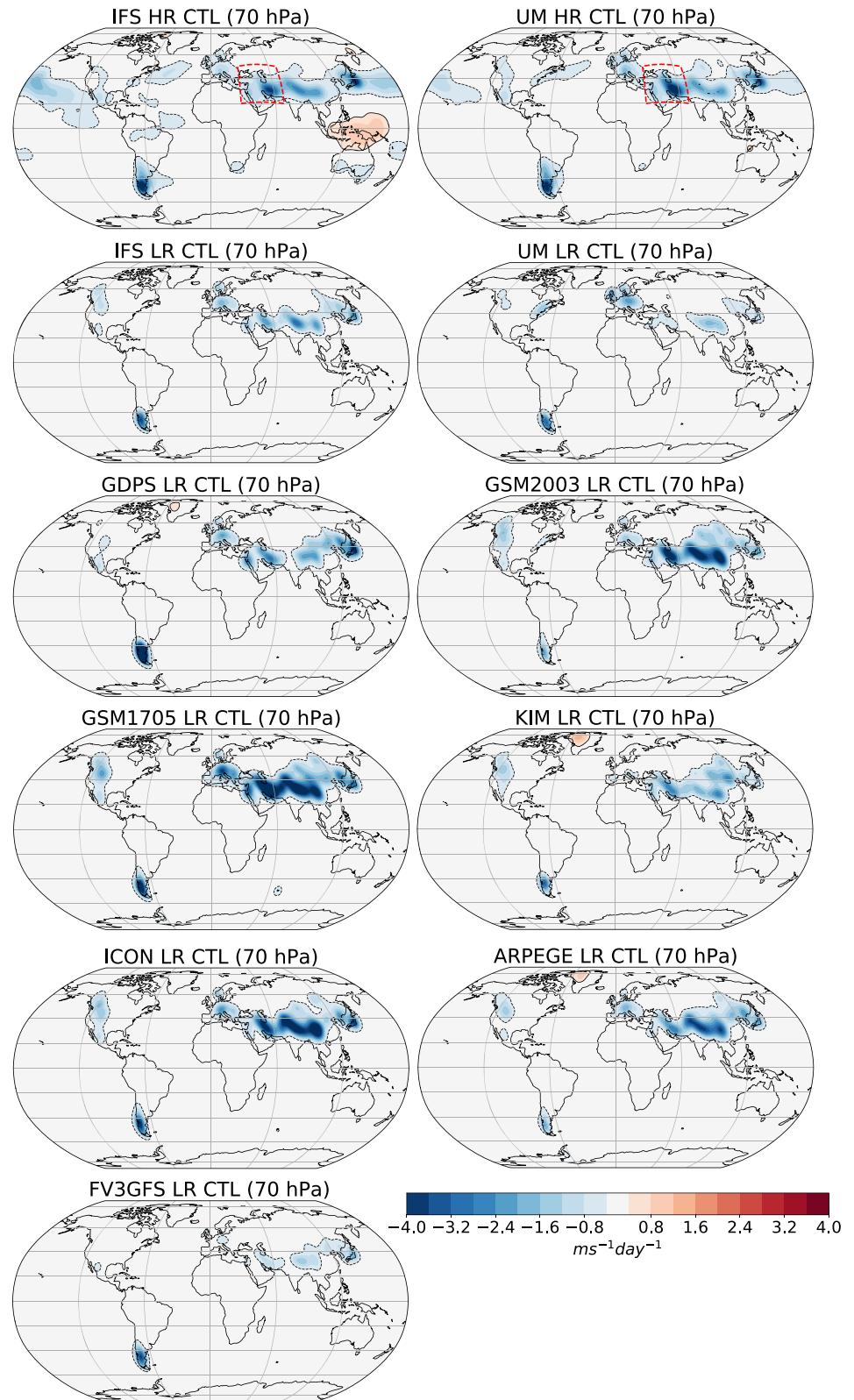
#### 3.1. Surface Stress

To compare the COORDE simulations, which are performed at climate and seasonal resolutions, to those in the WGENE Drag Project, which were at global NWP resolutions ( $\sim 16$  km), we begin by looking at the global distribution of the parametrized surface stress in the LR CTL and MR CTL experiments. Figure 1 shows the time-mean zonal-mean parametrized surface stress from the subgrid-scale orographic drag ( $\tau_{SSO}$ ), the boundary layer turbulent drag ( $\tau_{BL}$ ), and the sum of the two in LR CTL over land points only. This is comparable to Figure 3 of Sandu et al. (2019), but here we are considering simulations at much lower resolutions and over a different time period. Consistent with the results from the WGENE drag project, models with relatively less  $\tau_{SSO}$  over land tend to have more  $\tau_{BL}$ . The similarity in the total zonal mean parametrized surface stress is likely the result of repeated tuning exercises performed by each modeling group to minimize forecast errors in certain variables like mean sea level pressure, geopotential height, or surface winds, which are sensitive to surface drag and are key indicators for the large-scale circulation (Williams et al., 2020). Nonetheless, the models do still have quite large differences in their total parametrized surface stress over land, with a spread of about 0.05 Pa, approximately a third of the multimodel mean total, over the NH mid-latitudes. The differences in the total parametrized surface stresses are largest between 25°N and 50°N due to the difference in the latitudinal distribution of  $\tau_{BL}$  and  $\tau_{SSO}$ , since  $\tau_{BL}$  maximizes at higher latitudes compared with  $\tau_{SSO}$ .

The partitioning into  $\tau_{SSO}$  and  $\tau_{BL}$  over particular latitudinal bands, and their sensitivity to resolution, is shown in Figure 2. The model with the largest (and the smallest) total parametrized surface stress is different for each latitude band. This indicates that it is not a simple global scaling factor that controls the differences between the models but, instead, some regionally dependent features, such as stability or wind dependence of the schemes, the description of the subgrid orography, or land properties. For example, Elvidge et al. (2019) found that an important factor controlling the intermodel spread in  $\tau_{SSO}$  are the subgrid orographic fields used in the parametrization schemes, which will be highly dependent on the particular orographic region. Since the subgrid orographic fields depend on the mean orography, another factor that may lead to differences between the model parametrized stresses across regions is the type of grid used for each model and how its resolution varies with latitude.

In the NH high latitudes, the total parametrized surface stress over land is very much dominated by  $\tau_{SSO}$ , as opposed to  $\tau_{BL}$ , whereas the Southern Hemisphere (SH) high latitudes have a relatively equal contribution from both. This could be due to the fact that there are more large-scale mountains in the NH. In the NH subtropics and midlatitudes, the total parametrized surface stress varies by about 20% of the multimodel mean total parametrized stress, and its partitioning into  $\tau_{BL}$  and  $\tau_{SSO}$  varies greatly. GDPS, ICON, UM, and FV3GFS have larger  $\tau_{SSO}$ , while IFS, GSM1705, KIM, and ARPEGE have much larger  $\tau_{BL}$  between 25°N and 60°N.

Comparing LR CTL and MR CTL, one sees that the total parametrized surface stress generally decreases as the resolution is increased. This change with resolution is dominated by  $\tau_{SSO}$  as opposed to  $\tau_{BL}$ . This is an expected result since the SSO schemes account for orographic drag from orography with scales between  $\sim 5$  km and the model grid scale, whereas the BL schemes account for turbulent drag and turbulent orographic form drag from features with horizontal scales smaller than  $\sim 5$  km. When the resolution is increased from 100 to 40 km, the subgrid orography reduces because the grid-box mean orography increases, but the turbulent BL processes remain completely unresolved at both resolutions.  $\tau_{BL}$  does vary a little with resolution, but it could be argued that it increases at higher resolution in response to the reduction in  $\tau_{SSO}$ . Specifically, when  $\tau_{SSO}$  decreases with increasing resolution, the surface winds may increase, leading to larger  $\tau_{BL}$ . Because the resolution sensitivity is mostly from the SSO schemes, regions that are dominated by the SSO parametrization, such as the NH middle to high latitudes, also see a larger resolution sensitivity in their total parametrized surface stress. GSM1705 and KIM (over 25–60°S and 0–25°S only) are the only models



**Figure 4.** Resolved gravity wave drag at 70 hPa in global IFS HR CTL and UM HR CTL and parametrized gravity wave drag in LR CTL from all the models, as indicated by the titles. The parametrized gravity wave drag has been filtered using a low-pass filter with a length scale of 1,600 km, so as to make it comparable to the resolved gravity wave drag. Solid and dashed black contours separate positive and negative values, respectively.

that show an increase in  $\tau_{SSO}$ , and therefore the total parametrized surface stress, when the resolution is increased. The partitioning and resolution sensitivity of the SSO and BL schemes (or their underlying sub-grid orography and land fields) varies from region to region, a fact that needs to be considered when tuning or reformulating the parametrization schemes.

