# Characteristics of gravity waves from convection and implications for their parameterization in global circulation models

Claudia Stephan *

*Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado*

M. Joan Alexander

*NorthWest Research Associates, Inc., CoRA Office, Boulder, Colorado*

Jadwiga Richter

*National Center for Atmospheric Research, Boulder, Colorado*

* *Current affiliation:* National Centre for Atmospheric Science - Climate, Department of Meteorology, University of Reading, Reading, United Kingdom

*Corresponding author address:* Claudia Stephan, Department of Meteorology, University of Reading, P.O. Box 243, Reading RG6 6BB, United Kingdom

E-mail: c.c.stephan@reading.ac.uk

ABSTRACT

Characteristic properties of gravity waves from convection over the Continental United States are derived from idealized high-resolution numerical simulations. In a unique modeling approach, waves are forced by a realistic thermodynamic source based on observed precipitation data. The square of the precipitation rate and the gravity wave momentum fluxes both show log-normal occurrence distributions, with long tails of extreme events. Convectively generated waves can give forces in the lower stratosphere that at times rival orographic wave forcing. Throughout the stratosphere, zonal forces due to convective wave drag are much stronger than accounted for by current gravity wave drag parameterizations, so their contribution to the summer branch of the stratospheric Brewer-Dobson circulation is in fact much larger than models predict. A comparison of these forces to previous estimates of the total drag implies that convectively generated gravity waves are a primary source of summer hemisphere stratospheric wave drag. Furthermore, intermittency and strength of the zonal forces due to convective gravity wave drag in the lower stratosphere resembles analysis increments, suggesting that a more realistic representation of these waves may help alleviate model biases on synoptic scales. The properties of radar precipitation and gravity waves seen in this study lead to a proposed change for future parameterization methods that would give more realistic drag forces in the stratosphere without compromising mesospheric gravity wave drag.

# 1. Introduction

The meridional equator-to-pole Lagrangian-mean circulation in the stratosphere, the Brewer-Dobson circulation (BDC, Brewer (1949); Dobson (1956)), controls various dynamic and thermodynamic properties of the stratosphere. For instance, it plays a role in determining the temperature of the tropical tropopause, the amount of water vapor entering the stratosphere, the transport of aerosols, ozone and other trace gases, as well as the period of the tropical quasi-biennial oscillation.

The time scales of the BDC vary from several years in the upper stratosphere and mesosphere to just weeks right above the stratosphere. The slow overturning into the mid and upper stratosphere, often referred to as the "deep branch" of the BDC, is mainly present in the winter hemisphere (e.g. Birner and Boenisch (2011)), mostly driven by planetary waves (Plumb 2002) but partly by gravity waves (GWs). Okamoto et al. (2011) highlight the importance of orographic and non-orographic GWs in influencing the formation of the summer hemispheric upward branch of the winter circulation. The summer hemisphere branch and the seasonal variation in strength of the circulation is affected by small-scale GWs (Alexander and Rosenlof 1996). Alexander and Rosenlof (2003) show that smaller scale GWs also dominate the wave forcing in the spring-to-summer transition season in each hemisphere. In the lower stratosphere, the "lower branch" of the BDC is more symmetric between the hemispheres, mainly driven by synoptic- and planetary-scale waves and partly by GWs (Plumb 2002).

An intermodel comparison of the annual-mean upward mass flux at 70 hPa in comprehensive chemistry-climate models shows statistically significant agreement on the total strength of the circulation (Eyring et al. 2010). However, there is large variability in terms of the relative contributions of parameterized GWs versus resolved Rossby waves, ranging from close to zero to about half for GWs. The uncertainty is particularly large for non-orographic GWs. Furthermore, a 2.0-3.2% per decade acceleration of the BDC is seen across models but there again exists no consensus on the contributions from different waves types in driving this trend (Butchart et al. 2006, 2010; Cohen et al. 2014; Abalos et al. 2015). The differences in resolved versus gravity wave contributions reflect our poor ability to simulate gravity waves. Deficiencies, especially in parameterizations used for non-orographic GWs, remain a great motivation for improving our knowledge and understanding of atmospheric GWs, both through observations and numerical modeling (Alexander et al. 2010).

In this study an idealized version of the Weather Research and Forecasting (WRF) model is used to determine characteristic properties of GWs from Continental U.S. convection, in particular those quantities relevant to their parameterization in global models, for instance the amplitude spectrum and frequency of occurrence. The modeling approach is unique in that all simulations are carried out at a high horizontal resolution of 4 km and waves are forced by a realistic thermodynamic source based on observed precipitation data. At the same time the model is efficient enough to allow for long simulations on deep domains covering most of the Continental U.S. The numerical model and use of precipitation data are described in Section 2. In Section 3 the topic of wave intermittency is addressed, one of the most challenging aspects of non-orographic GW drag parameterizations. We will first show that the distribution of wave amplitudes over the summer U.S. agrees well with the universal shape of amplitude spectra observed and modeled in other regions of the globe. Secondly, we will compare the zonal wind tendencies from our model results to those in the Modern-ERa Retrospective Analysis for Research and Applications (MERRA) reanalysis and highlight deficiencies in their GW drag parameterizations. In Section 4 we compute the contribution of our simulated GWs to the forcing of the BDC and compare to the parameterized wave forcing in MERRA and the Community Atmosphere Model (CAM). Potential avenues for improving GW drag parameterizations in global models are discussed in Section 5, which also serves as a conclusion.

# 2. Experimental setup

### a. *A numerical model with a realistic gravity wave source*

This study uses the modeling approach described in Stephan and Alexander (2015), where a nonlinear idealized dry version of the WRF model is forced with high-resolution latent heating/cooling derived from precipitation observations over the Continental U.S. For several case studies, it was shown that this model produces an excellent quantitative comparison to waves observed by satellite.

Here, we simulate the entire month of June 2014 and an area covering most of the Continental U.S. at a high horizontal resolution of 4 km. Fig. 1 shows the arrangement of 10 sub-domains, each spanning 1000 km×1000 km. To exclude numerical artifacts close to the domain boundaries the idealized WRF model is run on slightly larger domains with a horizontal area of 1400 km×1400 km. Fig. 1 shows the centers of these domains, but there exist overlapping zones on each side of a domain that measure 200 km. This has the additional benefit of accounting for wave horizontal propagation: GWs that are triggered by convection close to a boundary and propagate out of their 1000 km×1000 km domain will be captured by the adjacent domain. Every 24 h independent model simulations are launched for each sub-domain.

Each sub-domain is initialized every day at 00:00 UTC with a one-dimensional daily-mean MERRA horizontal wind and potential temperature profile computed at the MERRA grid point closest to the center of the sub-domain. The 1000 km horizontal extent of the ten sub-domains corresponds to the lower limit of what are considered synoptic length scales. Therefore, large-scale background wind patterns, which are key for modeling wave-mean flow interactions, are adequately captured by our experimental setup. In terms of the vertical grid, there are 104 vertical levels with a spacing increasing linearly from 100 m at the surface to 600 m at 2400 m, and a constant separation of 600 m above 2400 m. The model top is at 65 km (0.1 hPa), with the upper 10 km consisting of a damping layer. For a detailed description of the model, see Stephan and Alexander (2015).

