Observed and Modeled Mountain Waves from the Surface to the Mesosphere Near the Drake Passage

[1]Christopher G. Kruse *, [1]M. Joan. Alexander, [2]Lars Hoffmann, [3]Annelize van Niekerk, [4]Inna Polichtchouk, [5]Julio T. Bacmeister, [1]Laura Holt, [6]Riwal Plougonven, [7]Petr Šácha, [8]Corwin Wright, [9]Kaoru Sato, [10]Ryosuke Shibuya, [11]Sonja Gisinger, [2]Manfred Ern, [2]Catrin Meyer, [2]Olaf Stein

[1]*NorthWest Research Associates, Boulder, Colorado, USA*

[2]*Forschungszentrum Jülich, Jülich Supercomputing Centre, Jülich, Germany*

[3]*Met Office, Exeter, UK*

[4]*ECMWF, Reading, UK*

[5]*Climate and Global Dynamics Laboratory, NCAR, Boulder, Colorado, USA*

[6]*Laboratoire de Météorologie Dynamique, Ecole Polytechnique, Palaiseau, France*

[7]*Earth Sciences, Department of Applied Physics, Universidade de Vigo, Ourense, Spain*

[8]*Centre for Space, Atmospheric and Oceanic Science, University of Bath, Bath, UK*

[9]*Department of Earth and Planetary Science, University of Tokyo, Tokyo, Japan*

[10]*Atmosphere and Ocean Research Institute, University of Tokyo, Tokyo, Japan*

[11]*Institute of Atmospheric Physics, Deutsches Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen, Germany*

Corresponding author address: NorthWest Research Associates, Boulder, Colorado

Four state-of-the-science numerical weather prediction (NWP) models were used to perform mountain wave- (MW) resolving hindcasts over the Drake Passage of a 10-day period with numerous observed MW cases in 2010. The Integrated Forecast System (IFS) and the Icosahedral Nonhydrostatic (ICON) were run in their current operational configurations at $\Delta x \approx 9$ and 13 km, globally. The Weather Research and Forecasting (WRF) model and the Met Office Unified Model (UM) were configured with a $\Delta x = 3$ km regional domain. All models were were deep, with domain tops near 1 Pa ($z \approx 80$ km). These deep domains allowed *quantitative* validation against Atmospheric InfRared Sounder (AIRS) observations, accounting for observation time, viewing geometry, and radiative transfer.

All models reproduced observed middle-atmosphere MWs with remarkable skill. Increased resolution improved validations. Still, all models underrepresented observed MW amplitudes, suggesting even at $\Delta x \approx 3$ km resolution, small-scale MWs are under-resolved and/or over-diffused. MW drag parameterizations are still necessary in NWP models at current operational resolutions of $\Delta x \approx 10$ km. Upper GW sponge layers in the operationally configured models signifiantly, artificially reduced MW amplitudes in the upper stratosphere and mesosphere. In the IFS, parameterized GW drags compensated for this effect, but still, total drags were ≈ 6 time smaller than that resolved at $\Delta x \approx 3$ km. Meridionally propagating MWs significantly enhance zonal drag over the Drake Passage. Interestingly, drag associated with lateral fluxes of zonal momentum were important; not accounting for these terms produces a drag at and below the polar night jet maxmimum with the wrong sign.

13 km is one possible grid configuration, but "operationally" a much finer resolution over Germany is being used. Maybe remove "operationally"?

InfraRed (?)

(?)

1. Introduction

Orographic gravity waves (GWs), or mountain waves (MWs), are internal gravity waves within the atmosphere forced by stratified flow over topography. MW generation is one way the atmosphere exchanges momentum with the Earth. Positive and negative pressure perturbations upstream and downstream of a mountain, respectively, exert a net force by the atmosphere on the mountain. An equal and opposite force is exerted by the mountain on the atmosphere. Often, a mountain wave is generated, which propagates the force by the mountain on the atmosphere upward. This force is ultimately exerted wherever the MW breaks, which can occur from the troposphere to the lower portion of the thermosphere (e.g. Fritts et al. 2016).

MWs have a variety of scales, ranging from $\lambda_h \approx 10$ km to hundreds of kilometers, depending on the terrain scales that force them. In the recent DEEPWAVE field campaign, aircraft-observed MWs with scales larger and smaller than 50 km were found to be equally important, fluxing the similar amounts of momentum (Smith and Kruse 2017). Given the effective resolutions of current numerical weather prediction (≈ 60 km) and climate (≈ 400 km) models, an important portion of the MW spectrum is under- or unresolved and therefore needs to be parameterized. GW drag parameterizations were first formulated in the early 1980s (e.g. Lindzen 1981; Holton 1983), and the addition of MW drag (MWD) parameterizations did improve weather and climate models (e.g. Palmer et al. 1986; McFarlane 1987; Miller et al. 1989) and were widely adopted thereafter. The momentum deposited by MW breaking is important at all levels of the atmosphere, though, its importance with respect to the atmosphere's zonally-averaged momentum budget generally increases with altitude, at least according to MWD parameterizations (e.g. Sato et al. 2018; Yasui et al. 2018). For an overview of the ≈100-year history of MW research, see Smith (2019).

Despite the rich literature and knowledge gained over the last century, current MWD parameterizations, and GW drag parameterizations in general, are still highly simplified and not well-constrained. Most current MWD parameterizations estimate one or two 2-D source MWs based on sub-grid-scale (SGS) terrain statistics. Then, these 2-D MWs are assumed to propagate strictly upward instantaneously, as treated by steady 2-D linear theory. The increase in amplitude with decreasing density is taken into account. When the 2-D parameterized wave reaches a large enough amplitude that some instability mechanism (e.g. overturning) is predicted to occur, the parameterized wave amplitude is not allowed to grow in amplitude further (e.g. Lindzen 1981) and is said to be "saturated." The reduction in amplitude below that predicted by linear theory results in a reduction of the momentum flux with height, from which the force on the flow is calculated.

Many aspects of current MWD parameterizations are inconsistent with current understanding of MW dynamics and observations. For example, terrain is 3-D and forces 3-D MWs. 3-D MWs can and do propagate laterally, spreading out with height (e.g. Eckermann et al. 2015). Lateral propagation spreads out the wave activity, reduces amplitudes, increases breaking heights, and also spreads out the drag relative to current parameterizations, which propagate MWs only vertically. Temporal evolution of the ambient wind and MWs can be important (e.g. Chen et al. 2005; Kruse and Smith 2018) and are also not represented in current parameterizations. Many more deficiencies could be listed.