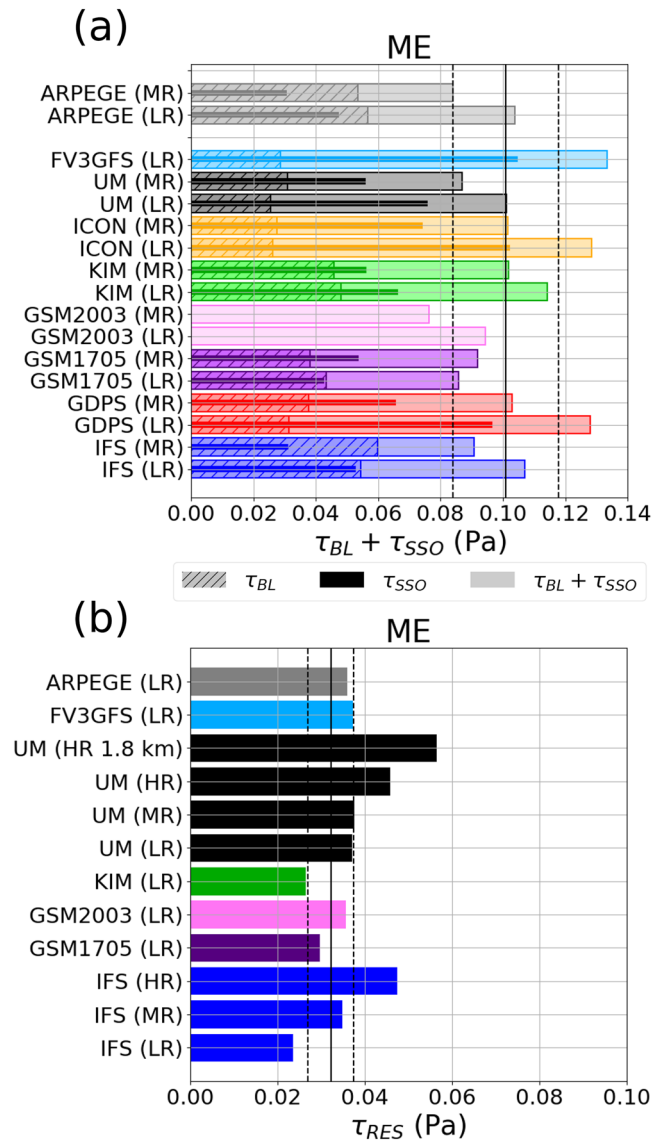
### 3.2. Gravity Wave Drag

What matters for the circulation is not only the regional distribution of the parametrized surface stress and its resolution sensitivity but also how it is distributed in the vertical, particularly since the orographic drag scheme includes a nonlocal vertically propagating gravity wave drag component. The vertical profiles of the parametrized gravity wave drag over the NH and SH stratosphere in LR CTL and MR CTL are shown in Figures 3a–3d. Over the NH, the spread in the gravity wave drag across the LR CTL models is large, but most models have a maximum just above the midlatitude jet (i.e., 70 hPa), where the waves are likely to break due to reduced windspeeds and density. The models do have some disagreement in the vertical distribution of gravity wave drag at higher altitudes (above 30 hPa), perhaps due to having different saturation criteria and different descriptions of subgrid mountain orientations or amplitudes. There are some notable outliers. In particular, GSM1705, GSM2003, and ICON have a larger maximum gravity wave drag magnitude ( $\sim -0.2$  to  $-0.3 \text{ ms}^{-1}\text{day}^{-1}$ ), and the FV3GFS model has a much smaller drag magnitude ( $\sim -0.05 \text{ ms}^{-1}\text{day}^{-1}$ ) over the NH in LR CTL. GSM2003 also has a peak drag that is at a lower altitude than in the other models. The peak drag in the UM's parametrized gravity wave drag is somewhat less pronounced than in the other models, which may be due to the fact that its parametrized gravity wave drag is applied over an approximated vertical wavelength of the wave. The gravity wave drag over the SH is much smaller than over the NH, since there is less orography here and the selected period is during SH summer. Unlike over the NH, GDPS LR CTL has the largest peak gravity wave drag over the SH, and KIM LR CTL has the smallest. Again, this implies that the parametrized gravity wave drag is strongly dependent on the region and not just on the scaling of the drag.

By comparing the parametrized and resolved gravity wave drag from the global simulations at resolutions of  $\sim 9$ ,  $\sim 40$ , and  $\sim 100$  km, we can determine the accuracy of the gravity wave saturation assumption, as well as the magnitude and resolution sensitivity, of the parametrized gravity wave drag. Figures 3e and 3f show the vertical distribution of resolved gravity wave drag averaged over the SH and NH stratosphere in the UM and IFS. The total (resolved plus parametrized) gravity wave drag in LR CTL and MR CTL in the UM and IFS is also shown. The peak magnitudes of the resolved gravity wave drag agree quite well in the two models at comparable resolutions, although there are some differences that are larger at seasonal resolutions, indicating that the resolved gravity wave drag does not depend much on the model used.

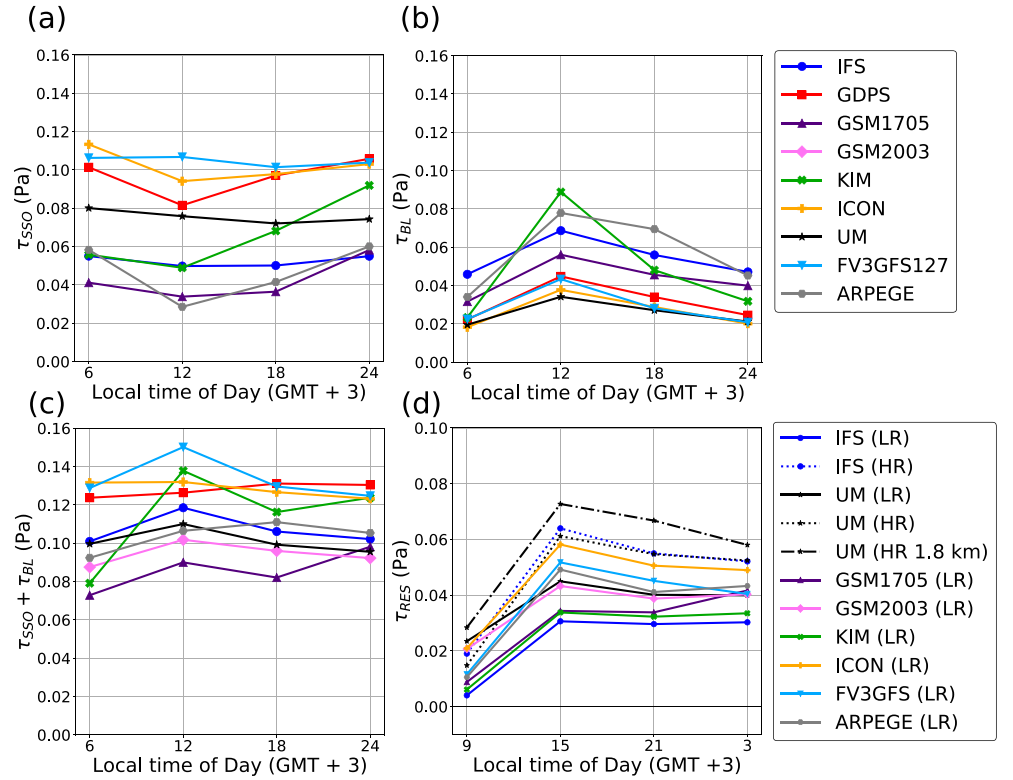
The total gravity wave drag in the UM and IFS LR CTL is about 30% less than the resolved gravity wave drag in the global 9 km HR CTL IFS and UM over the NH. In the MR CTL simulations, the total gravity wave drag in the UM and IFS seems to be reasonably well estimated, compared with the 9 km simulations, over the NH. Over the SH, the total gravity wave drag in UM and IFS is about 50% less in LR CTL than in the 9 km simulations and is about 30% less in MR CTL. The total gravity wave drag, therefore, seems to be more significantly underestimated over the SH than over the NH at both climate and seasonal resolutions. Assuming that the LR CTL and MR CTL simulations performed with the other models have similar resolved gravity wave drag to that in the UM and IFS at comparable resolutions, this implies that most of the models are underestimating the total gravity wave drag in the NH stratosphere at climate resolutions and in the SH stratosphere at both climate and seasonal resolutions. Models such as ICON, GSM1705, and GSM2003 do, however, seem to capture the peak magnitudes of gravity wave drag quite well over the NH, although GSM1705 LR CTL may be overestimating the drag. Only GSM1705, GDPS, and ICON, which have large parametrized drag, appear to be capturing the total gravity wave drag in the SH stratosphere. It is important to note that the 9 km simulations are not resolving the full spectrum of gravity waves and so the underestimation of the gravity wave drag in the LR and MR simulations is likely to be even larger than what has been quantified here, as will be discussed in section 4.2.

The resolved gravity wave drag is significantly increased when the resolution is increased from 100 to 40 km. Most of the parametrization schemes see a much smaller decrease in their parametrized gravity wave drag when the grid spacing is approximately halved. The parametrized gravity wave drag is, therefore, much less sensitive to resolution than the resolved gravity wave drag. Interestingly, both KIM and UM show an increase in parametrized gravity wave drag, although the increase in KIM is far more significant and only



**Figure 5.** Area-averaged zonal surface stress (land only) over the ME. (a) Parametrized surface stress from  $\tau_{BL}$ ,  $\tau_{SSO}$ , and their sum is shown, as indicated by the legend, for LR CTL and MR CTL. (b) Resolved surface stress  $\tau_{RES}$  in a subset of models at different resolutions. Solid and dashed vertical lines indicate the mean and standard deviation, respectively, of (a) the total parametrized surface stress ( $\tau_{BL} + \tau_{SSO}$ ) in LR CTL and MR CTL and (b) the resolved surface stress in LR CTL across the subset of models. Note that only the sum of  $\tau_{BL} + \tau_{SSO}$  is shown for GSM2003 in panel (a).

noticeable over the SH. The resolution sensitivity in KIM can be explained through an analysis of the subgrid orographic fields that are used as input to their gravity wave drag parametrization. The factor that determines the orographic asymmetry and which partly determines the gravity wave drag surface stress magnitude (see Choi & Hong, 2015 for details) increases and becomes more positive with increasing resolution. This leads to an increased scaling of the parametrized gravity wave drag at higher resolutions. The resolution sensitivity of the UM is also consistent with the results of Vosper et al. (2019), who found that the UM's parametrized gravity wave drag component was relatively insensitive to resolution over the Rocky mountains. This was shown to be as a result of the parametrization formulation. More precisely, the mountain wave amplitude at launch level (effective mountain height) is determined through a low-level Froude number:

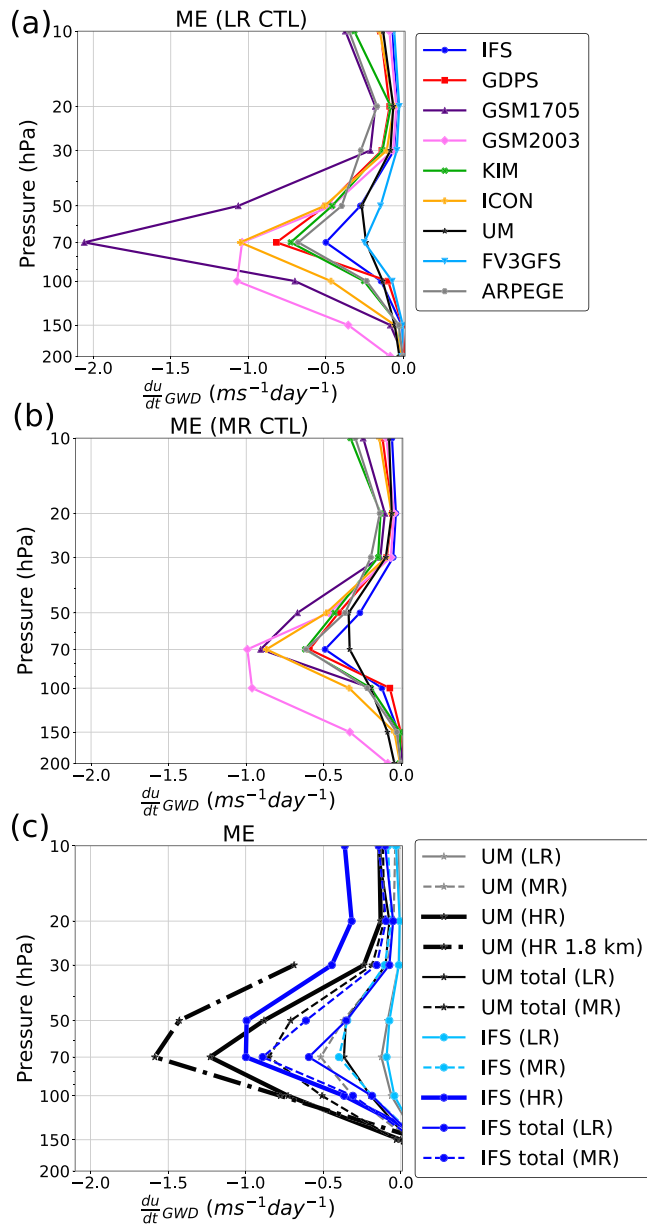


**Figure 6.** Mean diurnal cycle of the area-averaged zonal surface stress over the ME from (a)  $\tau_{SSO}$ , (b)  $\tau_{BL}$ , (c)  $\tau_{SSO} + \tau_{BL}$ , and (d) resolved  $\tau_{RES}$ . Panels (a)–(c) show results from LR CTL experiments. In panel (d), resolutions are indicated by the legend. Note that the time axes in panels (a)–(c) are different from those in panel (d) because the parametrized stresses are 6-hourly means, whereas  $\tau_{RES}$  is computed from 6-hourly instantaneous values.

$$h_{eff} = \min\left(h, \frac{U}{NF_c}\right), \quad (5)$$

where  $h_{eff}$  is the mountain wave amplitude,  $h$  is the subgrid mountain height, which is proportional to the standard deviation of the subgrid orography,  $U$  and  $N$  are the winds and stability over the height of the subgrid mountain, and  $F_c$  is a critical Froude number. This means that, in the blocked flow regime, the wave amplitude is given by  $U/(NF_c)$ , which does not vary much with resolution. The parametrized gravity wave momentum fluxes can, therefore, only vary with resolution in proportion to the wavenumber (inverse horizontal scale) of the subgrid orography. This horizontal scale is defined from inputs such as subgrid orographic gradients and standard deviation, computed within a model grid-box using a high-resolution orography data set. Since the dominant scale of the subgrid orography becomes larger at coarser resolutions, the scheme does not accurately capture the full spectrum of subgrid orography when resolution is varied. For both UM and KIM, the resolution sensitivity over the NH and SH can, therefore, be partly explained by changes in the “subgrid” orography at different resolutions.

GSM1705 is the only model for which the parametrized gravity wave drag decreases by a significant amount over the NH and SH when the resolution is increased. This is due to the fact that, although the gravity wave drag parametrization reduces the mountain wave amplitude at low Froude numbers using an expression very similar to that of (5), the critical Froude number ( $F_c$ ) in GSM1705 is set to a relatively small value of 1.5. This means that the flow is less likely to be in a blocked flow regime, allowing the gravity wave surface stress to vary proportionally with the square of the standard deviation of the subgrid orography, which increases at larger grid spacing. What is more, unlike most of the other gravity wave drag parametrizations, GSM1705 does not scale the gravity wave surface stress by a wavenumber derived from the subgrid orography, which becomes smaller at larger vegrid spacing.



**Figure 7.** Area-averaged gravity wave zonal wind drag over the ME stratosphere from (a, b) parametrized gravity wave drag in all the models and (c) resolved/total gravity wave drag in the UM and IFS. Panel (a) shows LR CTL experiments, and panel (b) shows MR CTL experiments. In panel (c), model resolutions are indicated by the legend.

The geographical distribution of the resolved gravity wave drag at 70 hPa from the 9 km IFS and UM (HR CTL) and parametrized gravity wave drag from the LR CTL experiments is shown in Figure 4. The parametrized and resolved gravity wave drag “hotspots” are quite well collocated. The regions with the largest drag are the ME region, the Himalayas, Japan, and the Southern Andes. Several models also show parametrized gravity wave drag over the Northern Rocky mountains, which is not as visible in the HR CTL IFS and UM. There also appears to be some resolved drag located over the oceans, which may be a result of nonorographic gravity waves generated over the storm tracks or lateral propagation of orographic gravity waves. Since none of the parametrizations used in the participating models account for lateral propagation and we are only interested in orographic gravity waves, sea points are masked out in Figure 3. Most of the models clearly have less parametrized gravity wave drag over land, particularly over orographic regions, than that present in the 9 km simulations. The models with a larger peak drag in Figure 3 show a distinct maximum over the ME and Himalayas in the NH and over the Southern Andes in the SH.

#### 4. Regional Drag Distribution

Given the large gravity wave drag seen over the ME region (indicated by a red box in Figure 4) within the upper troposphere to lower stratosphere, the following sections go on to investigate this region of complex orography in more detail. We make use of both the global high-resolution simulations (performed at 6–9 km) and the km-scale regional simulations (performed at 1.8–3 km) to better validate the parametrized orographic drag in the low-resolution simulations.

##### 4.1. Surface Drag Partitioning, Resolution Sensitivity, and Diurnal Cycle

We start by using the high-resolution simulations to evaluate the orographic surface stresses in the LR and MR simulations over the ME region. Insights about what the magnitude of the parametrized orographic surface stresses ( $\tau_{SSO}$ ) should be and how they should vary with resolution can be gained by comparing them with the resolved surface stresses ( $\tau_{RES}$ ; as given by Equation 2) at progressively higher resolutions. Figure 5a shows  $\tau_{SSO}$  and  $\tau_{BL}$  in LR CTL and MR CTL for the different models over the ME. The total parametrized surface stress ( $\tau_{SSO} + \tau_{BL}$ ) varies between the models over this region, with most of the models having a dominant contribution from  $\tau_{SSO}$ . It is only IFS and ARPEGE that have a dominant contribution from  $\tau_{BL}$ . Figure 5b then shows  $\tau_{RES}$  over the ME at different resolutions of the UM, the IFS, and in LR CTL for some of the models. In the LR CTL experiments,  $\tau_{RES}$  is quite similar across the models, as indicated by the vertical lines showing the standard

deviation, but the parametrized  $\tau_{SSO}$  in Figure 5a is not. The standard deviation of  $\tau_{RES}$  across the LR CTL experiments is  $\sim 0.008$  Pa, whereas the standard deviation of  $\tau_{SSO}$  in LR CTL is  $\sim 0.023$  Pa. This suggests that, at comparable resolutions, the parametrized orographic drag is more uncertain than the resolved orographic drag. There is an increase in  $\tau_{RES}$  in the UM and IFS when the grid spacing is reduced, although the increase from LR CTL to MR CTL is larger in the IFS.

The resolved surface stress in UM 1.8 km HR CTL is an indication of what the total surface stress from low-level orographic drag and gravity wave drag should be in the low-resolution simulations. The resolved  $\tau_{RES}$  increases by about  $0.028 \pm 0.008$  Pa when the resolution is increased from 100 to 1.8 km. Similarly, it increases by about 0.02 Pa from 40 to 1.8 km. This indicates that parametrized  $\tau_{SSO}$



should be  $\sim 0.03$  Pa in LR CTL and 0.02 Pa in MR CTL over this region. However,  $\tau_{SSO}$  in LR CTL and MR CTL is much larger than this in most models, suggesting that some of the models are significantly overestimating  $\tau_{SSO}$  at seasonal and climate resolutions.