The heating algorithm for converting rain rates to latent heating/cooling is developed, tested and described in detail in Stephan and Alexander (2015). The algorithm is derived from the precipitation and latent heating field of a full-physics WRF simulation which includes the developing, mature and decaying stages of typical continental convection. It relates 10-min surface precipitation rates averaged over an area of 4 km×4 km that exceed a convective threshold of 1 mm/10 min to the average profile of latent heating and cooling. The amplitudes and depths of the heating/cooling profiles are linear functions of precipitation rate. In Stephan and Alexander (2015) the idealized model was run with the original heating and cooling field and with the algorithm-derived heating/cooling to show that employing a convective threshold and using average profiles instead of original profiles does not have a large impact on the generated GW momentum flux spectrum. The idealized modeling approach reproduced the shape of full-physics GW momentum flux well and the total integrated flux was within $\pm 20\%$.

The heating algorithm is suitable for precipitation data with a horizontal resolution of 4 km×4 km and a temporal resolution of 10 min. Model runs over extended periods of time and large areas require a gridded precipitation data set. In this study we use the National Centers for Environmental Prediction/Environmental Modeling Center's (NCEP/EMC) 4 km gridded Stage IV precipitation data to derive the time-varying heating/cooling field. The Stage IV analysis is based on the multi-sensor hourly Stage III analysis produced by the 12 River Forecast Centers in the Continental U.S. After a manual quality control performed at the River Forecast Centers it is made into a national product. The horizontal resolution of the idealized WRF model is chosen to match the Stage IV horizontal grid. The total precipitation for June 2014 is shown in Fig. 1 as colors.

While the horizontal resolution of the Stage IV analysis is appropriate for modeling GW generating convective cells, the temporal resolution of 1 h is not high enough to capture the intermittency of localized intense cells that have been observed as intense GW sources.

Therefore, we have developed a statistical method to construct 10 min precipitation data from the hourly data.

### *b. From hourly to 10 min precipitation values*

#### 1) Derivation of the precipitation algorithm

Our goal is to compute the probability $P(P10|P60)$, i.e. the probability of 10 min values of precipitation $P10$ given a 60 min value $P60$. The six 10 min values are not independent as their sum needs to equal P60.

Statistics that describe how hourly accumulation values break down into 10 min accumulation values can be inferred from analyzing precipitation data with an original temporal resolution of 10 min. To this end, we obtain the Storm Total Rainfall Accumulation Product (STP) for individual Next-Generation Radar (NEXRAD) stations. The STP product provides radar-estimated rainfall accumulations within 230 km of the radar in polar coordinates with a resolution of 2 km$\times 1^\circ$. Data from several stations are interpolated in space and time to obtain a 10 min 4 km$\times$4 km mosaic. In this process we average overlapping arrays from different stations to obtain smooth maps. This procedure is carried out for areas of 2000 km$\times$2000 km, time periods of 24 h and for 5 different storms: A mesoscale convective complex (20 June 2007), a squall line (5 June 2005), a mesoscale convective system (13 June 2013), and two events of intense convection with a smaller degree of organization, one in the Southeast (8 June 2014), and one in the Midwest (19 June 2014).

The purple histograms, labeled original data, in Fig. 2 are the distributions of 4 km$\times$4 km 10 min rain rates greater than zero for the 5 storm cases. The 99th and 95th percentiles are shown in each panel. Solid lines are lognormal distributions with the same mean and standard deviation as the data.

The original 10 min accumulations, denoted $P10$, are next integrated to obtain hourly accumulations, $P60$. Then, for each 10 min interval that was used in computing $P60$, we calculate the factor $m = P10/P60$, where $0 \leq P10 \leq P60$ and $0 \leq m \leq 1$. A value of $m = 1$ corresponds to all precipitation falling within 10 min. The goal is to compute the probability distribution $P(m|P60)$ of the factors $m$ given a value for $P60$: The higher the value of $P60$, the higher is the probability that it rained for a longer period of time and the probability distribution becomes more strongly peaked around $m = 1/6$. For small $P60$ the probability that all rain fell within only 10 min increases and larger values of $m$ occur more frequently.

For use in the algorithm, we combine the data from all storms and separate it into five categories based on the values of the hourly accumulation: 0 mm/h$<$ $P60$ $<$10 mm/h, 10 mm/h$\leq$ $P60 < 20$ mm/h, 20$\leq$ $P60$ $<$30 mm/h, 30 mm/h$\leq$ $P60$ $<$40 mm/h and 40 mm/h$\leq$ $P60$. For each category let $\nu$ denote the probability that no rain fell within a 10 min interval ($m = 0$). The values of $\nu$ are given in Table 1. As expected, the likelihood that no rain falls within some fraction of the hour decreases with increasing hourly accumulations. The probability distributions $P(m|P60)$ for $m > 0$ can be approximated by lognormal distributions with mean values $\mu$ and standard deviations $\sigma$, also given in Table 1:

$$P(m|P60) = \frac{1}{m\sigma\sqrt{2\pi}} e^{\frac{-(\ln(m)-\mu)^2}{2\sigma^2}} \tag{1}$$

Indeed, as argued earlier, $\mu$ decreases with larger $P60$, which means that small values of $m$ become more likely. This translates to $P60$ being more equally distributed over the hour.

The algorithm for deriving 10 min values from an hourly value $P60$ works as follows. First, the precipitation strength category is determined. If $P60 \geq 40$ mm/h we assign $P10 = P60/6$ for all six 10 min intervals that make up this hour. Otherwise we use the appropriate values for $\nu$, $\mu$ and $\sigma$ from Table 1 and loop through five of the six 10 min intervals. These five intervals do not correspond to the first 50 minutes of the hour but are chosen randomly to ensure that precipitation statistics are identical for all 10 min intervals within the hour. For each of the 5 randomly chosen 10 min intervals $1 \leq j \leq 5$ we determine whether rain fell or not in a binomial trial, where the probability that rain fell is $p = (1 - \nu)$. If rain fell, the lognormal distribution given by $\mu$ and $\sigma$ is randomly sampled to obtain $m_j$

and we assign $P10_j = m_j \times P60$ to time interval $j$. Should for some time $j > 1$, $\sum_{i=1}^{j} m_i > 1$, the random sampling of the lognormal distribution is repeated. For the last interval, $j = 6$, we assign $m = 1 - \sum_{j=1}^{5} m_j$ to ensure that $P60$ is matched exactly.

The green histograms in Fig. 2 show the distributions of 10 min precipitation values reconstructed from the hourly data. A two-sided Kolmogorov-Smirnov test is performed to quantify the similarity of the two histograms shown in each panel and the significance is shown at the bottom. Overall there is excellent agreement. The worst match is found for the squall-line case (5 June 2005). We suspect this can be attributed to the fast propagation speed of this storm and/or to this storm having particularly high precipitation rates.

The precipitation algorithm accurately reproduces the statistical distributions of 10 min precipitation values. A good match of overall precipitation amount and of intense rain events found in the tails of the distributions is essential for triggering a realistic GW spectrum in the idealized model. However, there are additional factors that can affect the shape of the GW spectrum above the storm, for example the horizontal distribution and organization of precipitation cells and the frequency distribution of the heating in time. When applying the precipitation algorithm outlined above, these variables are partially constrained because the precipitation algorithm is designed to exactly reproduce the hourly accumulation value at each grid point. The sub-hourly distribution of precipitation on the other hand is left to chance. Therefore, additional validation of the GWs generated by the precipitation algorithm is required.