Additionally, MWD parameterizations are not well constrained by observations. Numerous datasets collected on numerous platforms do contain MW signals; however, these data are often infrequent, are sparce, or incompletely sample the MW field (e.g. data from field campaigns, radiosondes, commercial aircraft). These limitations prevent use of these data to constrain the inevitable tuning parameters within MWD parameterizations used by global weather prediction and climate models. Noteworthy possible exceptions are data sets collected by satellite-borne nadir

(x) Wright et al., ACP, 2017 discuss lateral propagation / wave gravity into the stratospheric jet

we might want to include reference to Corwin's and Neil Hindley's paper on 3-D S-Transform analysis of AIRS data (ask Corwin for reference)

and limb infrared radiometers, which can be sensitive to the temperature perturbations caused by MWs. Recent efforts (e.g. Alexander et al. 2009; Ern et al. 2017) have developed methods of using these satellite data to estimate the momentum flux vectors at stratospheric altitudes, which potentially have the spatial and temporal sampling needed to constrain MWD parameterizations. However, these sensors have limitations as well, being only able to sense a portion of the horizontal and vertical spectrum of scales.(x) The inconsistencies of MWD parameterizations with current physical understanding and their lack of observational constraints motivate additional research.

The overall objective of this work is to provide high-quality estimates of mountain wave momentum fluxes and drags from the troposphere to the mesosphere that are anchored to reality. This objective is met via a model intercomparison involving four state-of-the-science NWP models that are validated against stratospheric AIRS observations. With such a collection of NWP model output, a second route to reality is to evaluate methods of computing MW momentum fluxes from satellite data via method intercomparison and an observing system simulation experiment (OSSE). For example, momentum fluxes from synthetic AIRS temperature perturbation "data," computed as if the AIRS sensor were viewing the modeled atmosphere, can be compared to the true momentum fluxes within the model computed from wind component perturbations. Figure 1 provides an overview of the two approaches to reality. The model validation and intercomparison is presented here. The method intercomparison and OSSE are left for a future paper.

The NWP models used and the case studied are described in Section 2. The models are validated by comparing synthetic AIRS "data", computed as if the AIRS senser were viewing the modeled atmospheres, with actual AIRS data in Section 3. MW diagnostics are compared between the participating NWP models in Section 4. In Section 5, a previous version and a current MWD parameterization within the Community Atmosphere Model (CAM) are compared with each other and the validated models. The enhancement of zonal drag over the Drake Passage by meridionally

(x) Catrin Meyer et al., AMT, 2018 might be a nice reference for this...

propagating MWs from the north and south is presented in Section 6. Finally, a summary is provided in Section 6.

2. Case Studied, Participating Models and Their Configurations

a. Region and Period of Interest

The region of focus is centered on the Drake Passage, including the mountains of the Southern Andes, Antarctic Peninsula, and a handful of islands in between and to the east (e.g. South Georgia). The time period of interest is 9 Oct 2010 through 20 Oct 2010, where many significant MW events were observed by the Atmospheric Infrared Sounder (AIRS), as shown in Fig. 2.

This region is of interest for a number of reasons. The southern Andes mountains are one of the most significant generators of deeply propagating MWs in the world, and certainly the most significant such region in the southern hemisphere (e.g. Hoffmann et al. 2013). The Drake Passage is a region where lateral propagation of these deeply propagating waves into the stratospheric polar night jet (PNJ) is significant (e.g. Jiang et al. 2013; Hendricks et al. 2014; Amemiya and Sato 2016; Jiang et al. 2019). In a broader context, this region is of interest as many climate models contain the so called "cold-pole problem" (e.g. Eyring et al. 2010), where lower-stratospheric polar temperatures are too cold in the southern hemisphere winter and spring, the meridional temperature gradient is too strong, and (via thermal wind balance) the stratospheric PNJ is also too strong, all of which have detrimental effects on the simulated climate. McLandress et al. (2012) hypothesized this problem might be caused by missing GW drag in the stratosphere around 60^{o}S. The sources of these missing GWs and drag in climate models, and whether or not they are responsible for the cold-pole problem, are still debated. Possible sources include missing MWs launched from mountainous islands in this latitude belt (e.g. Alexander and Grimsdell 2013), too

little non-orographic GW activity associated with tropospheric jetstream and frontal imbalances (e.g. Jewtoukoff et al. 2015), and lateral propagation of MWs into the PNJ (e.g. Amemiya and Sato 2016).

↳ Initially we selected the region + time period for overlap with the Concordiasi campaign. We donot use the Concordiasi balloon data here, but might still be worthwhile mentioning?

b. Description of AIRS Observations

Infrared radiances measured by the Atmospheric Infrared Sounder (AIRS) (Aumann et al. 2003; Chahine et al. 2006) on the Aqua satellite are used in this study to evaluate the realism of the stratospheric gravity wave patterns in the simulations. Aqua was launched in May 2002 into a nearly polar, sun-synchronous, low earth orbit at 705-km altitude 100^o inclination and with a 100 min orbital period. Equator crossings occur at 01:30 (descending) and 13:30 (ascending) local times. By scanning across-track, AIRS observes 1780-km wide image swaths with 90 across-track footprints and resolution varying from 14 x 14 km^2 at nadir to 21 x 42 km^2 at the swath edge. Adjacent scans are separated by 18-km along-track distance. For observing stratospheric gravity waves, we average channels over three sets of infrared frequencies sensing the dry stratosphere where clouds, surface emissions, and reflected sunlight do not affect the radiances. Weighting functions for these channel sets (gray lines) and channel averages (colored lines) are shown in Fig. 3. Two of the channel sets lie in the 15 μm band and one in the 4.3 μm band (Alexander and Barnet 2007; Hoffmann et al. 2013, 2017). The vertical width of these weighting functions limit AIRS sensitivity to gravity waves with vertical wavelengths longer than ↳ about 15 km, with sensitivity increasing for longer vertical wavelengths. Radiances in each channel set are averaged to obtain low-noise brightness temperature data products. Noise varies with scene temperature and channel frequency, but channel-averaged noise varies from 0.1-0.3 K for October at these high southern latitudes (Hoffmann et al. 2014).

maybe repeat (Hoffmann et al., 2017) as a reference for 6-th order poly fit.

Gravity wave anomalies are derived as the residual from a cross-track polynomial fit to each scan. Removal of limb-brightening effects is achieved by subtraction of a 4th order polynomial fit to the higher altitude "15 μm (high)" and 4.3 μm sets, and a 6th order fit to the lower altitude "15 μm (low)" set. Gravity wave brightness temperature anomalies from AIRS at 4.3 μm is shown in Fig. 2 for all relevant swath segments during the October 9-20, 2010 focus period. Corresponding swath segments for both 15 μm sets at the same locations are also used in the model validation presented in Section 3.

c. Models and Configurations

Four state-of-the-science NWP models were configured to attempt to reproduce ~~the~~ the MWs observed by AIRS over the period of interest (Fig. 2). Two regional models (the Weather Research and Forecasting (WRF) model and the UK Met Office's Unified Model (UM)) and two global models (ECMWF's Integrated Forecast System (IFS) and the German Weather Service's Icosohedral Nonhydrostatic (ICON) model) were used to reproduce the observed waves. Details of individual models' configurations are briefly given below. All model initial conditions and boundary conditions were provided by the operational IFS analyses run at $\Delta x \approx 16$ km during this 2010 period.