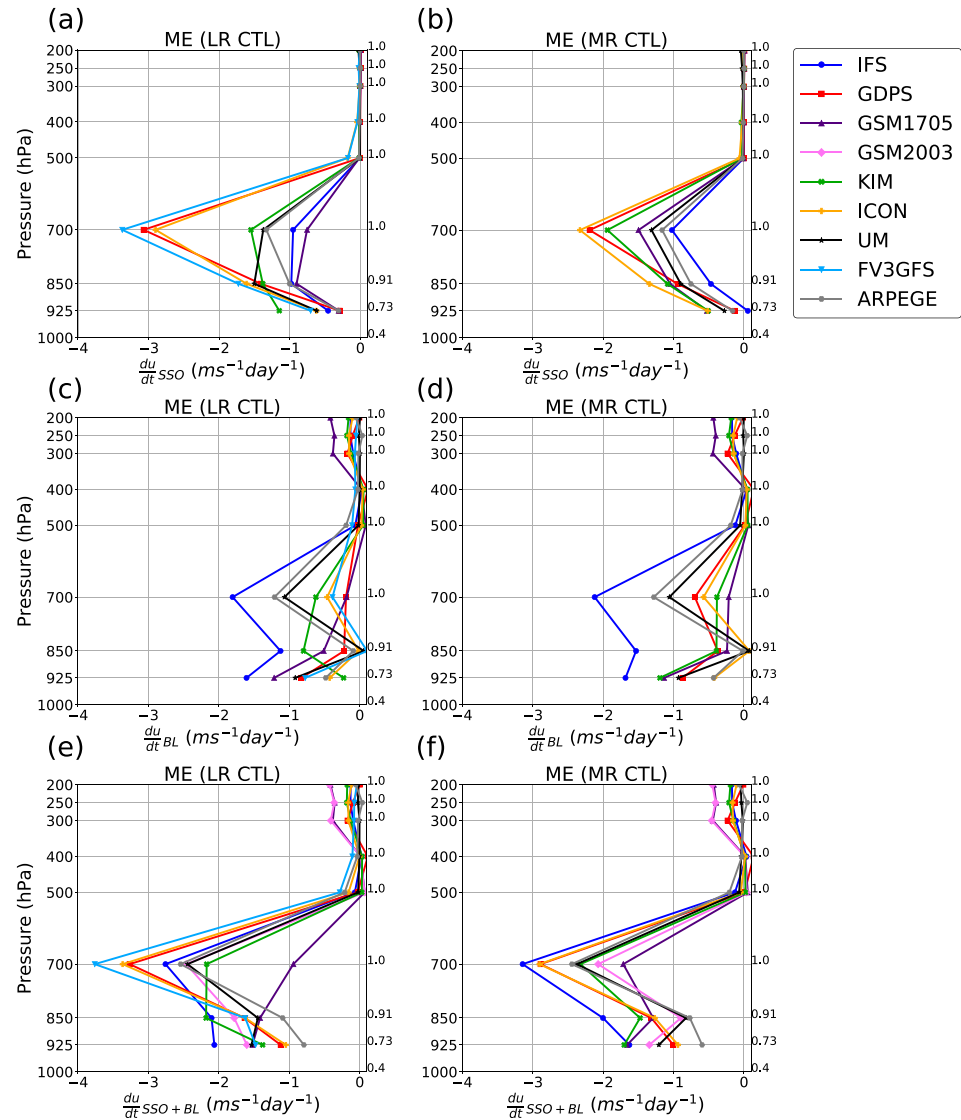
As has already been shown, there are large uncertainties in the partitioning of the drag into BL and SSO among models. Since these are processes that inherently act at different spatial scales and have different stability and wind dependencies, the diurnal cycles of the parametrized surface stresses produced by these schemes may differ. Figures 6a and 6b indeed shows that  $\tau_{BL}$  and  $\tau_{SSO}$  have very different diurnal cycles. In the middle of the day,  $\tau_{BL}$  maximizes when there is increased turbulent mixing and surface wind speeds. In contrast,  $\tau_{SSO}$  generally minimizes at midday because the reduced stability and increased winds are conducive to vertically propagating gravity waves instead of low-level flow blocking. The KIM model has a very large diurnal cycle in its total parametrized surface stress (Figure 6c) when compared with the other models. This is partly due to the contribution from  $\tau_{BL}$ , which increases during the day and is relatively small at night. Compared with models that have similar time-mean  $\tau_{BL}$ , such as the IFS or ARPEGE, the diurnal cycle in KIM is larger. This suggests that it is the dependence of  $\tau_{BL}$  on stability or winds that makes it stand out, particularly the turbulent orographic form drag component (Koo et al., 2018). Unlike  $\tau_{SSO}$ , the resolved  $\tau_{RES}$  (Figure 6d) gets larger during the daytime and smaller during the evening in all of the models. This difference between  $\tau_{SSO}$  and  $\tau_{RES}$  could either be as a result of interaction between the parametrized  $\tau_{BL}$ ,  $\tau_{SSO}$  and  $\tau_{RES}$  or because  $\tau_{SSO}$  does not have the correct dependence on winds and/or stability. There is some indication that models with larger  $\tau_{BL}$  also have smaller  $\tau_{SSO}$  during the daytime. This is potentially because large  $\tau_{BL}$  weakens the near-surface winds, and since  $\tau_{SSO}$  depends on wind, this results in smaller  $\tau_{SSO}$  during the daytime. This interaction between the two schemes can lead to difficulties in correctly constraining their contributions. The differences in the diurnal cycles of the parametrized orographic drag and the resolved orographic drag shown here, and the choice of partitioning among the models, are likely to play a role in the diurnal cycle of the surface wind.

#### 4.2. Gravity Wave Drag Profiles

The resolved gravity wave drag in the UM 1.8 km HR CTL is now compared with the parametrized gravity wave drag in LR CTL and MR CTL experiments over the ME. This will allow us to gauge the validity of the conclusions drawn from the global (9 km) HR CTL simulations in section 3.2. Figure 7 shows the parametrized and resolved gravity wave drag in the upper troposphere to lower stratosphere in all the models, averaged over the ME. The spread of the parametrized gravity wave drag at  $\sim 70$  hPa in LR CTL and MR CTL is representative of what is seen over the NH in Figure 3, which is not surprising given that a large proportion of the NH gravity wave drag at those altitudes is coming from the ME region during this period (Figure 4).

The resolved gravity wave drag is evidently increased in UM 1.8 km HR CTL, compared with the global HR CTL IFS and UM. In the UM, the peak resolved drag increases by about 25% when the resolution is increased from 9 to 1.8 km. This implies that the underestimation of the total gravity wave drag over the NH and SH in the LR CTL and MR CTL models is even more substantial than what was identified in section 3.2. The resolved gravity wave drag in the UM and IFS are not as similar at comparable resolutions (particularly in MR CTL and HR CTL) over this region compared with the entire NH (Figure 3), since the IFS has slightly less resolved gravity wave drag than the UM. This is compensated by the IFS having slightly more parametrized gravity wave drag over this region. This is opposite to what was found in Vosper et al. (2019) over the Rocky mountains, where the resolved/parametrized gravity wave drag was larger/smaller in the IFS compared with the UM in global 16 km simulations. It is possible that this difference is due to: variability, since we are looking at a slightly different time period; regional difference in resolved orography; difference in resolution, since we are looking at 100 and 40 km resolutions instead of 16 km simulations; or/and possible interaction between the resolved and parametrized orographic drag.

The peak deceleration from the regional UM 1.8 km HR CTL over the ME is  $\sim 1.5 \text{ ms}^{-1}\text{day}^{-1}$ , whereas the total from resolved and parametrized gravity wave drag in most of the models is  $\sim 0.5 \pm 0.25 \text{ ms}^{-1}\text{day}^{-1}$  in LR CTL and  $\sim 0.875 \text{ ms}^{-1}\text{day}^{-1}$  in MR CTL. Most of the MR CTL models have about half the total gravity wave drag (resolved plus parametrized) and the LR CTL models have about a third of the total gravity wave drag compared with the resolved in the regional UM 1.8 km HR CTL. The profile and magnitude of the parametrized gravity wave drag in GSM1705 are more similar to the 1.8 km UM although somewhat

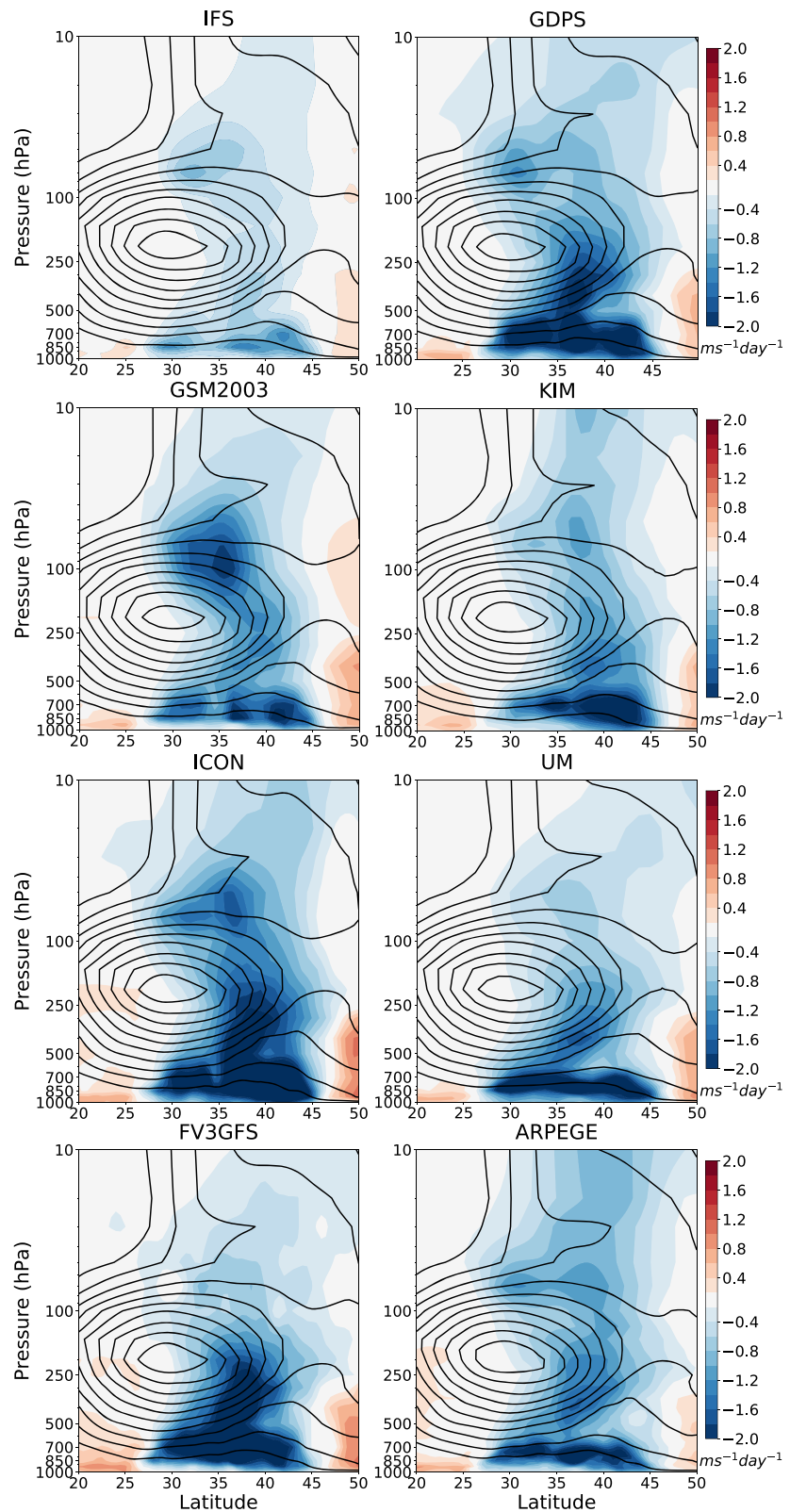


**Figure 8.** Area-averaged zonal wind drag profiles over the ME troposphere interpolated onto pressure levels from (a, b)  $\frac{du}{dt}_{SSO}$ , (c, d)  $\frac{du}{dt}_{BL}$  and (e, f) the sum of  $\frac{du}{dt}_{SSO} + \frac{du}{dt}_{BL}$ . Panels (a), (c), and (e) show LR CTL, and panels (b), (d), and (f) show MR CTL. Since the drag is interpolated to pressure levels for each model, the pressure levels that are lower than the mean orography are masked out. Values quoted on the right-hand axes indicate the fraction of area that is not masked out.