#### 2) Validation of waves generated by the precipitation algorithm

To validate GWs generated by the precipitation algorithm we perform a total of four simulations using the configuration described in Section 2 but domain sizes of 2000 km$\times$2000 km. Two simulations are carried out for the mesoscale convective complex (20 June 2007) and two for the squall line (5 June 2005) case. For each case one simulation is based on the original 10 min precipitation data set and the other on the reconstructed 10 min data.

We selected these two storms because in terms of the distributions shown in Fig. 2 they represent the best and worst match of reconstructed and original 10 min data.

Fig. 3 shows the absolute GW momentum flux spectra at 15 km as a function of phase speed and propagation direction for the 4 runs. The spectra are computed from 24 h of horizontal and vertical wind velocities saved every 10 min, using the method described in Stephan and Alexander (2014). The white line in each panel corresponds to the 700 hPa wind and the black dashed lines to the winds at levels between 700 hPa and 15 km. In agreement with theory, the black dashed lines coincide well with regions of dissipation, as critical level filtering occurs when a wave approaches a level where the phase speed equals the wind speed. Overall, the similarity between the runs based on the original and the reconstructed data is remarkable. For the squall line case there is some flux missing in the direction of the 700 hPa wind. The 700 hPa wind is commonly used for estimating the propagation direction and speed of the convective cells. The fact that the difference between the simulations is largest in this direction supports the assertion that it is the higher-than-average propagation speed of this storm which causes the relatively poor match found in the analysis of Fig. 2.

The spectra in Fig. 3 represent daily averages over a very large area and do not contain information about instantaneous and local magnitudes of momentum flux. The amplitude of GWs above convection is strongly tied to the strength of the underlying heating cells, which remain subgrid-scale in most climate models and represent one of the most difficult parameters to constrain in GW drag parameterizations (Richter et al. 2010). Knowledge of the local, instantaneous wave amplitudes is crucial because they determine the breaking levels of GWs.

The benefit of the modeling approach introduced in Stephan and Alexander (2015) is that the heating magnitude is directly related to observed precipitation. To verify that the realism of local wave amplitudes is not suffering from constructing 10 min precipitation data from hourly data, Fig. 4 shows the probability distributions of 100 km$\times$100 km instantaneous flux magnitude at 15 km (left) and 35 km (right) height, for the squall line case (top) and the mesoscale convective complex (bottom). The flux magnitudes are derived by first computing $\hat{u}(x, y, k, l)$, $\hat{v}(x, y, k, l)$ and $\hat{w}(x, y, k, l)$ every 10 min using a two-dimensional S-transform (Stockwell et al. 1996). Here, $\hat{u}$, $\hat{v}$ and $\hat{w}$ are the zonal, meridional and vertical wind component amplitudes, $x$ and $y$ denote the horizontal grid coordinates, $k$ and $l$ the zonal and meridional wave numbers. Fig. 4 shows values of momentum flux up to the 90th percentile, obtained by integrating $\sqrt{(\hat{u}\hat{w}^* + \hat{v}\hat{w}^*)}$ over all $k$ and $l$ and areas of 100 km$\times$100 km and multiplication by the air density $\rho$. Here, $\hat{w}^*$ denotes complex conjugation. The range of fluxes is shorter at the higher altitude because the largest amplitude waves have dissipated below. The similarity of the purple (original resolution of 10 min) and green histograms (reconstructed data) is assessed with a two-sided Kolmogorov-Smirnov test. The largest discrepancy occurs at 35 km for the squall line case as should be expected from the previous discussion. In general the differences between the distributions are small and the agreement very good.

# 3. Intermittency in simulated gravity wave spectra

Previous modeling efforts as well as observational studies with stratospheric balloons and satellites emphasize the high intermittency of the GW field (e.g., Hertzog et al. (2012);Hertzog et al. (2013)). This has implications for GW parameterizations in global models. A given average flux produced by a large number of small-amplitude wave events will produce drag at much higher altitudes than the same average flux carried by a small number of high-amplitude wave packets. As argued in the previous section, the idealized model uses a precipitation field with a highly realistic variability as input. In this section we quantify the intermittency of the GW momentum flux spectrum over the Continental U.S. for the month of June 2014.

### *a. Momentum flux amplitude*

The top panel of Fig. 5 shows probability density functions of simulated absolute zonal momentum flux amplitudes averaged over 100 km×100 km and 3 h for different altitudes. The average is computed from an accumulated value of $u'w'$, which is updated every 15 seconds. Cloud-resolving models predict that convectively generated GWs typically have time periods ranging from 10 min to several hours (e.g. Piani et al. (2000)). The 3 h interval for averaging is chosen to include contributions of waves with a large range of frequencies while minimizing the effect of wave cancellation that can occur when waves propagating in opposite directions overlap: The speed at which average storms travel is small enough to produce an approximately concentric wave field of waves propagating out and away from the source.

The mean value, 90th and 99th percentiles as well as the percentages of flux associated with values larger than the percentiles are also indicated. The black dashed line is a log-normal distribution with the same mean and standard deviation as the spectrum at 15 km. Lognormal distributions have been found to describe well the spectra of GW momentum flux in other regions of the world. Hertzog et al. (2012) examined Vorcore balloon and High Resolution Dynamics Limb Sounder (HIRDLS) satellite observations of absolute zonal momentum flux between 50°S and 65°S at 20 km over the Southern Ocean and found that both data sets are well approximated by lognormal distributions. In their study of deep tropical convection, Jewtoukoff et al. (2013) also obtained lognormal distributions of absolute momentum flux from balloon observations during the PreConcordiasi campaign. In particular, they found a typical mean momentum flux value of 5 mPa in the tropics at 20 km during the months of February to May, which is close to our mean value of 6 hPa at 20 km.

Another feature that our results share with previous findings is self-similarity. The 90th and 99th percentiles of momentum flux distributions explain about the same proportions of the total flux at different altitudes, 50% for the 90th percentile and 10% for the 99th percentile. Hertzog et al. (2012) reported self-similarity with identical fractions in their WRF simulations over Antarctica, examining different heights. These findings are in agreement with work by de la Camara et al. (2014), who encountered these same proportions in their analysis of a multiwave stochastic parameterization of non-orographic GWs tuned and tested against Concordiasi observations. Specifically, their analysis suggests that this self-similarity holds independent of season, latitude and height.

Furthermore, de la Camara et al. (2014) suggest that the lognormality of the GW momentum flux source spectra may be related to a lognormal behavior of the squared precipitation probability density function. This quantity is shown as histograms in the bottom panel of Fig. 5 for the Stage IV data, labeled observations (green), and the reconstructed precipitation data (black), which we use in the heating algorithm for forcing the idealized WRF model. The resolution has been degraded to 100 km$\times$100 km and 3 h to match that of the momentum flux amplitudes. The dashed lines are lognormal distributions with the same mean and standard deviation. Indeed, the lognormal curves represent the precipitation strength distributions very accurately up to the 99th percentiles. They tend to slightly overestimate the occurrence frequencies of large precipitation rates, but this is also true for the momentum flux amplitudes in the upper panel.