1) WRF CONFIGURATION

A single $\Delta x \approx 3$ km domain was configured over the region of interest. The domain was deep, with 180 vertical levels extending up to 1 Pa ($z \approx 80$ km). The vertical resolution was varied with height, with the highest resolution near the surface and gradually reduced to a maximum of $\Delta z = 600$ m. Model terrain was derived from the 30 arc-second Global Multi-Resolution Terrain Elevation Data (GMTED) digital elevation model (DEM) (Danielson and Gesch 2011), except for

over Antarctica, where the ≈ 200 m resolution Radarsat Antarctic Mapping Project, Version 2 (RAMP2) DEM was used (Liu et al. 2015). The model was initialized once at 12 UTC on 8 Oct 2010 and then integrated for 11 days. During the integration, only boundary conditions from the 6-hourly operational IFS analyses guided the simulated atmosphere. No interior nudging was used within the 3900 km by 3900 km domain. The WRF model (version 4.1.0) required modification to stably integrate beyond a handful of time steps with a domain top this high. For a brief description of these modifications, see Appendix A.

→ maybe add a sentence about the sponge layer at the model top?

2) UM CONFIGURATION

A global UM simulation, with a grid-spacing of $\Delta x \approx 16$ km and with 85 vertical levels extending to $z = 85$ km initialized every 24 hours from IFS analyses, was used to provide the boundary conditions to a 3000 km by 3000 km regional simulation with a grid-length of $\Delta x \approx 3$ km over the Drake Passage region. Both the global and regional domains were run with UM version 11.1 and used the Global Atmosphere 6 (Walters et al. 2017) and the Regional Atmosphere 1 (Bush et al. 2020) physics configurations, respectively. The $\Delta x \approx 3$ km UM nest was initialized once at 12 UTC on 8 August 2010 and run continuously for 12 days, guided by hourly boundary conditions from the periodically reinitialized global run. The regional UM used a rotated latitude longitude grid, such that the grid-spacing is approximately equal throughout the domain. The vertical resolution was significantly increased within the stratosphere compared with the global simulations, employing 242 vertical levels up to 85 km.

Similar to the regional WRF simulations, the regional model orography is generated from a blend between the 30 arc-second GMTED DEM and a 1 km version of the RAMP2 DEM was used. No filtering was performed on the mean orography and a weak implicit Rayleigh damping of the vertical winds (similar to Klemp et al. 2008) is applied from $z \approx 60$ km upward. The

4. Model Intercomparison

a. Intercomparison of MW Forcing

In the previous section, the models were primarily compared to AIRS observations. Here, the simulations from the different models are compared to one another within the three black-boxed regions in Fig. 7. Prior to all analyses here, fields from each model were regridded onto the WRF grid. Perturbation fields were then computed via the spectral filtering method of Kruse and Smith (2015) on the same WRF grid (except for Figs. 7 and 15, where fields were high-pass filtered on the native UM grid to make use of its larger domain), retaining scales smaller than 500 km. Then, fields were further regridded onto regular latitude, longitude grids within each sub-domain, with grids being $\Delta x \approx 3$ km at the northern boundaries. The Southern Andes, Antarctic Peninsula, and South Georgia regions are denoted by "SA", "AP", and "SG". The 10-day average absolute temperature perturbations at $z = 40$ km in the UM simulation are color shaded in Fig. 7, overviewing the regional MW hot spots and lateral propagation of these MWs in the upper stratosphere.

Terrain spectra and slope distributions within each model and sub-domain are compared in Fig. 8. In all sub-domains, all model terrain spectra agree very well at scales larger than 100 km ($k < 10^{-2}$). All models largely follow the expected -2 slope. The notable drop offs of spectra from this slope, and from spectra of the higher-resolution models, give indication of the effective terrain resolution within each model. The ICON model, which has the coarsest grid resolution ($\Delta x \approx 13$ km), has the coarsest effective terrain resolution as well. The IFS has a slightly higher effective terrain resolution, most easily diagnosed by the precipitous drop off its terrain spectra. The local maxima at smaller scales in the IFS spectra is spurious. The terrain is represented spectrally, with scales smaller than $4\Delta x \approx 36$ km removed. However, IFS terrain were provided on a

TCo2559 grid. The spectral synthesis onto the higher-resolution grid introduces the small-scale noise in the terrain field. The WRF and UM models have the same grid resolution ($\Delta x \approx 3$ km). However, they clearly a have different effective terrain resolutions, with WRF smoothing the the grid-mean orography while UM did not. The slope distributions are largely consistent with the spectra.

The subdomain-averaged 10-m zonal wind time series in each simulation are shown in Fig. 9, along with the subdomain-averaged 10-m winds in the IFS analyses used to for initial and/or boundary conditions in all models. These time series of winds that largely force the MWs seen aloft show that despite the different ways each model incorporated IFS analyses (see Section 2c for details) and domain sizes, the models largely agree on these low-level winds and do not drift too far from the IFS analyses, at least over the SA and AP domains. Larger differences are notable in the smaller, more remote SG domain to the east.

b. Overview of Period of Interest

An overview of the 10-day period of interest is provided in Fig. 10, where subdomain-average zonal and meridional winds, vertical fluxes of zonal momentum, and zonal MWD are shown for each subdomain. Here, the vertical fluxes of horizontal momentum are defined by

$$\begin{aligned} MF_{zx} &= \overline{\rho}\,\overline{u'w'}, \\ MF_{zy} &= \overline{\rho}\,\overline{v'w'}, \end{aligned} \tag{1}$$

and MWDs as

$$\begin{aligned} MWD_{zx} &= -\tfrac{1}{\overline{\rho}}\tfrac{\partial MF_{zx}}{\partial z}, \\ MWD_{zy} &= -\tfrac{1}{\overline{\rho}}\tfrac{\partial MF_{zy}}{\partial z}. \end{aligned} \tag{2}$$

GW-absorbing upper sponges, beginning at $z \approx 40$ km (i.e. 1 hPa) and $z = 44$ km in the IFS and ICON runs, respectively.

The salient feature in the meridional drag profiles is the the reversal of meridional drag, with southward drag occuring at $z \approx 45$ km, with northward drag further aloft (Figs. 12-13). This may be a result of the strong terrain anisotropy of South Georgia Island, where the predominantly westerly low-level winds impinging on the northwest-to-southeast-oriented terrain would tend to launch southward-propagating MWs with larger amplitudes than those propagating northward. This asymmetry would result in southward drag being deposited at lower altitudes.

The x-y-time-averaged profiles in Fig. 13 allow more quantitative comparison between the resolved momentum fluxes and drags over the period of interest. Again, resolved fluxes and drags in the $\Delta x \approx 3$ km ($\Delta x \approx (10 km)$) models. Overall, fluxes and drags in the $O(10 km)$ NWP models are significantly less than the $\Delta x \approx 3$ km limited-area domains. For example, lower-stratospheric resolved zonal momentum fluxes (MF_{zx}) in the NWP models are ≈ 2 times smaller than higher-resolution limited area models (LAMs) over the SA and AP subdomains, and ≈ 3 times smaller in the SG subdomain. Significant disagreement is apparent in the net meridional momentum fluxes, where different resolutions produce different signs of net flux in the troposphere and lower stratosphere!