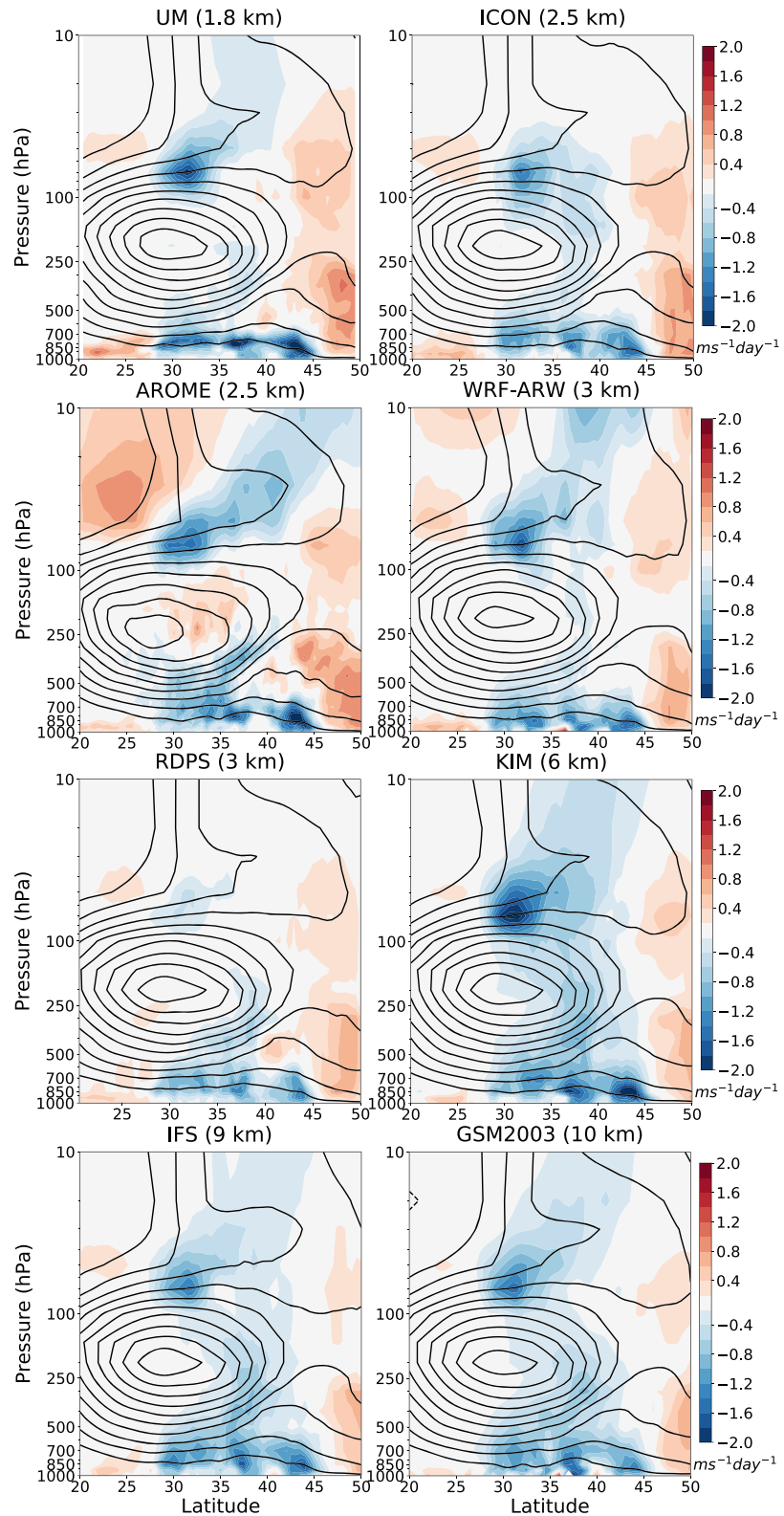
overestimated in LR CTL. As was found over the NH, the resolution sensitivity of the parametrized gravity wave drag (Figures 7a and 7b) is much weaker than that of the resolved gravity wave drag (Figure 7c) in most of the models. The IFS's parametrized gravity wave drag is almost identical in LR and MR. Consistent with the discussion in section 3.2, the UM's parametrized gravity wave drag increases with increasing resolution over this region. In contrast, the parametrized gravity wave drag in GSM1705 and GDPS has a resolution sensitivity comparable to that of the resolved gravity wave drag.

### 4.3. Tropospheric Drag Profiles

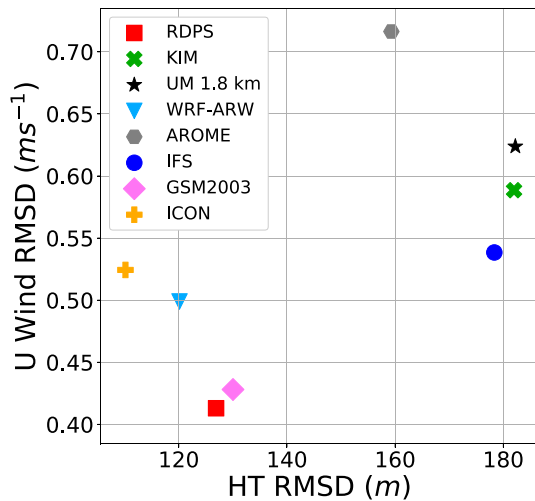
It is much more difficult to evaluate the vertical profile of the drag within the troposphere using high-resolution simulations than it is to evaluate the stress at the surface and the gravity wave drag in the free atmosphere. Resolved momentum fluxes and drag within the troposphere are difficult to diagnose because separating the drag processes from their impact on circulation near the surface is complicated. Nonetheless, it is of interest to look at the parametrized drag profiles within the troposphere so as to



**Figure 9.** Impact of parametrized orographic drag (LR CTL minus LR NOSO) on the zonal wind at the end of the 24 hr forecasts, averaged over the set of 14 forecasts and longitudinally averaged over the ME (shading). Thick black contours are the mean zonal wind in LR CTL, with a  $5 \text{ ms}^{-1}$  contour interval.



**Figure 10.** Impact of resolved orographic drag (HR CTL minus HR LROR) on the zonal wind at the end of the 24 hr forecasts, averaged over the set of 14 forecasts and longitudinally averaged over the ME (shading). Thick black contours are the mean zonal wind in HR CTL, with a  $5 \text{ ms}^{-1}$  contour interval.



**Figure 11.** Relationship between the RMSD between the HR CTL and HR LROR experiments for the zonal wind (U) and mean orographic height (HT) averaged over the ME region. The zonal wind differences are averaged between 1,000 and 10 hPa.

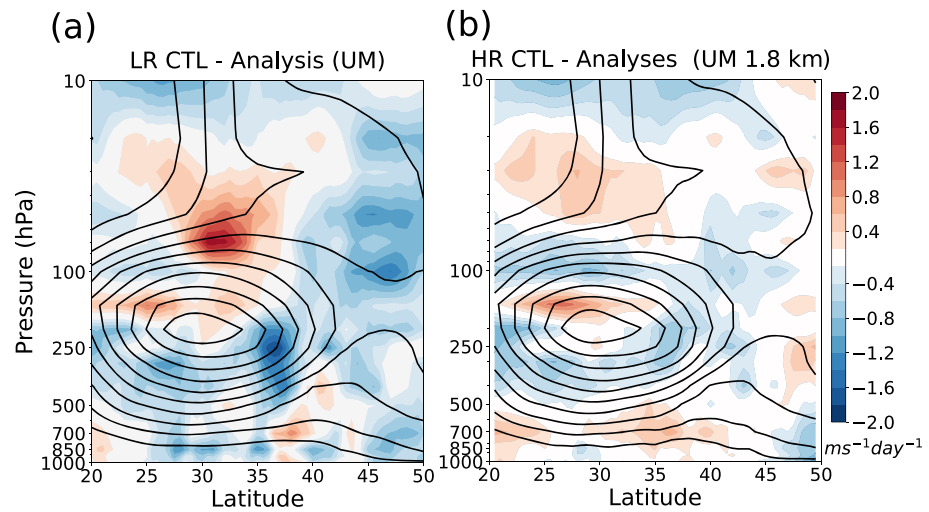
determine to what extent the vertical distribution of the total drag is influenced by the choice of parametrization and partitioning between processes. Considering first the LR CTL simulations (Figure 8), we can see that the SSO drag profiles are quite different from the BL drag profiles. The SSO drag has a large maximum at about 700 hPa with small values near the surface. This is because the levels below 700 hPa are mostly subterranean (and so are masked out) and because the orographic flow blocking scheme depends on the vertical profile of the winds, which are increasing with height. The BL drag also has a maximum at about 700 hPa and a secondary maximum at 925 hPa. FV3GFS has considerably more SSO drag at 700 hPa and less BL drag throughout the troposphere, leading to large total parametrized drag overall. GSM1705 has much less SSO drag but also much less BL drag. It is clear that the profiles of both the SSO and BL drag are different in GSM1705 in the sense that they lack a peak deceleration at 700 hPa. This is most likely due the fact the vertical profile of stress from the low-level wave breaking parametrization is assumed to be a quadratic function of pressure rather than having a vertical dependence on windspeed, as is the case for most orographic flow blocking schemes.

Both KIM and GSM1705 employ a low-level wave breaking parametrization, and although KIM also has a flow blocking parametrization, the low-level wave breaking makes up  $> \sim 60\%$  of the drag in the troposphere in KIM (not shown). In both these models, the SSO drag increases with increasing resolution at 850–500 hPa. The resolution sensitivity in KIM is likely related to the fact that, while the total surface stress reduces at higher resolution, the altitude of momentum deposition is higher and therefore imparts larger deceleration of the zonal winds (as the density is decreased there). In GSM1705, this resolution dependence is a result of resolution-dependent tuning parameters in the low-level wave breaking scheme, designed to optimize higher resolution forecasts. Other models show an expected resolution sensitivity in their SSO drag, with decreasing parametrized orographic drag at higher resolution and little sensitivity in their BL drag profiles. In some models, namely, IFS and GDPS, the BL drag increases with resolution at 850–700 hPa, suggesting that this is a response to the SSO drag decreasing with resolution. Overall, the choice in partitioning and parametrized process significantly affects the profile of parametrized drag and, as shall be seen, the zonal wind within the troposphere.