Also shown are the precipitation strength distributions for MERRA reanalysis data during June 2014 and the CAM5 model. The CAM5 precipitation data used in this plot are composed of different years of CAM5 runs, as will be explained in detail in section 4.*a*. We notice that both MERRA and CAM5 underestimate stronger precipitation rates and do not follow lognormal distributions, as can be seen by comparing the histograms to the corresponding dashed lines. This has implications for the potential of improving the parameterizations of non-orographic GWs in these models.

### *b. Zonal wind tendencies in the stratosphere*

Next, we will examine the GW drag in the idealized WRF model and compare to MERRA reanalysis. Orographic waves are stationary and break at lower levels, whereas the non-orographic spectra include a range of phase speeds. Orographic GW drag in MERRA is parameterized using the scheme by McFarlane (1987) and non-orographic wave effects are based on Garcia and Boville (1994). In MERRA history files, orographic and non-orographic GW drag are combined and saved in one field. To compare to the non-orographic component of the forcing, we select regions 2, 7 and 8 (see Fig. 1), because the contribution of orographic waves is negligible there. The top panel of Fig. 6 shows the WRF daily mean zonal forcing, which is given by $F_z = -\frac{1}{\rho}\frac{\delta}{\delta z}\left[\overline{\rho u'w'}\right]$, where $\overline{\rho u'w'}$ is the momentum flux as computed in Section $2.b.2$. To facilitate the comparison to MERRA, values have been interpolated to MERRA pressure levels. The blue line below the top panel shows the time evolution of 10 min Stage IV precipitation averaged over sub-domains 2, 7 and 8. Daily mean MERRA precipitation averaged over these 3 sub-domains is almost identical to Stage IV. The panel in the center shows the MERRA GW drag and the bottom panel the MERRA GW drag plus analysis zonal wind increments. During a 6h-update cycle, the analysis corrections (observation-minus-background departures) are applied to the forecast model through an additional tendency term in the model equations (Rienecker et al. 2011). The panels on the right show the monthly mean forcing (solid purple line) plus/minus one standard deviation (dotted green lines).

Comparing the WRF and MERRA GW drag, it is apparent that the forcing in WRF is at least one order of magnitude stronger. This can be attributed to waves with large amplitudes that are triggered by intense convection and break in the stratosphere. The GW source spectrum in MERRA is not tied to the underlying convection, misses these high-amplitude waves completely, and therefore exhibits a very homogeneous behavior in time in the stratosphere. Analysis wind increments in the middle atmosphere are thought of as partially correcting for missing GW drag in coarse models (e.g, McLandress et al. (2012)), and when considering MERRA GW drag plus analysis increments the temporal intermittency in the lowermost stratosphere below 50 hPa compares much better to WRF. This suggests that a more realistic representation of convectively generated GWs may help alleviate model biases near the tropopause on synoptic scales and represents a problem worth further investigation in the future.

# 4. The contribution to the Brewer-Dobson circulation

In this section we quantify the role of GWs from Continental U.S. convection in driving the Brewer-Dobson circulation by comparing to the forcings in MERRA and CAM5.

## *a. CAM5 data*

For a detailed description of the CAM5 model used in this section see Richter et al. (in review) and the references therein. The model has 46 vertical levels with a model top at 0.3 hPa and a horizontal resolution of  100 km. The parameterization of non-orographic GWs follows Richter et al. (2010) and includes a frontal GW drag scheme as well as a convective GW drag scheme.

The convective GW drag scheme is a so-called source parameterization based on Beres (2004). Source parameterizations link characteristics of GWs to the underlying wave source, namely the convective heating field in the model. One key parameter in the Beres scheme is the convective heating rate, which determines the amplitude of the waves. However, this quantity is only known as an average over a model grid box. To estimate a heating rate representative of individual convective cells, it is assumed that convection takes up 5% of the area of a grid box. Wave amplitude, specified as momentum flux, is proportional to the square of this local heating rate. As a consequence, the amplitude of the waves is the least certain aspect of this parameterization.

In addition to wave amplitude, wave horizontal phase velocities and propagation directions need to be estimated. These are primarily affected by the depth of the heating and by the mean tropospheric winds. Once amplitude and propagation characteristics are determined, the parameterization launches waves at the top of the convective heating. Wave drag is created at levels where the upward propagating waves dissipate above the wave breaking level according to the Lindzen-McFarlane parameterization method (see Garcia et al. (2007)).

Given this sensitivity to the heating and the background wind profile, for a comparison between the GW drag in WRF versus CAM5 it would be ideal if both models had identical mean background wind profiles and similar precipitation characteristics. Since this is generically not the case, we find a corresponding June from the 10 year CAM5 simulation that most closely matches the zonal wind and precipitation strength in the WRF simulations, separately for each domain. Fig. 7 shows the corresponding June monthly mean zonal wind and precipitation strength distributions for WRF (solid lines) and CAM5 (dashed lines). For all domains the monthly mean value of 100 km $\times$100 km average precipitation rate, shown in the panel, is smaller for CAM5. In addition, as noted beforehand in the discussion of Fig. 5, CAM5 as well as MERRA underestimate stronger precipitation rates.

### *b. Missing convective GWD*

The left panel of Fig. 8 displays the zonal wind tendencies averaged over June 2014 and the area covered by the 10 model domains for the WRF simulations (purple), all CAM5 GW drag schemes combined (orange) and MERRA GW drag plus analysis corrections (green). The middle panel shows how the CAM5 tendencies break down into forcing from convective, frontal and orographic GWs, and the panel on the right distinguishes between MERRA GW drag and analysis increments. Recall that the values for the CAM5 model are composed of different years of simulations as described in the previous paragraph.

Alexander and Rosenlof (1996) computed the contribution of small-scale waves (wavelengths $\lesssim$ 1000 km) to the forcing of the BDC as the residual difference between total and resolved forcing estimates for data sets from the National Meteorological Center and the UK Meteorological Office, and the Upper Atmosphere Research Satellite (UARS). For June they obtained typical values of -1 m/s/day at 10 hPa and +4 m/s/day at 1 hPa (their Figure

1). Table 2 lists the contribution of the modeled area to the zonal mean forcing for all available components of GW drag and in three different altitude layers representing lower (100-10 hPa) and upper (10-1h Pa) stratosphere and stratopause (1-0.4 hPa), i.e. the forcing averaged over the simulated area multiplied by 0.16, because our simulations cover 16% of the total area in the latitude band 25.7-48.5°N. Assuming that the remaining 84% of the latitude band provide a similar wave driving, the WRF GW drag tendencies averaged over 10-1 hPa and 1-0.4 hPa constitute a fairly good match of the values reported by Alexander and Rosenlof (1996). There is evidence that this assumption may be valid because the precipitation averaged over the area of our study is similar to precipitation averaged over the full latitude band. This comparison of the WRF GWD to Alexander and Rosenlof (1996) provides further evidence that our model of convectively generated waves is realistic, and that these waves can provide all of the unresolved stratospheric forcing needed to drive the Brewer-Dobson circulation at these latitudes.