Evaluating the effect of resolution on drag is more difficult here, due to the strong GW-damping upper sponges applied in the global models. Despite significantly smaller fluxes into the stratosphere in the global models, the resolved drags are more comparable to the limited-area models at and below $z \approx 45$ km. It is unclear if the drags in the upper stratosphere are due to model dynamics and physics or sponge damping. Further aloft, the upper sponges in the global models have clearly reduced resolved drags, particularly in the IFS output. In the IFS, the orographic and non-orographic GWD parameterizations were active and do compensate for this lack of resolved

drag, but still, the total (resolved + parameterized GWD) in the IFS is ≈6 times smaller than the LAMs in the mesosphere. Sponge or no sponge, MWD parameterizations are still necessary even at the current operational NWP resolutions of $\Delta x \approx 10$ km.

↳ for SG, it might be good to refer to some of more recent modelling/observational papers here (or in the introduction): Hindley et al., ACP, 2021; Jackson et al, BAMS, 2018 (and papers of Simon Vosper?)

d. Model Intercomparison of MW Spectra

The analyses presented in the previous subsection show the area net momentum fluxes across the entire MW spectrum. However, broad-spectrum orography forces a broad-spectrum of MWs. The momentum flux and MWD co-spectra are presented in Fig. 14 at selected altitudes within each subdomain. These spectra were computed by using 2-D FFTs on the subdomain lat/lon grids to compute 2-D MF_{zx} cospectra. These 2-D cospectra were then binned by wavenumber magnitude into 1-D cospectra. These 1-D MF cospectra are plotted such that the area-integral in the spectral space plotted is proportional to the area-averaged fluxes in physical space (e.g. $\frac{1}{A}\int_{-\infty}^{\infty} \frac{\partial MF_{zx}}{\partial \ln\lambda}\, \partial \ln\lambda = \frac{1}{A}\iint \overline{\rho}\, u'w'\, dxdy$). The drag co-spectra are computed by taking the vertical derivative of the MF_{zx} co-spectra at ±5 km from the altitudes indicated in the panels. These panels show what scales are attenuated and responsible for depositing drag on 10-km depths of atmosphere at the selected levels.

The comparison of momentum flux and drag spectra in Fig. 14 highlights the breadth of the flux-carrying MW spectrum and how much of this spectrum is not resolved even at the relatively high ≈ 10-km resolution global models. The smaller-scales resolved in the LAMs contribute significantly to the total fluxes over all altitudes and sub-regions shown. Generally, fluxes in the coarser NWP models are smaller than in the LAMs. Curiously, fluxes at larger scales are occasionally larger in the coarser NWP models, compensating for the lack of flux at smaller scales somewhat.

The previous parameterization is relatively simple, and based on the McFarlane (1987) scheme (for details, see Neale et al. 2010). This parameterization is uses an isotropic source formulation, specifying the source MWs and MF to be parallel and opposite to the source-level flow. Source wave amplitudes are specified as twice the standard deviation of sub-grid-scale (SGS) orography, unless this wave amplitudes is predicted to overturn, in which case the source amplitude is reduced to the saturation amplitude.

The new parameterization, currently implemented in CAM6, varies significantly in how the SGS terrain is represented and how the source MWs are specified. It fits 2-D ridges to the SGS terrain, finding dominant ridges, along with their widths, lengths, heights, and orientations, in the SGS terrain data. This method allows more realistic, larger ridge heights to be used in estimating source MW amplitudes and allows terrain orientation and anisotropy to be taken into account. Further details are provided in Appendix B.

Parameterized momentum fluxes and drags are quantitatively compared against those resolved in the four high-resolution models presented here in Figs. 11-13. Both the zonal and meridional components of momentum flux and drag are improved, though not perfectly, in the new parameterization. The zonal momentum fluxes and drags are much lower in the old parameterization (cf. Figs. 11 i, k and j, l, respectively). The enhancement in MF_{zx} by the new parameterization is likely due to larger source amplitudes being specified and produce a better comparison to WRF and UM. More striking differences, and improvements, are seen in the meridional components (Fig. 12). The previous isotropic parameterization produces nearly zero meridional fluxes and drags, as the source-level winds are primarily westerly. However, the new parameterization accounts for the strong northwest-southeast terrain orientation, producing southward momentum fluxes and drags aloft. While these meridional drags compare much better with WRF and UM, the southward drag does appear a little low and the northward drags noted further aloft in WRF and UM are non-

existent. Overall, the new parameterization is a signficant physical and quantitative improvement over the previous one, there is still much room for improvement.

6. Meridional Propagation of MWs and Drag

A primary motivation for studying the Drake Passage region was the previous work pointing to missing stratospheric GWD near 60^{o}S (McLandress et al. 2012) and the density of stratospheric MW hotspots (e.g. Hoffmann et al. 2013) in this region. The source of this missing drag is not settled, with good hypotheses including under, or no, representation of MWs and drag by small islands, non-orographic GWs, and lateral (i.e. meridional) propagation of MWs into these latitudes of interest.

While the orographic sources in the Drake Passage region are not at 60^{o}S, MWs laterally-launched by 3-D terrain and meridional refraction of MWs by the deep meridional shear of zonal wind on the flanks of the polar night jet (e.g. Jiang et al. 2013, 2019) both contribute to the presence of GWs over and east of the Drake Passage (e.g. Fig. 7). Still, it is unclear how much zonal momentum is fluxed by these meridionally-propagating waves and whether the resulting zonal drags are significant.

Here, the 3-D fluxes of zonal momentum and the influences of the time- and zonal-mean zonal wind were computed within the UM model. These quantities were computed within the two red boxes in Fig. 7, one upstream and one over and downstream of the Drake Passage. The 3-D zonal MF vector,

$$\mathbf{MF}_x = \langle MF_{xx} = \overline{\rho}\,\overline{u'^2}, MF_{yx} = \overline{\rho}\,\overline{u'v'}, MF_{zx} = \overline{\rho}\,\overline{u'w'} \rangle, \quad (3)$$

the influences on the zonal-mean zonal winds by each flux component,

Fig. 15c) more than compensates the acceleration at and below the PNJ max, resulting a net drag on the PNJ (MWD_{tot}, Fig. 15d).

To summarize, meridional propagation of MWs does significantly increase zonal momentum fluxes and drag by a factor of 2-4 over the Drake Passage. The meridional divergence of meridional flux of zonal momentum term was actually important in the total impact of the small-scales on the large-scale zonal flow at and below the PNJ max. In fact, if only the vertical divergence of vertical flux of zonal momentum were considered, as is conventionally done, the sign of the drag on the PNJ max would be incorrect! These results motivate further research and parameterization development considering lateral propagation of MWs and suggest that lateral momentum fluxes and resulting drags may need to be considered.