#### 4.4. Resolved Versus Parametrized Orographic Drag

As discussed in section 4.3, diagnosing the vertical profile of resolved orographic drag within the troposphere is difficult. However, the method used in van Niekerk et al. (2018) provides a way of determining the impact of resolved and parametrized orographic drag on the winds and, thus, an indirect way of evaluating parametrized orographic drag within the troposphere. Figure 9 shows the impact of parametrized orographic drag in the low-resolution simulations (i.e., LR CTL - LR NOSSO) on the zonal winds longitudinally averaged over the ME at the end of the 24 hr forecasts (averaged over the 14 forecasts). These plots, therefore, represent the change in the winds over the 24 hr period of the experiment that result from the SSO parametrization schemes (including the subgrid orography fields) and the response in the model dynamics to those schemes. Since these are differences at short lead times, the impacts remain local to the proximity of the mountain chain and are, therefore, predominantly due to the ME mountains.

The diversity of the impact on the zonal wind (Figure 9) is a reflection of the various parametrization schemes outlined in Table 1 and differences in partitioning between SSO and BL drag, as well as in the subgrid orographic fields. The vertical depth of the impact on the zonal wind varies between the models. Most of the models show a very deep deceleration of the winds, extending throughout the troposphere, which comes primarily from the low-level drag components of the parametrizations and the response from the dynamics, as was demonstrated in van Niekerk et al. (2018). Clearly, models with large SSO drag within the troposphere (e.g. FV3GFS, GDPS, and ICON in Figure 8a) show a larger deceleration of the zonal wind there. In the lower stratosphere (between 150 and 10 hPa), there is an indication of deceleration from the gravity wave drag component of the schemes, with the magnitude and distribution (in the vertical and horizontal) varying between



**Figure 12.** Zonal wind error (model minus analysis) in (a) UM LR CTL and (b) UM 1.8 km HR CTL at the end of the 24 hr forecasts, averaged over the set of 14 forecasts and longitudinally averaged over the ME (shading). Thick black contours are the mean zonal wind in the model, with a  $5 ms^{-1}$  contour interval.

the models, consistent with Figure 7a. In particular, GSM2003 has a large deceleration in the lower stratosphere that is at a lower altitude than in the other models. The UM and FV3GFS both have a very weak deceleration in the stratosphere, corroborating the discussion in section 4.2.

Our ensemble of high-resolution experiments with high-resolution (HR CTL) and low-resolution (HR LROR) orography at grid spacing ranging from 1.8 to 3 km regionally, and at grid spacing ranging from 6 km to 10 km globally, can be used to quantify the impact from resolved orography (shown in Figure 10). This impact represents the change in the zonal winds resulting from resolving the effects of orographic features with scales smaller than 100 km and up to the model resolved scales, which would need to be parameterized at 100 km resolutions. The impact of resolved orography on the zonal wind is largest in the lower stratosphere due to resolved orographic gravity wave breaking occurring at the tropopause, where there is a rapid decrease in density and wind speed with height. A large impact can also be seen near the surface due to low-level orographic drag effects.

The intermodel differences in the resolved orographic impacts deduced from the high-resolution experiments, seen in Figure 10, can be partly explained by the differences in the mean orography between the HR CTL and HR LROR experiments. One would expect that a larger difference in the mean orography between the HR CTL and HR LROR experiments would lead to a larger impact on the zonal winds. Indeed, Figure 11 suggests that there is some relationship between the root-mean-squared differences (RMSD) in the mean orography and the impact on the winds. The largest orographic HT RMSD is in UM 1.8 km and KIM, which also show large wind responses to the resolved orography. GSM2003 and RDPS have relatively smaller HT RMSD, and correspondingly, the wind differences are smaller. ICON, WRF-ARW, and AROME do not entirely fit this relationship, however. For example, ICON has one of the highest resolutions for the HR runs (2.5 km), and yet the orographic HT RMSD is lower than in IFS HR experiments (at 9 km) while the wind response is similar. The higher HT RMSD between the IFS HR experiments, despite their lower resolution (9 km), compared to the other HR experiments could be explained by the fact that the mean orography is less filtered in IFS than in other global models (Elvidge et al., 2019). The similarity of the wind responses in IFS and ICON suggests, however, that the change in horizontal resolution between the HR CTL and HR LROR orography is not the only factor determining the wind response. Numerical aspects, such as the damping applied to vertical winds, time step, advection scheme, and the vertical or horizontal resolution of the models are likely to also play a role in the magnitude of the impact of the resolved orography on the wind field.

Despite the differences in the magnitude (shown in Figure 11), the spatial distribution of the zonal wind response to the resolved orography is quite similar across the models compared with that from

parametrized orographic drag, particularly at the location of maximum gravity wave breaking within the stratosphere. This gives us some confidence in using these high-resolution simulations for validation of the parametrized orographic drag at coarser resolutions. It also reinforces the analysis performed and conclusions drawn from the 1.8 km UM simulations in section 4.2.

To place these impacts from resolved and parametrized orographic drag into context and to demonstrate the fidelity of the high-resolution simulations, Figure 12 shows the zonal wind errors after 24 hr relative to analysis in the UM LR CTL and UM 1.8 km HR CTL averaged over the ME. These fields give some indication of the systematic errors over the considered period. The impact from resolved and parametrized orographic drag (Figures 9 and 10) is of the same magnitude as the model errors in UM LR CTL in Figure 12, highlighting the dominant role that the orography plays in circulation accuracy over complex terrain. The improvements gained by increasing the resolution are clear, since the negative wind errors near the surface and positive wind errors in the lower stratosphere seen in UM LR CTL (Figure 12b) are almost completely removed in UM 1.8 km HR CTL (Figure 12a). There is still some indication that the near-surface winds are too strong in UM 1.8 km HR CTL, which is likely to be alleviated if the orographic drag parametrization was turned on. Nonetheless, the small errors in Figure 12b illustrate that the high-resolution simulations, even with parametrized orographic drag turned off, are close to our best guess of the atmospheric state (the analysis), further supporting their use for the validation of the parametrized orographic drag and its impacts on circulation.

The impact on the zonal winds from resolved orography in the HR simulations can then be compared with that from parametrized orographic drag in the LR simulations as a means of validating the representation of orographic drag at climate resolutions. There are clear differences. The impact from the parametrized orographic drag in Figure 9 near the surface is far larger than what is seen from the resolved orography in Figure 10. While some impact on the zonal winds from the resolved orography can be seen throughout the troposphere, it is far less pronounced than that induced by parametrized drag. This is an indication that the parametrized orographic drag near the surface is too large in most of the models, excluding perhaps ARPEGE and IFS, at the low resolutions. This is consistent with the conclusions drawn from the resolved and parametrized surface stresses in Figure 5 and the negative near-surface zonal wind errors in UM LR CTL (Figure 12a). In most of the high-resolution simulations, the impact on the wind from resolved gravity wave breaking peaks at a height of 150–30 hPa and between the latitudes 30°N and 35°N (Figure 10), whereas that from the parametrized gravity wave drag peaks at a latitude between 35°N and 40°N (Figure 9). It is only in IFS, GDPS, and ARPEGE that the impact from parametrized gravity wave drag peaks at roughly the same latitude as that from resolved orography within the stratosphere. In van Niekerk et al. (2018), this northward displacement of the impact from parametrized gravity wave drag was found to be partly due to the response of the resolved dynamics to the parametrized drag, which is also likely to be the case here.

The MR CTL experiments also generally show positive zonal wind errors within the lower stratosphere (indicative of insufficient gravity wave drag) but have much smaller near-surface wind errors (not shown) compared with LR CTL. This is, again, consistent with the LR CTL simulations having too large low-level drag, since this component generally reduces when the resolution is increased. It is, of course, plausible that other modeled processes are the cause of these errors, but given the magnitude and location of the errors and their response to resolution, orographic drag is likely to be the first-order process responsible.

## 5. Conclusions

Building on previous WGNE drag model comparisons (Elvidge et al., 2019; Zadra et al., 2013) and on two recent studies which use km-scale simulations to evaluate parametrized orographic drag (van Niekerk et al., 2018; Vosper et al., 2019), the COORDE project further investigates the uncertainties in the representation of orographic drag processes at resolutions typical of seasonal forecasting and climate projections. The multimodel comparison performed in COORDE has revealed the variety of orographic drag parametrizations employed in models used by several operational centers. The low-level drag schemes, which account for nonhydrostatic gravity waves and orographic flow blocking, were found to vary much more in their formulation than those that represent vertically propagating gravity waves. This is related to the fact that tropospheric processes overlap and, therefore, isolating the relevant orographic drag processes that may be

operating within the boundary layer is far more difficult compared with identifying orographic gravity wave drag in the free atmosphere.

Despite the different formulations, the models' drag parametrizations have been tuned, either implicitly or explicitly, to produce reasonable large-scale circulation. This is evident from the fact that the zonal mean parametrized surface stresses were found to be relatively similar at comparable resolutions, although differences still exist over particular regions. The tuning of the orographic drag schemes, guided mostly by the minimization of forecast errors in large-scale circulation metrics of, for example, mean sea level pressure, tropospheric geopotential height, or surface winds, has resulted in drag from different processes (e.g., low-level flow blocking, low-level wave breaking, upper level gravity waves, or turbulent drag) being used interchangeably. Drag from a certain scheme can be erroneously substituted with drag from another scheme while trying to minimize large-scale circulation errors, which may not have a unique solution in terms of optimal uncertain parameters. The COORDE model comparison demonstrates that this substitution can affect local aspects of the circulation, such as the diurnal cycle in near-surface wind speed, due to the fact that the various drag processes depend on stability or winds in different ways and have different regional and vertical drag distributions. For these reasons, it is necessary to go beyond the traditional forecast-skill-based tuning exercise and to provide process-based constraints for the different orographic drag processes, following the avenues suggested in Sandu et al. (2019).