It is particularly noteworthy that the GW drag in our simulations, which is purely convective, is larger than the CAM5 orographic GW drag averaged over 100-10 hPa. The changes of GW drag with altitude seen in Fig. 8 and Table 2 highlight a common misconception that it is primarily only orographic GW drag that is relevant in the lower stratosphere owing to its large-amplitude waves that break at lower levels, while non-orographic GW drag, as it is currently parameterized, primarily only affects high altitudes. The middle panel of Fig. 8 as well as the numbers in lines 2 and 3 of Table 2, showing separately the convective and orographic GW drag contributions in CAM5 at different levels, illustrate this condition in CAM5. In reality convectively generated GWs can have large amplitudes and therefore also break in the lower stratosphere. Even the more advanced source parameterization in CAM5 underestimates high-amplitude waves, which results in missing GW drag in the stratosphere. We also note that the MERRA GW drag and analysis increments combined (last line of Table 2) are of similar magnitude compared to the WRF tendencies in the lower stratosphere (100-10 hPa), even though their structure with height (Fig. 8) is quite different.

# 5. Discussion and conclusion

We analyzed observed precipitation data and GWs in high-resolution simulations of June 2014 over the Continental U.S. In an idealized version of the WRF model, waves were forced by a realistic thermodynamic source based on observed precipitation data. At horizontal scales of 100 km$\times$100 km we found that the probability distribution of momentum flux amplitudes above the storms and the square of precipitation rate both follow lognormal distributions, a characteristic that has been reported for simulations, as well as observations, in multiple other regions of the globe. An important feature of lognormal distributions is the long tail consisting of rare and extreme values. Not capturing this high degree of variability in wave amplitudes has important implications for GW drag parameterizations, as the wave amplitudes determine the height at which waves break and deposit their momentum.

Comparing the daily mean wave forcing in our simulations to GW drag in MERRA reanalysis data, we found the parameterization in MERRA is underestimating both the variability and the magnitude of the GW drag throughout the stratosphere. This result was somewhat expected because the GW source spectrum in MERRA is homogenous in space and time and therefore does not include high-amplitude wave events. The intermittency and magnitude of zonal wind tendencies stemming from MERRA analysis increments in the lowermost stratosphere are more similar to the simulations.

Lastly, we examined monthly mean zonal wind tendencies in the simulations to evaluate their contribution to the Brewer-Dobson circulation, and compare these to MERRA, and the CAM5 model. The CAM5 model includes an orographic and frontal GW drag scheme, as well as a convective GW source parameterization. However, neither the GW drag scheme in MERRA nor the more advanced source parameterization in CAM5 are including enough high-amplitude waves. This results in missing GW drag, particularly in the stratosphere. Previous studies found similar deficiencies in the tropics. For example, Schirber et al. (2014) showed that aspects of the quasi-biennial oscillation can be improved by using a GW source parameterization instead of assuming constant spectra in the WACCM model. Lott and Guez

(2013) also found a more intermittent spectrum caused wave breaking at lower altitudes, and this helped to decouple the quasi-biennial oscillation from the annual cycle. Bushell et al. (2015) tested a version of the Met Office global models spectral nonorographic scheme with enhanced source intermittency at the launch level and report an improved representation of the quasi-biennial oscillation.

An important aspect of our modeling approach is that we use a statistical method to derive 10 min precipitation values from an hourly data set. Given that precipitation characteristics exhibit a universal behavior it seems conceivable that a similar method could be applied to grid point precipitation values in global models, possibly providing a way to estimate a spatial sub-gridscale variability in addition to the temporal statistical refinement. As a result one could obtain a realistic distribution of cloud-scale precipitation rates. Further, by using a heating algorithm similar to Stephan and Alexander (2015), these precipitation rates could be converted to local heating amplitudes, the most uncertain parameter in current parameterizations. Combining the Beres (2004) parameterization with a stochastic approach by randomly choosing from this heating amplitude distribution has several benefits: A constant convective fraction of 5% would no longer need to assumed. A realistic intermittency in wave amplitudes could be obtained with some waves breaking in the lower stratosphere. Most waves will still have fairly small amplitudes such that it is unlikely to cause aggravating effects on the mesospheric wave forcing as a result of these suggested parameterization changes. This more physically based approach could potentially come at no extra computational cost and adapt naturally to changes in climate.

*Acknowledgments.*

The authors would like to thank the three anonymous reviewers for their constructive comments to improve the manuscript. We would like to thank Dr. Julio Bacmeister, Dr. Robert Stockwell and Dr. Laura Holt for insightful discussions, and Prof. Gretchen Mullendore for pointing us to the NCEP Stage IV precipitation data set. This work was supported by grant number AGS–1318932 from the National Science Foundation programs in Physical and Dynamic Meteorology and Climate and Large-scale Dynamics. Data used in this study are Next-Generation Radar Level-III products available from the National Oceanic and Atmospheric Administration Satellite and Information Service (http://www.ncdc.noaa.gov/data-access/radar-data), the National Centers for Environmental Prediction/Environmental Modeling Center's 4 km gridded Stage IV precipitation data (http://data.eol.ucar.edu/codiac/dss/id=21.093) and the Modern-ERa Retrospective Analysis for Research and Applications (MERRA) reanalysis (http://gmao.gsfc.nasa.gov/research/merra/). The Weather Research and Forecasting model (http://wrf-model.org/index.php) and the Community Atmosphere Model (http://www.cesm.ucar.edu/models/cesm1.0/cam/) are supported by the National Center for Atmospheric Research.

# REFERENCES

Abalos, M., B. Legras, F. Ploeger, and W. J. Randel, 2015: Evaluating the advective Brewer-Dobson circulation in three reanalyses for the period 1979-2012. *J. Geophys. Res. Atmos.*, **120**, doi:10.1002/2015JD023182.

Alexander, M. J. and K. H. Rosenlof, 1996: Nonstationary gravity wave forcing of the stratospheric zonal mean wind. *J. Geophys. Res.*, **101**, 23 465–23 474.

Alexander, M. J. and K. H. Rosenlof, 2003: Gravity wave forcing in the stratosphere: Observational constraints from UARS and implications for parameterization in global models. *J. Geophys. Res.*, **108**, doi:10.1029/2003JD003373.

Alexander, M. J., 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. *Quart. J. Roy. Meteor. Soc.*, **136**, 1103–1124, doi:10.1002/qj.637.

Beres, J. H., 2004: Gravity wave generation by a three-dimensional thermal forcing. *J. Atmos. Sci.*, **61**, 1805–1815.

Birner, T. and H. Boenisch, 2011: Residual circulation trajectories and transit times into the extratropical lowermost stratosphere. *Atmos. Chem. Phys.*, **11**, 817–827, doi:10.5194/acp-11-817-2011.

Brewer, A. W., 1949: Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere. *Quart. J. Roy. Meteor. Soc.*, **75**, 351–363, doi:10.1002/qj.49707532603.

Bushell, A. C., N. Butchart, S. H. Derbyshire, D. R. Jackson, G. J. Shutts, S. B. Vosper, and S. Webster, 2015: Parameterized gravity wave momentum fluxes from sources related to convection and large-scale precipitation processes in a global atmosphere model. *J. Atmos. Sci.*, **72**, 4349–4371, doi:http://dx.doi.org/10.1175/JAS-D-15-0022.1.