7. Summary and Conclusions

In this article, four state-of-the-science NWP models were used to recreate the atmosphere around the Drake Passage during the 10-20 October 2010 period, where numerous MW events were observed by AIRS within the period. The ICON and IFS models were run at ≈ 13 km and ≈ 9 km resolution in their operational configurations. The WRF and UM models were run in a regional configuration at $\Delta x = 3$ km resolution. All models were run with a model top near $z = 80$ km in order to contain the entire life-cycle of the MWs (i.e. from generation to dissipation) and to allow quantitative comparison with AIRS observations. These deep domains allowed most of the AIRS instrument weighting functions for the selected channels to be entirely contained within the domains and be largely unaffected by the necessary upper sponge layers, at least in the WRF and UM domains. Viewing geometries, observation times, and radiative transfers for the AIRS channels/wavelengths used were all taken into account...to make this comparison / evaluation as accurate/exact as possible.

The model validation showed that all models had, overall, remarkable skill at reproducing the AIRS-observed MWs within the middle atmosphere. This was remarkable, given the numerous differences between the models (e.g. different dynamical cores, parameterizations, grids, initializations, boundary condition methods) and the deep vertical distances the MWs had to propagate in order to be observed by AIRS. Still, the validations were not perfect. All models appeared to have the most skill reproducing MWs from the Andes, which had the largest scale, and best resolved, orography. At the other end of the orographic spectrum, the largest differences between the models and between models and observations occured for MWs generated by South Georgia Island. Here, it was clear the coarser resolution global NWP models struggled. Despite remarkable qualitative validations, all models under-represented observed MW amplitudes, even after accounting for effective model resolutions and instrument noise, suggesting even at $\Delta x \approx 3$ km, models still under-resolve and/or over-diffuse MWs.

Resolved momentum fluxes and drags, averaged over the Southern Andes, Antarctic Peninsula, and South Georgia subdomains, were also intercompared between the four models in detail. The IFS and ICON models, run globally at current operational resolutions of $\Delta x \approx 10$ km, generally had significantly less momentum fluxes and drags than the $\Delta x = 3$ km resolution WRF and UM output. The upper sponges in both the IFS and ICON models had clearly significantly and artificially reduced fluxes and drags in the upper stratosphere and mesosphere. While the orographic and non-orographic GWD parameterizations compensated for this reduction somewhat, the total drag was still ≈ 6 times smaller than that resolved by the $\Delta x = 3$ km models.

The intercomparison of time-averaged zonal momentum flux and drag cospectra highlighted the breadth of the flux-carrying MW spectrum and how much of the spectrum these ≈ 10-km global models are missing relative to the ≈ 3-km regional models. These model spectra show the continued need for MW drag parameterizations at current operational NWP resolutions. The

systematic differences in MW momentum fluxes at small, grid-scale resolutions between WRF and UM, presumably due to different numerical diffusions, also suggest the entire flux-carrying MW spectrum is not yet resolved even at $\approx$ 3-km resolution. These results motivate spectral approaches for MW drag parameterizations (e.g. van Niekerk and Vosper 2021).

A significant motivation for this research was to evaluate and ultimately improve MW drag parameterizations. The previous and current MW drag parameterizations in CAM were compared with each other and against the MW-resolving models via CAM runs continuously nudged to the ERA-Interim reanalysis. The current parameterization, which better represents SGS terrain heights and accounts for terrain anisotropy, does seem to be a significant improvement over the previous parameterization, particularly for components of drag perpendicular to the forcing flow. Still, there is clearly lots of room for improvement; there is too much parameterized drag at too low an altitude and too little further aloft. An interesting feature in the 3-km models was the reversal of meridional MW drag above South Georgia Island. The parameterization did have a representation of the expected southward MW drag, but did not represent resolved drag reversal. If the resolved drag reversal is due to southward-propagating MWs being larger and breaking lower than those coming off the northern slow of South Georgia, then this result further motivates a spectral approach to MW drag parameterization.

Finally, a brief analysis of how meridional propagation of MWs over the Drake Passage influence zonal winds was presented. Meridionally-propagating MWs enhanced the vertical fluxes of zonal momentum over the Drake Passage by about a factor of four. Zonal GWDs were enhanced by a factor of 2-4 over the Drake Passage. An interesting result from this analysis was that the meridional divergence of the meridional flux of zonal momentum was important, particularly at and below the PNJ max. In fact, here this term is larger than and opposite to the drag inferred from the vertical divergence of vertical flux of zonal momentum, which is typically how GWD is

defined and computed. Not accounting for this term would result in the wrong direction of drag being inferred. This result suggests parameterizations that do account for lateral propagation may need to also account for lateral fluxes of horizontal momentum.

Acknowledgments. Numerous institutions and funding agencies supported this international collaborative work. Support was provided by the International Space Science Institute (ISSI), the Stratosphere-troposphere Processes And their Role in Climate (SPARC) project, and by the Advanced Study Program postdoctoral fellowship at the National Center for Atmospheric Research (NCAR), which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977. High-performance computing was performed on the Cheyenne supercomputer (ark:/85065/d7wd3xhc) with support provided by NCAR's Computational and Information systems Laboratory, also sponsored by the National Science Foundation. Additionally, this work used JASMIN, the UK collaborative data analysis facility. Finally, David Gill, Jimy Dudhia, Jordan Powers, Kevin Manning, and Joe Klemp, all within MMM at NCAR, were essential in getting WRF to run in the deep configuration presented here.

APPENDIX A

Deep WRF Configuration

Through many failed attempts, it was learned that WRF, as distributed, cannot be run deeper than ≈ 1 hPa ($z \approx 45$ km) for more than a handful of time steps in the realistic configuration. Much trial and error revealed three issues that had to be overcome.

The first issue was that the default, high-order horizontal interpolators produced intersecting vertical levels after interpolation. David Gill realized that these interpolators are not monotonic, occasionally allowing interpolated values above or below the nearest four analysis points in regions

of complex terrain. This caused an intersection of the tightly-spaced vertical levels of the IFS analysis. Using only four point, monotonic horizontal interpolators prevented this issue.

The second issue was that the vertical interpolation fails at and above ≈ 35 km. This is due to legacy if logic that threw out analysis levels that were too close in pressure. Setting the namelist parameter "zap_close_levels" to 0.1 Pa, from 500 Pa, prevented this problem.

The final issue that prevented stable integrations with a deep domain was instability in the lateral relaxation zone that blends interior resolved fields with analyses at the lateral boundaries. This instability, often 2 or 3 grid points from from the boundary, would cause the solution to blow up in ≈ 20 time steps or less whether or not there was terrain in the domain. Removing the lateral relaxation zone by setting "spec_bdy_width" to 1 and modifying WRF's existing nudging code to increasingly nudge WRF's lateral boundaries to the same analysis used for boundary conditions allowed WRF to run stably up to 1 Pa ($z \approx 80$ km). This modification to the nudging code essentially serves the same purpose as the default relaxation zone. Why this modification worked and the default lateral relaxation is an unsolved puzzle. Still, this configuration produced very skillful results, as presented above.