This study makes progress in this direction by using high-resolution simulations to perform a process-based evaluation of the parametrized low-level orographic drag and orographic gravity wave drag at resolutions typically used for seasonal and climate integrations. The methodology used here not only compares parametrized and resolved surface stresses and gravity wave drag in high- and low-resolution simulations, as is typically done (c.f. Vosper et al., 2016, 2019), but also makes use of novel high-resolution simulations with high- and low-resolution orography to estimate the impact of resolved orography on the flow. While this method assumes that the high-resolution simulations are a good proxy for observations, this assumption has been shown to be justified not only here but also in previous studies (e.g., Holt et al., 2017; Stephan et al., 2019). Our approach has helped to demonstrate that parametrized low-level orographic drag is, generally, overestimated at climate resolutions and stratospheric orographic gravity wave drag is underestimated at both climate and seasonal resolutions. These experiments have, thus, provided new constraints on: the vertical and regional distribution of parametrized and resolved orographic gravity wave drag; the magnitude of the orographic surface stresses and gravity wave drag in the upper troposphere to lower stratosphere; the resolution sensitivity of orographic surface stress and gravity wave drag; the diurnal cycle of orographic surface stresses; and orographic drag throughout the troposphere, which is notoriously difficult to diagnose directly from either observations or high-resolution simulations. It is hard to imagine how such information could have been obtained from a typical forecast-skill-based tuning exercise, even through repeated iterations.

Aside from making qualitative statements about the models' parametrization schemes, this study has provided a quantitative comparison to act as a reference for model developers. The magnitude of the parametrized gravity wave drag within the lower stratosphere was found to vary across the models and regions, even at comparable resolutions. The total gravity wave drag (resolved plus parameterized) over the NH was found to be underestimated by at least 30% in most of the models at climate resolutions, with the exception of GSM1705, GSM2003, and ICON, when compared with resolved gravity wave drag in global 9 km simulations. Over the SH, the total gravity wave drag was found to be underestimated by at least 50% and 25% at climate and seasonal resolutions, respectively. This underestimation is consistent with the SH "cold-pole" bias, which refers to climatological stratospheric temperature biases in excess of around 10 K over the SH polar stratosphere (Garcia et al., 2017), that are seen among models. Since the significant gravity wave drag underestimation is present when only land points are included in the resolved and parametrized gravity wave drag profiles (Figure 3), this work shows that a large part of the missing wave drag that contributes to this bias is likely to be from orographic rather than nonorographic gravity waves. Furthermore, the fact that the resolved gravity wave drag increases by ~25% over the ME region when the resolution is increased from 9 to 1.8 km suggests that the underestimation of the total gravity wave drag over the SH and NH in the LR and MR simulations is in fact even larger than what was estimated using the 9 km global simulations. Indeed, over the ME region, the total gravity wave drag in the LR and MR simulations is underestimated by approximately 70% and 40%, respectively, compared with the regional UM 1.8 km simulations.



Diagnosis of the resolved surface stresses over the ME indicate that parameterized low-level orographic drag is significantly overestimated in most of the models at climate resolutions. This is further supported by comparison between the impacts from parametrized orographic drag and resolved orography on the winds, which showed that the parametrized orographic drag leads to larger deceleration of the winds throughout the troposphere compared with resolved orography. The resolved orographic surface stress was also shown to have a different diurnal cycle compared with the SSO stress, indicating that either the SSO dependence on wind/stability may be incorrect or there is some interaction between the parametrized drag processes.

Given that the parametrized surface stress is much larger than that predicted from the high-resolution simulations but the parametrized gravity wave drag in the stratosphere is much smaller than that in high resolution simulations, a rebalancing of the two may be necessary. It is possible that the low-level drag has been increased to account for the lack of gravity wave drag in the upper atmosphere, since the two have a similar impact on the large-scale circulation in the troposphere. For example, both lead to an increase in the pressure over the pole (Palmer et al., 1986; Williams et al., 2020) and a reduction in the speed of the barotropic jet (McFarlane, 1987; Pithan et al., 2016; van Niekerk et al., 2017). While tuning the schemes to produce a reasonable circulation within the troposphere, one may be inadvertently substituted for the other. What is more, the handover between resolved and parametrized orographic drag when resolution is varied is not well handled in most of the models. For example, as was also identified for the UM in Vosper et al. (2019), most of the models exhibited a weak resolution sensitivity in their parametrized gravity wave drag when compared with the resolved gravity wave drag. It is, therefore, also likely that the models have been tuned for a particular resolution, for example, those used in short-range or seasonal forecasting, since the total drag is more accurately estimated at those resolutions. The same orographic drag parametrization can also produce very different resolution sensitivity, total drag and vertical drag partitioning (e.g., the Lott & Miller, 1997 scheme in the IFS and ICON), further highlighting the uncertainty in current orographic drag parametrizations. This work, thus, emphasizes that current orographic drag parametrizations are not ideal for use across model resolutions without further improvements (e.g., through reformulation of the gravity wave drag schemes to account for the full unresolved orographic spectrum, see Vosper et al., 2019).

## Data Availability Statement

The FV3GFS simulations for COORDE are stored at NOAA/HPSS data archive system. Model data used in the production of figures can be found at <https://doi.org/10.5281/zenodo.3763463>.

## Acknowledgments

The authors would like to thank Michel Roch for performing experiments with the Canadian model for this analysis. A. van Niekerk would like to acknowledge discussions at the ISSI Bern with members of the “New Quantitative Constraints on Orographic Gravity Wave Stress and Drag Satisfying Emerging Needs in Seasonal-to-Subseasonal and Climate Prediction.” A. van Niekerk was supported by the Met Office Weather and Climate Science for Service Partnership (WCSSP) India as part of the Newton Fund. The work of V. A. Yudin was supported by NOAA/NWS Award NA18NWS4680050 to University of Colorado (CIRES) with computing resources provided by NOAA/RDHP. V. A. Yudin thanks Svetlana Karol, Jun Wang, and Fanglin Yang for performing and discussing the COORDE simulations in the research version of FV3GFS (127 levels) extended into the mesosphere. The authors would also like to express their gratitude to three anonymous reviewers for their helpful suggestions, which led to a much improved manuscript.

## References

- Alexander, M. J., Geller, M., McLandress, C., Polavarapu, S., Preusse, P., Sassi, F., et al. (2010). Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models. *Quarterly Journal of the Royal Meteorological Society*, *136*(650), 1103–1124. <https://doi.org/10.1002/qj.637>
- Alpert, J. C. (2004). Sub-grid scale mountain blocking at NCEP. 20th Conf. WAF/16 Conf. NWP P2.4.
- Beljaars, A. C. M., Brown, A. R., & Wood, N. (2004). A new parametrization of turbulent orographic form drag. *Quarterly Journal of the Royal Meteorological Society*, *130*(599), 1327–1347. <https://doi.org/10.1256/qj.03.73>
- Berckmans, J., Woollings, T., Demory, M.-E., Vidale, P.-L., & Roberts, M. (2013). Atmospheric blocking in a high resolution climate model: Influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, *14*(1), 34–40. <https://doi.org/10.1002/asl2.412>
- Butler, A., Charlton-Perez, A., Domeisen, D. I. V., Garfinkel, C., Gerber, E. P., Hitchcock, P., et al. (2019). Chapter 11—Sub-seasonal predictability and the stratosphere. In A. W. Robertson & F. Vitart (Eds.), *Sub-seasonal to seasonal prediction* (pp. 223–241). Elsevier. <https://doi.org/10.1016/B978-0-12-811714-9.00011-5>
- Catry, B., Geleyn, J.-F., Bouysse, 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. <https://doi.org/10.1127/0941-2948/2008/0272>
- Choi, H.-J., & Hong, S.-Y. (2015). An updated subgrid orographic parameterization for global atmospheric forecast models. *Journal of Geophysical Research: Atmospheres*, *120*, 12,445–12,457. <https://doi.org/10.1002/2015JD024230>
- Crook, N. A. (1988). Trapping of low-level internal gravity waves. *Journal of the Atmospheric Sciences*, *45*(10), 1533–1541. [https://doi.org/10.1175/1520-0469\(1988\)045<1533:TOLLIG>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<1533:TOLLIG>2.0.CO;2)
- ECMWF (2019). IFS documentation CY46R1.
- Elvidge, A. D., Sandu, I., Wedi, N., Vosper, S. B., Zadra, A., Boussetta, S., et al. (2019). Uncertainty in the representation of orography in weather and climate models and implications for parameterized drag. *Journal of Advances in Modeling Earth Systems*, *11*, 2567–2585. <https://doi.org/10.1029/2019MS001661>
- Fritts, D. C., & Alexander, M. J. (2003). Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics*, *41*(1), 1003. <https://doi.org/10.1029/2001RG000106>
- Garcia, R. R., Smith, A. K., Kinnison, D. E., Cámara, A., & Murphy, D. J. (2017). Modification of the gravity wave parameterization in the whole atmosphere community climate model: Motivation and results. *Journal of the Atmospheric Sciences*, *74*(1), 275–291. <https://doi.org/10.1175/JAS-D-16-0104.1>