Butchart, N., et al., 2006: Simulations of anthropogenic change in the strength of the BrewerDobson circulation. *Climate Dyn.*, **27**, 727–741, doi:10.1007/s00382-006-0162-4.

Butchart, N., et al., 2010: Chemistry-climate model simulations of twenty-first century stratospheric climate and circulation changes. *J. Climate*, **23**, 5349–5374, doi:10.1175/2010JCLI3404.1.

Cohen, N. Y., E. P. Gerber, and O. Buehler, 2014: What drives the Brewer-Dobson circulation? *J. Atmos. Sci.*, **71**, 3837–3855, doi:http://dx.doi.org/10.1175/JAS-D-14-0021.1.

de la Camara, A., F. Lott, and A. Hertzog, 2014: Intermittency in a stochastic parameterization of nonorographic gravity waves. *J. Geophys. Res. Atmos.*, **119**, 11,905–11,919, doi:10.1002/2014JD022002.

Dobson, G. M. B., 1956: Origin and distribution of the polyatomic molecules in the atmosphere. *Proc. R. Soc. Lond. A*, **236**, 187–193, doi:10.1098/rspa.1956.0127.

Eyring, V., T. G. Shepherd, and D. W. Waugh, 2010: Chemistry-climate model validation. SPARC Rep. 5 WCRP-30, WMO/TD-40, 426 pp. [Available online at http://www. sparc-climate.org/publications/sparc-reports/sparc-report-no5/].

Garcia, R. R. and B. A. Boville, 1994: Downward control of the mean meridional circulation and temperature distribution of the polar winter stratosphere. *J. Atmos. Sci.*, **51**, 2238–2245.

Garcia, R. R., D. R. Marsh, D. E. Kinnison, B. A. Boville, and F. Sassi, 2007: Simulation of secular trend in the middle atmosphere. *J. Geophys. Res.*, **112**, 1950–2003, doi:10.1029/2006JD007485.

Hertzog, A., M. J. Alexander, and R. Plougonven, 2012: On the intermittency of gravity wave momentum flux in the stratosphere. *J. Atmos. Sci.*, **69**, 3433–3448, doi:10.1175/JAS-D-12-09.1.

Hertzog, A., M. J. Alexander, and R. Plougonven, 2013: Global observations of gravity wave intermittency and its impact on the observed momentum flux morphology. *J. Geophys. Res. Atmos.*, **118**, 10,980–10,993, doi:10.1002/jgrd.50869.

Jewtoukoff, V., R. Plougonven, and A. Hertzog, 2013: Gravity waves generated by deep tropical convection: Estimates from balloon observations and mesoscale simulations. *J. Geophys. Res. Atmos.*, **118**, 9690–9707, doi:10.1002/jgrd.50781.

Lott, F. and L. Guez, 2013: A stochastic parameterization of the gravity waves due to convection and its impact on the equatorial stratosphere. *J. Geophys. Res. Atmos.*, **118**, 8897–8909, doi:10.1002/jgrd.50705.

McFarlane, N. A., 1987: The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. *J. Atmos. Sci.*, **44**, 1775–1800.

McLandress, C., T. G. Shepherd, S. Polavarapu, and S. R. Beagley, 2012: Is missing orographic gravity wave drag near 60s the cause of the stratospheric zonal wind biases in chemistry-climate models? *J. Atmos. Sci.*, **69**, 802–818, doi:10.1175/JAS-D-11-0159.1.

Okamoto, K., K. Sato, and H. Akiyoshi, 2011: A study on the formation and trend of the Brewer-Dobson circulation. *J. Geophys. Res.*, **116**, doi:10.1029/2010JD014953.

Piani, C., D. Durran, M. J. Alexander, and J. R. Holton, 2000: A numerical study of three-dimensional gravity waves triggered by deep tropical convection and their role in the dynamics of the QBO. *J. Atmos. Sci.*, **57**, 3689–3702.

Plumb, R. A., 2002: Stratospheric transport. *J. Meteorol. Soc. Jpn.*, **80**, 793–809, doi:10.2151/jmsj.80.793.

Richter, J. H., C. Deser, and L. Sun, in review: Effects of stratospheric variability on El Nino teleconnections. *Environ. Res. Lett.*

Richter, J. H., F. Sassi, and R. R. Garcia, 2010: Toward a physically based gravity wave source parameterization in a general circulation model. *J. Atmos. Sci.*, **67**, 136–156.

Rienecker, M. M., et al., 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Schirber, S., E. Manzini, and M. J. Alexander, 2014: A convection-based gravity wave parameterization in a general circulation model: Implementation and improvements on the QBO. *J. Adv. Model. Earth Syst.*, **6**, 264–279, doi:10.1002/2013MS000286.

Stephan, C. C. and M. J. Alexander, 2014: Summer season squall line simulations: Sensitivity of gravity waves to physics parameterization and implications for their parameterization in global climate models. *J. Atmos. Sci.*, **71**, 3376–3391, doi:10.1175/JAS-D-13-0380.1.

Stephan, C. C. and M. J. Alexander, 2015: Realistic simulations of atmospheric gravity waves over the Continental U.S. using precipitation radar data. *J. Adv. Model. Earth Syst.*, **07**, 823–835, doi:10.1002/2014MS000396.

Stockwell, R. G., L. Mansinha, and R. P. Lowe, 1996: Localization of the complex spectrum: the S transform. *IEEE Transactions on Signal Processing*, **44**, 998–1001.

# List of Tables

TABLE 1. Values for the three parameters needed to derive 10 min precipitation rates from hourly precipitation rates for the 4 precipitation categories. Values of $P60$ are given in units of mm/h. Please refer to Section 2.$b$.1 for a description of the parameters.

| category: | $0 < P60 < 10$ | $10 \leq P60 < 20$ | $20 \leq P60 < 30$ | $30 \leq P60 < 40$ |
|---|---|---|---|---|
| $\nu$ | 0.58 | 0.33 | 0.23 | 0.15 |
| $\mu$ | $-1.29$ | $-1.76$ | $-1.86$ | $-1.90$ |
| $\sigma$ | 0.97 | 0.98 | 0.96 | 0.88 |

TABLE 2. Contribution of the area covered by WRF domains to the zonal mean wind tendency for different components of GW drag. Because the simulation area covers 16% of the latitude band 25.7-48.5°N, the numbers below are obtained by multiplying the simulated wave driving by 0.16. Numbers are given in units of m/s/day and are averages of the acceleration in m/s/day at individual pressure levels over the pressure ranges indicated in the top row.

| pressure range: | $100 - 10$ hPa | $10 - 1$ hPa | $1 - 0.4$ hPa |
|---|---|---|---|
| WRF gwd | $-0.072$ | $-0.156$ | 0.639 |
| CAM conv. gwd | $-0.003$ | $-0.014$ | 0.044 |
| CAM oro. gwd | $-0.009$ | $< 0.001$ | 0.001 |
| CAM fro. gwd | $-0.008$ | $-0.035$ | $-0.030$ |
| CAM tot. gwd | $-0.019$ | $-0.048$ | 0.015 |
| MER gwd | $-0.009$ | $-0.029$ | 0.025 |
| MER gwd+ana | $-0.052$ | $-0.037$ | $-0.183$ |