APPENDIX B

Overview of current MWD parameterization in CAM

The orographic gravity wave (OGW) drag scheme in CAM6 has been modified in two ways from that used in earlier versions of the model (i.e. McFarlane 1987). First, additional drag from flow blocking and high-drag configurations has been incorporated following the approach of Scinocca and McFarlane (2000) (SM2000). Second, the generation of subgrid scale (SGS) forcing data for OGW has been substantially modified from that used in earlier versions of CAM. OGW orientation

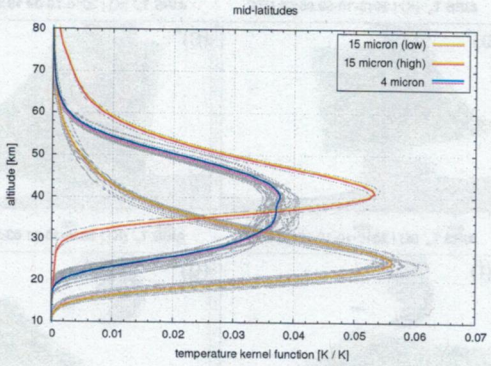
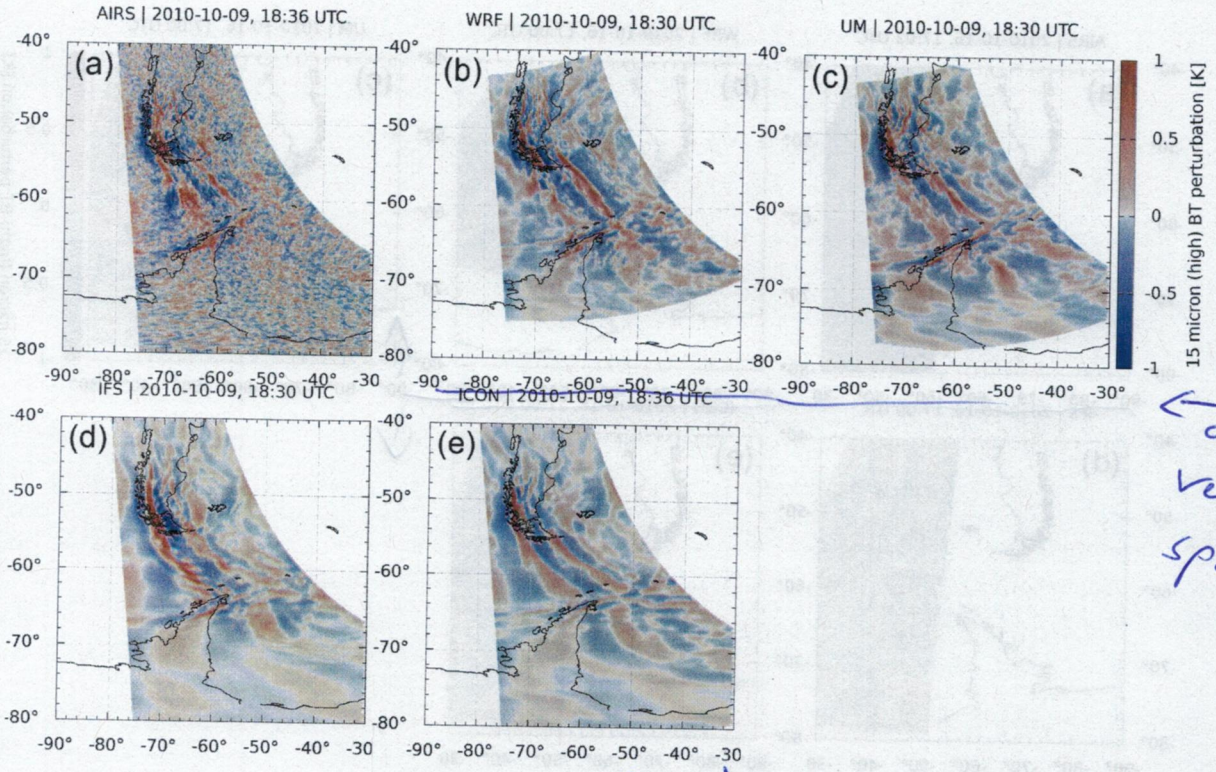


FIG. 3. Individual and average AIRS instrument weighting functions for midlatitude winter.

for different channel sets of the instrument (see plot key/legend).



add vertical space

(a) (b-e)

934 FIG. 4. A qualitative comparison of observed and modeled 15-μm-high (z≈40 km) AIRS brightness tem-
935 perature perturbations. In this comparison, all models were remarkably skillful in reproducing the observed
936 middle-atmopshere mountain waves.