- Garratt, J. R. (1994). Review: The atmospheric boundary layer. *Earth-Science Reviews*, 37(1), 89–134. [https://doi.org/10.1016/0012-8252\(94\)90026-4](https://doi.org/10.1016/0012-8252(94)90026-4)
- Giorgetta, M. A., Brokopf, R., Crueger, T., Esch, M., Fiedler, S., Helmert, J., et al. (2018). ICON-A, the atmosphere component of the ICON earth system model: I. Model description. *Journal of Advances in Modeling Earth Systems*, 10, 1613–1637. <https://doi.org/10.1029/2017MS001242>
- Holt, L. A., Alexander, M. J., Coy, L., Liu, C., Molod, A., Putman, W., & Pawson, S. (2017). An evaluation of gravity waves and gravity wave sources in the Southern Hemisphere in a 7 km global climate simulation. *Quarterly Journal of the Royal Meteorological Society*, 143(707), 2481–2495. <https://doi.org/10.1002/qj.3101>
- Hong, S.-Y., Kwon, Y. C., Kim, T.-H., Esther Kim, J.-E., Choi, S.-J., Kwon, I.-H., et al. (2018). The Korean Integrated Model (KIM) system for global weather forecasting. *Asia-Pacific Journal of Atmospheric Sciences*, 54, 267–292. <https://doi.org/10.1007/s13143-018-0028-9>
- Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., et al. (2017). The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, 98(3), 589–602. <https://doi.org/10.1175/BAMS-D-15-00135.1>
- Iwasaki, T., Yamada, S., & Tada, K. (1989). A parameterization scheme of orographic gravity wave drag with two different vertical partitionings. *Journal of the Meteorological Society of Japan. Ser. II*, 67(1), 11–27. [https://doi.org/10.2151/jmsj1965.67.1\\_11](https://doi.org/10.2151/jmsj1965.67.1_11)
- JMA (2019). Outline of the operational numerical weather prediction at the Japan meteorological agency. [https://www.jma.go.jp/jma-eng/jma-center/nwp/outline2019-nwp/pdf/outline2019\\_03.pdf](https://www.jma.go.jp/jma-eng/jma-center/nwp/outline2019-nwp/pdf/outline2019_03.pdf)
- Kanehama, T., Sandu, I., Beljaars, A., van Niekerk, A., & Lott, F. A. (2019). Which orographic scales matter most for medium-range forecast skill in the Northern Hemisphere winter? *Journal of Advances in Modeling Earth Systems*, 11, 3893–3910. <https://doi.org/10.1029/2019MS001894>
- Kim, Y.-J., & Arakawa, A. (1995). Improvement of orographic gravity wave parameterization using a mesoscale gravity wave model. *Journal of the Atmospheric Sciences*, 52(11), 1875–1902. [https://doi.org/10.1175/1520-0469\(1995\)052<1875:IOGWP>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<1875:IOGWP>2.0.CO;2)
- Kim, Y.-J., & Doyle, J. D. (2005). Extension of an orographic-drag parameterization scheme to incorporate orographic anisotropy and flow blocking. *Quarterly Journal of the Royal Meteorological Society*, 131(609), 1893–1921. <https://doi.org/10.1256/qj.04.160>
- Koo, M.-S., Choi, H.-J., & Han, J.-Y. (2018). A parameterization of turbulent-scale and mesoscale orographic drag in a global atmospheric model. *Journal of Geophysical Research: Atmospheres*, 123, 8400–8417. <https://doi.org/10.1029/2017JD028176>
- Kruse, C. G., & Smith, R. B. (2015). Gravity wave diagnostics and characteristics in mesoscale fields. *Journal of the Atmospheric Sciences*, 72(11), 4372–4392. <https://doi.org/10.1175/JAS-D-15-0079.1>
- Lin, S.-J., Putman, W., & Harris, L. (2017). The GFDL finite-volume cubed-sphere dynamical core. <https://www.gfdl.noaa.gov/wp-content/uploads/2020/02/FV3-Technical-Description.pdf>
- Lott, F., & Miller, M. J. (1997). A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quarterly Journal of the Royal Meteorological Society*, 123(537), 101–127. <https://doi.org/10.1002/qj.49712353704>
- Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., et al. (2012). Tuning the climate of a global model. *Journal of Advances in Modeling Earth Systems*, 4, M00A01. <https://doi.org/10.1029/2012MS000154>
- 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. [https://doi.org/10.1175/1520-0469\(1987\)044<1775:TEOEG>2.0.CO;2](https://doi.org/10.1175/1520-0469(1987)044<1775:TEOEG>2.0.CO;2)
- McTaggart-Cowan, R., Vaillancourt, P. A., Zadra, A., Chamberland, S., Charron, M., Corvec, S., et al. (2019). Modernization of atmospheric physics parameterization in Canadian NWP. *Journal of Advances in Modeling Earth Systems*, 11, 3593–3635. <https://doi.org/10.1029/2019MS001781>
- Palmer, T. N., Shutts, G. J., & Swinbank, R. (1986). Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization. *Quarterly Journal of the Royal Meteorological Society*, 112(474), 1001–1039. <https://doi.org/10.1002/qj.49711247406>
- Pierrehumbert, R. T. (1986). An essay on the parameterization of orographic gravity wave drag. <https://www.ecmwf.int/node/11673>
- 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, 7231–7240. <https://doi.org/10.1002/2016GL069551>
- Prill, F., Reinert, D., Rieger, D., Zängl, G., Schröter, J., Förstner, J., et al. (2019). Working with the ICON model practical exercises for NWP mode and ICON-ART.
- Richter, J. H., Sassi, F., & Garcia, R. R. (2010). Toward a physically based gravity wave source parameterization in a general circulation model. *Journal of the Atmospheric Sciences*, 67(1), 136–156. <https://doi.org/10.1175/2009JAS3112.1>
- Roehrig, R., Beau, I., Saint-Martin, D., Alias, A., Decharme, B., Guérémy, J.-F., et al. (2020). The CNRM global atmosphere model ARPEGE-climat 6.3: Description and evaluation. *Journal of Advances in Modeling Earth Systems*, 12, e2020MS002075. <https://doi.org/10.1029/2020MS002075>
- Sandu, I., Bechtold, P., Beljaars, A., Bozzo, A., Pithan, F., Shepherd, T. G., & Zadra, A. (2016). Impacts of parameterized orographic drag on the Northern Hemisphere winter circulation. *Journal of Advances in Modeling Earth Systems*, 8, 196–211. <https://doi.org/10.1002/2015MS000564>
- Sandu, I., van Niekerk, A., Shepherd, T. G., Vosper, S. B., Zadra, A., Bacmeister, J., et al. (2019). Impacts of orography on large-scale atmospheric circulation. *Nature Partners Journal Climate and Atmospheric Science*, 2, 10. <https://doi.org/10.1038/s41612-019-0065-9>
- Shepherd, T. G. (2014). Atmospheric circulation as a source of uncertainty in climate change projections. *Nature Geoscience*, 7, 703–708.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., et al. (2008). A description of the advanced research WRF version 3 (no. NCAR/TN-475+STR). <https://doi.org/10.5065/D68S4MVH>
- Smith, S. A., Doyle, J. D., Brown, A., & Webster, S. (2005). Sensitivity of resolved mountain drag to model resolution for map case studies. *Hrvatski Meteorološki Časopis*, 132(40), 264–267. <https://doi.org/10.1256/qj.05.67>
- Stephan, C. C., Strube, C., Klocke, D., Ern, M., Hoffmann, L., Preusse, P., & Schmidt, H. (2019). Intercomparison of gravity waves in global convection-permitting models. *Journal of the Atmospheric Sciences*, 76(9), 2739–2759. <https://doi.org/10.1175/JAS-D-19-0040.1>
- van Niekerk, A., Sandu, I., & Vosper, S. B. (2018). The circulation response to resolved versus parameterized orographic drag over complex mountain terrains. *Journal of Advances in Modeling Earth Systems*, 10, 2527–2547. <https://doi.org/10.1029/2018MS001417>
- 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. <https://doi.org/10.1175/JAS-D-17-0085.1>
- van Niekerk, A., Shepherd, T. G., Vosper, S. B., & Webster, S. (2016). Sensitivity of resolved and parameterized surface drag to changes in resolution and parametrization. *Quarterly Journal of the Royal Meteorological Society*, 142(699), 2300–2313. <https://doi.org/10.1002/qj.2821>

- Vannière, B., Demory, M.-E., Vidale, P. L., Schiemann, R., Roberts, M. J., Roberts, C. D., et al. (2019). Multi-model evaluation of the sensitivity of the global energy budget and hydrological cycle to resolution. *Climate Dynamics*, *52*(11), 6817–6846.
- Vosper, S. B. (2015). Mountain waves and wakes generated by south Georgia: Implications for drag parametrization. *Quarterly Journal of the Royal Meteorological Society*, *141*(692), 2813–2827. <https://doi.org/10.1002/qj.2566>
- 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*, 3128–3137. <https://doi.org/10.1002/qj.2894>
- Vosper, S. B., van Niekerk, A., Elvidge, A., Sandu, I., & Beljaars, A. (2019). What can we learn about orographic drag parametrization from high-resolution models? A case study over the rocky mountains. *Quarterly Journal of the Royal Meteorological Society*, 979–995. <https://doi.org/10.1002/qj.3720>
- 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. <https://doi.org/10.1002/qj.49710946202>
- Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., et al. (2019). The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations. *Geoscientific Model Development*, *12*(5), 1909–1963. <https://doi.org/10.5194/gmd-12-1909-2019>
- Webster, S., Wells, H., Vosper, S., & Smith, S. (2020). Unified model documentation paper 022: Gravity wave drag. [https://code.metoffice.gov.uk/doc/um/latest/papers/umdp\\_022.pdf](https://code.metoffice.gov.uk/doc/um/latest/papers/umdp_022.pdf)
- Williams, K. D., van Niekerk, A., Best, M. J., Lock, A. P., Brooke, J. K., Carvalho, M. J., et al. (2020). Addressing the causes of large-scale circulation errors in the Met Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, *146*, 2597–2613.
- Wood, N., Brown, A. R., & Hewer, F. E. (2001). Parametrizing the effects of orography on the boundary layer: An alternative to effective roughness lengths. *Quarterly Journal of the Royal Meteorological Society*, *127*(573), 759–777. <https://doi.org/10.1002/qj.49712757303>
- Wood, N., & Mason, P. (1993). The pressure force induced by neutral, turbulent flow over hills. *Quarterly Journal of the Royal Meteorological Society*, *119*(514), 1233–1267. <https://doi.org/10.1002/qj.49711951402>
- Yonehara, H., Matsukawa, C., Nabetani, T., Kanehama, T., Tokuhira, T., Yamada, K., et al. (2020). Upgrade of JMA's operational global model. *WGNE Research Activities in Earth System Modelling*, *6*, 6–19.
- Zadra, A., Antonopoulos, S., Archambault, B., Beaulne, A., Bois, N., Buehner, M., et al. (2015). Improvements to the Global Deterministic Prediction System (GDPS) (from version 2.2.2 to 3.0.0), and related changes to the Regional Deterministic Prediction System (RDPS) (from version 3.0.0 to 3.1.0). [https://collaboration.cmc.ec.gc.ca/cmc/cmoi/product\\_guide/docs/lib/op\\_systems/doc\\_opchanges/tech-note\\_gdps300\\_20130213\\_e.pdf](https://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/lib/op_systems/doc_opchanges/tech-note_gdps300_20130213_e.pdf)
- Zadra, A., Bacmeister, J., Bouysse, F., Brown, A., Lock, A., Figueroa, S. N., et al. (2013). WGNE drag project. [https://collaboration.cmc.ec.gc.ca/science/rpn/drag\\_project/](https://collaboration.cmc.ec.gc.ca/science/rpn/drag_project/)
- Zadra, A., Roch, M., Laroche, S. A., & Charron, M. (2003). The subgrid-scale orographic blocking parametrization of the GEM model. *Atmosphere-Ocean*, *41*(2), 155–170. <https://doi.org/10.3137/ao.410204>