# List of Figures

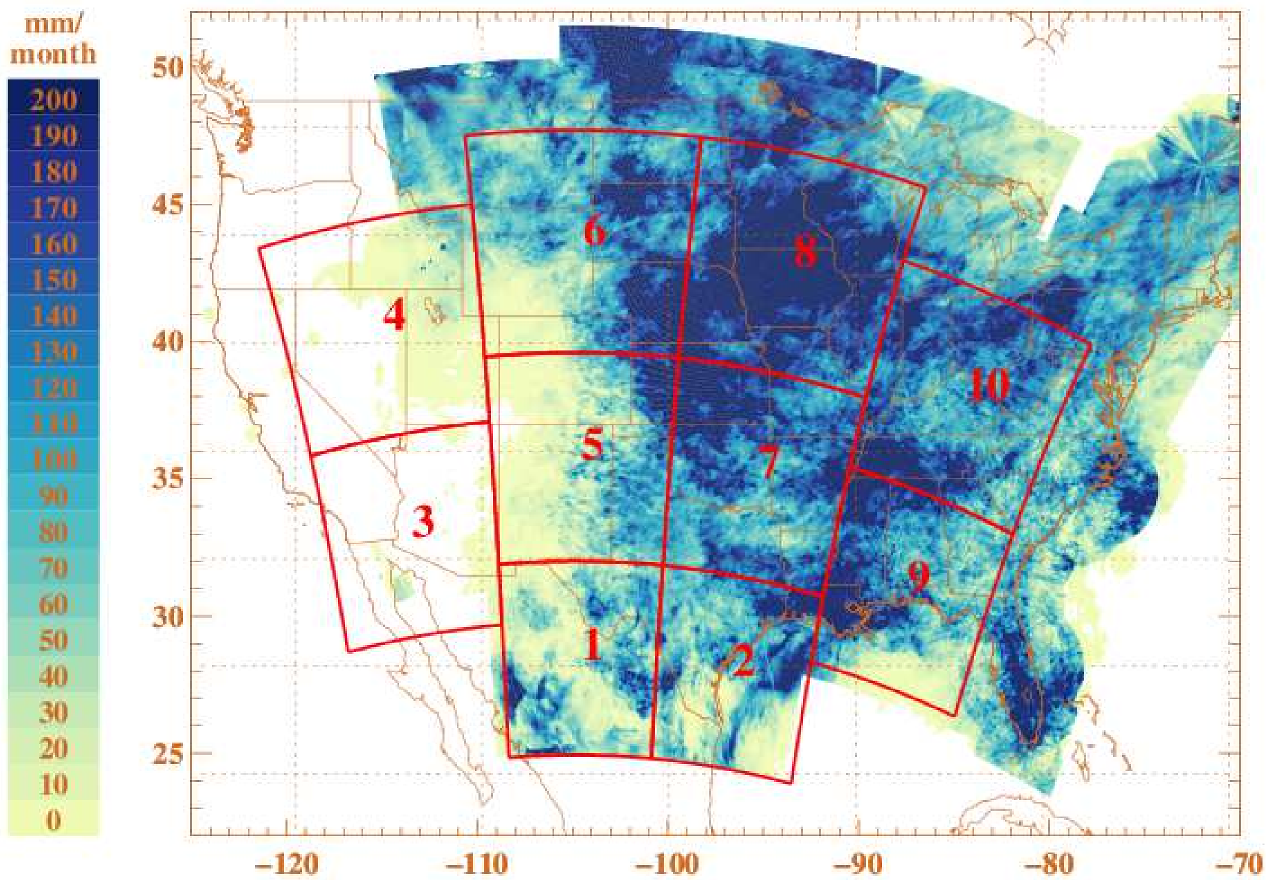


FIG. 1. Map of the locations of the ten 1000 km×1000 km WRF domains that are evaluated in this study. Colors indicate the NCEP 4 km Stage IV total precipitation for June 2014. Latitude and longitude are shown on the axes.

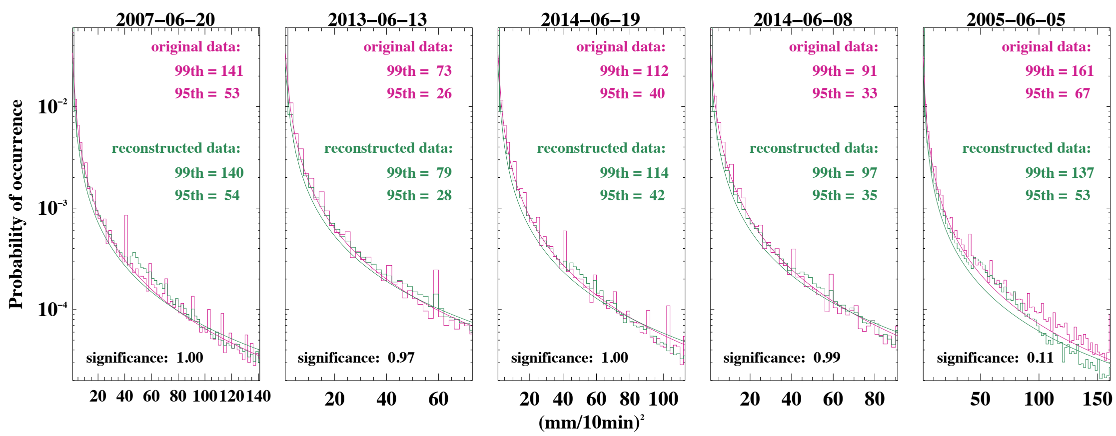


FIG. 2. Histograms of (10-min precipitation rate)$^2$ based on 24 h of data in an area of 2000 km$\times$2000 km, showing occurrence frequencies at a horizontal resolution of 4 km. The 5 panels correspond to 5 different storms. Violet colors denote data with an original temporal resolution of 10 min and green colors are values obtained after degrading the data to an hourly resolution and reconstructing it using the algorithm described in the text. The 99th and 95th percentiles of the distributions are indicated, as well as the probability that both histograms are statistically identical. The solid lines are lognormal distributions with the same mean and standard deviation as the data.