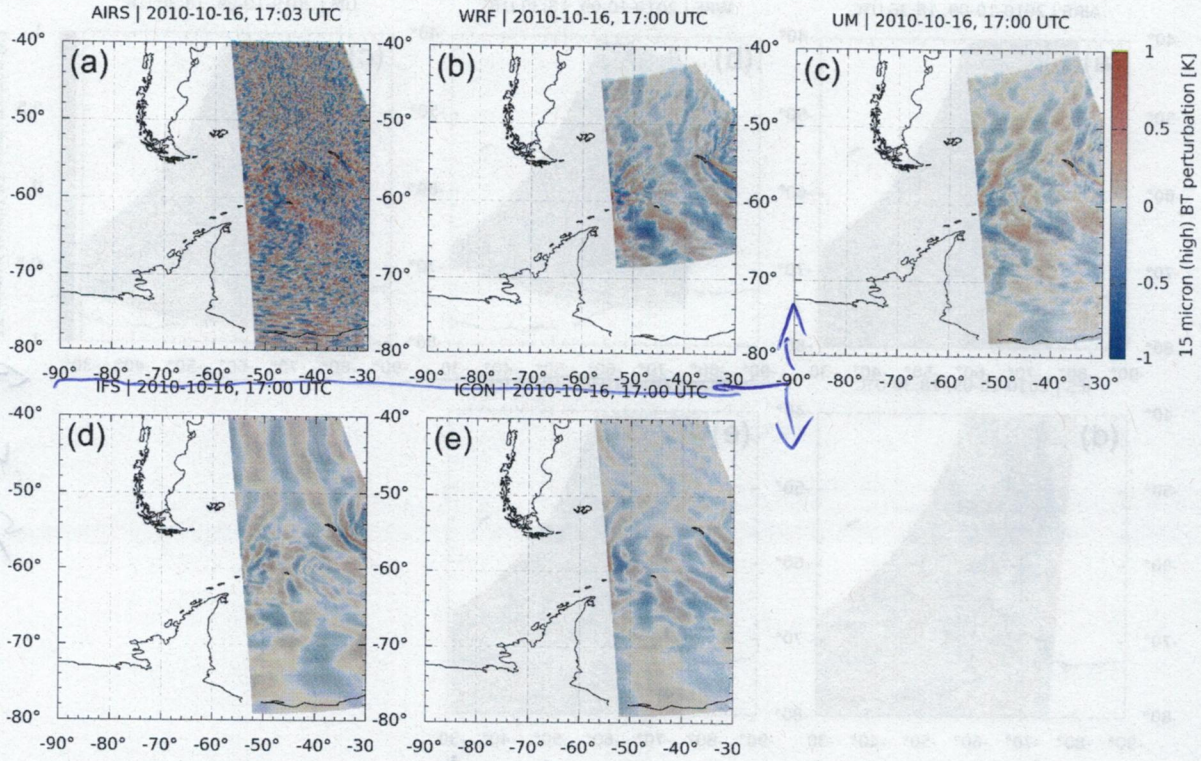
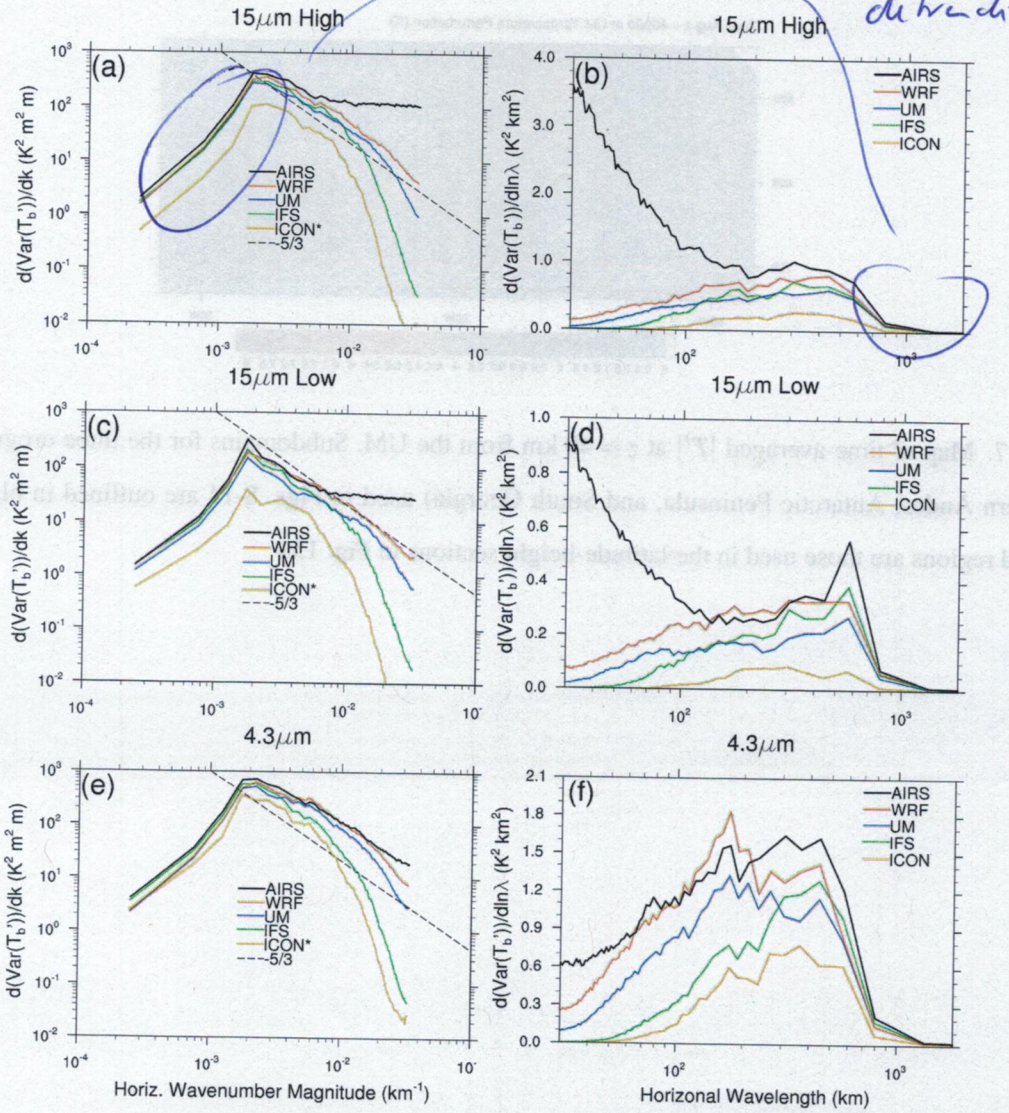
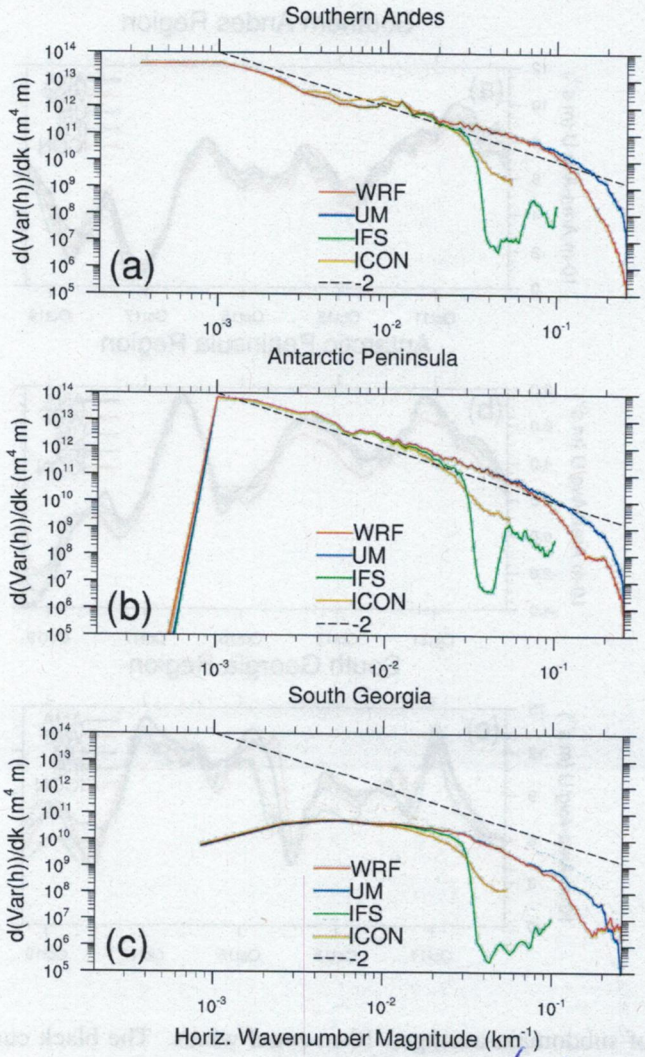


FIG. 5. Same as Fig. 4, but for an overpass where the models had moderate skill.

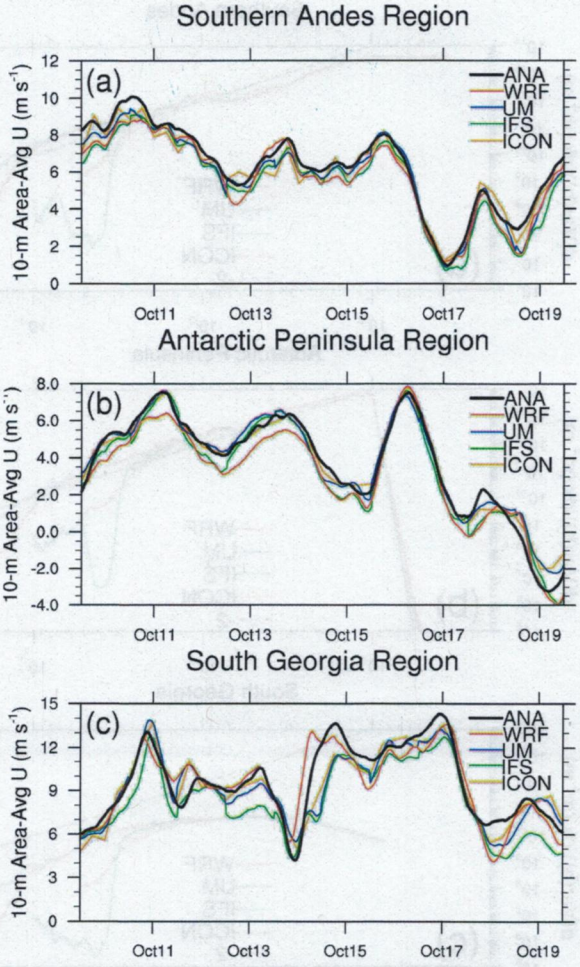


937 FIG. 6. Two-dimensional spectra binned into 1-D spectra for each overpass, then summed up over all over-
938 passes plotted in two ways. Area under the curves in (b), (d), and (f) are proportional to area-integrated T_b' in
939 overpasses in Fig. 2, and further summed over all overpasses (i.e. $\propto \sum_{N_{op}} Var(T_b')$, where $Var(T_b') = \iint T_b' \, dx \, dy$
940 and N_{op} is the number of overpasses).



of terrain height (?)

FIG. 8. 2-D FFT spectra binned by wavenumber for each subdomain for the four participating models. Prior to analysis, terrains of all models were regridded onto a common lat/lon grid with a resolution of ≈ 3 km. Note that $Var(h) = \iint h^2\,dx\,dy$.



947 FIG. 9. Comparison of subdomain-averaged 10-m zonal winds. The black curve labeled "ANA" shows the
948 subdomain-averaaged 10-m winds in the operational IFS analyses used to force all models.

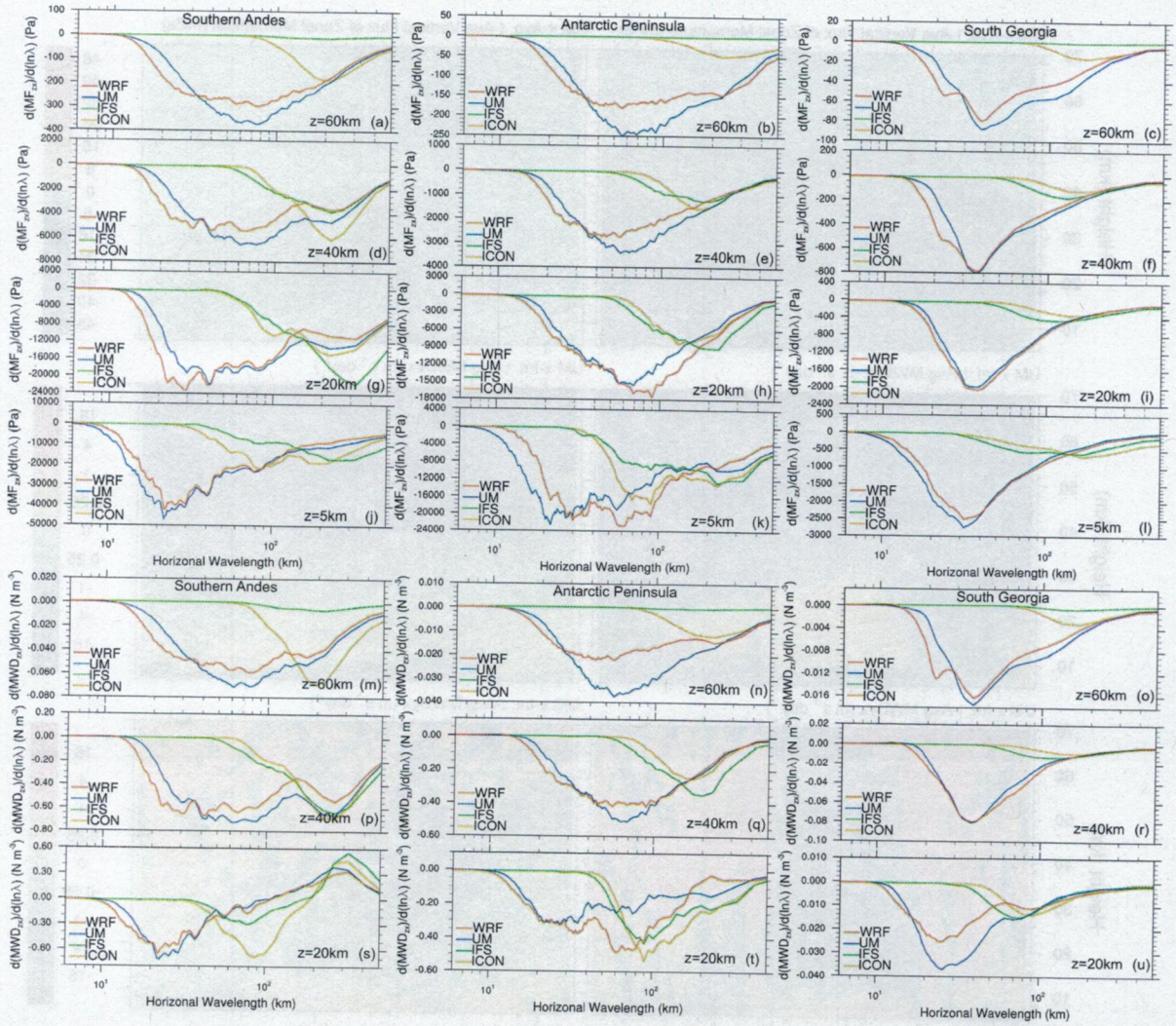


FIG. 14. Time-averaged zonal MF and MWD spectra intercomparison. MWD spectra differenced overe ±5 km.