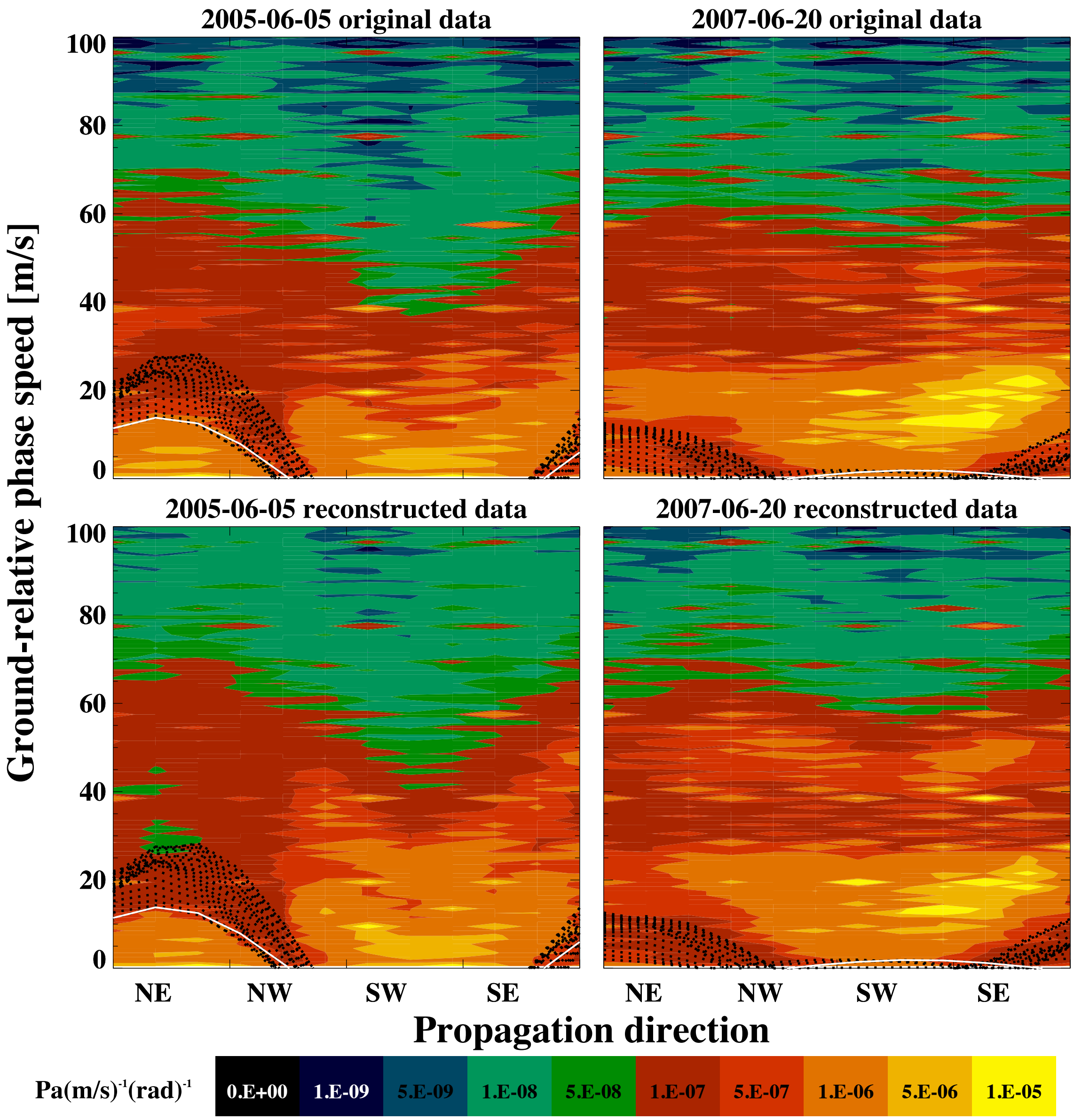


FIG. 3. Total momentum flux spectra for the squall line case (left) and mesoscale convective complex (right) at 15 km altitude as a function of propagation direction and ground-relative phase speed obtained from 24 h simulations on 2000 km×2000 km large domains with a horizontal resolution of 4 km. Labels indicate whether the model is based on the original 10 min data set or the reconstructed data. White lines are the 700 hPa steering level winds and black dashed lines the winds at levels between 700 hPa and 15 km.

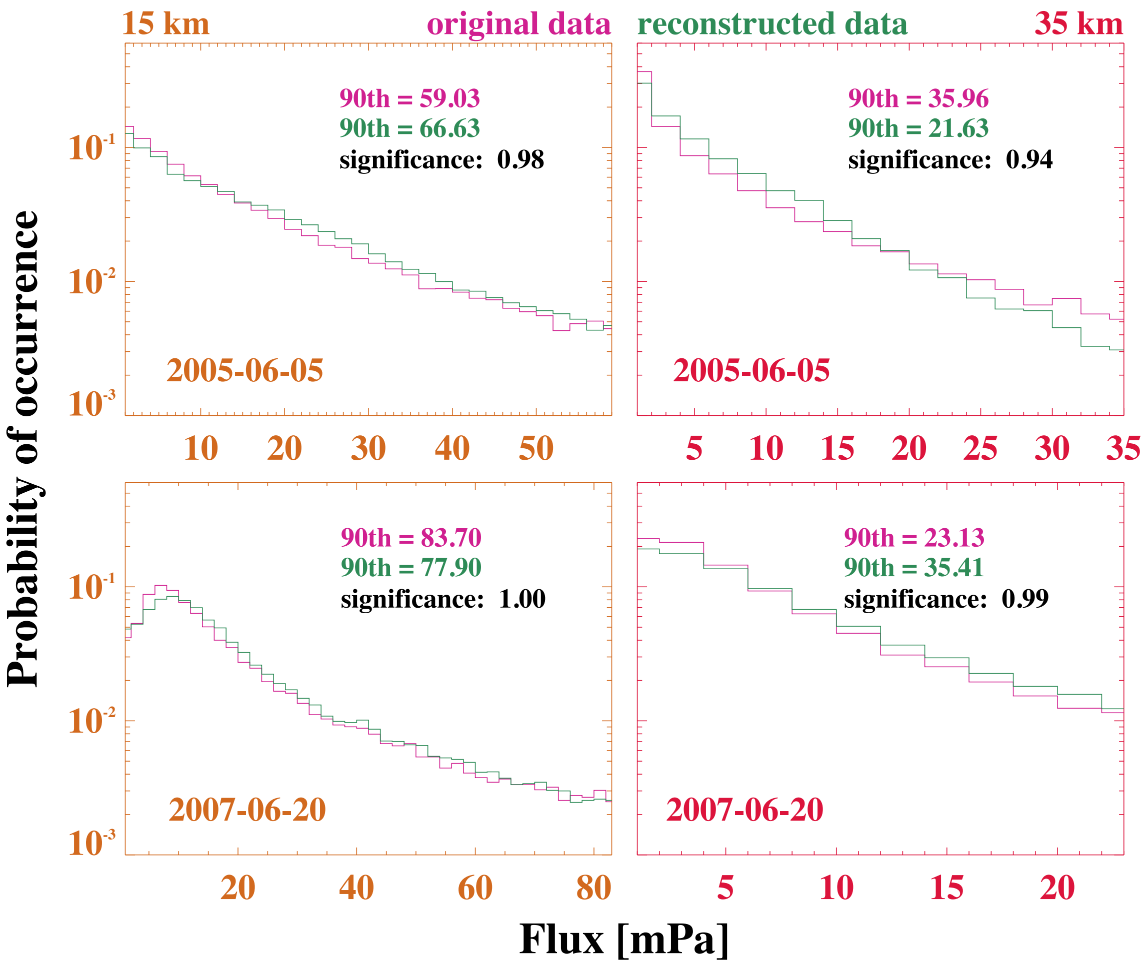


FIG. 4. Probability distributions of simulated 100 km×100 km instantaneous total flux amplitude at 15 km (left) and 35 km (right), for the squall line case (top) and the mesoscale convective complex (bottom), comparing simulations based on the original 10 min data (purple) and the reconstructed data (green). Values of momentum flux up to the 90th percentile are displayed.

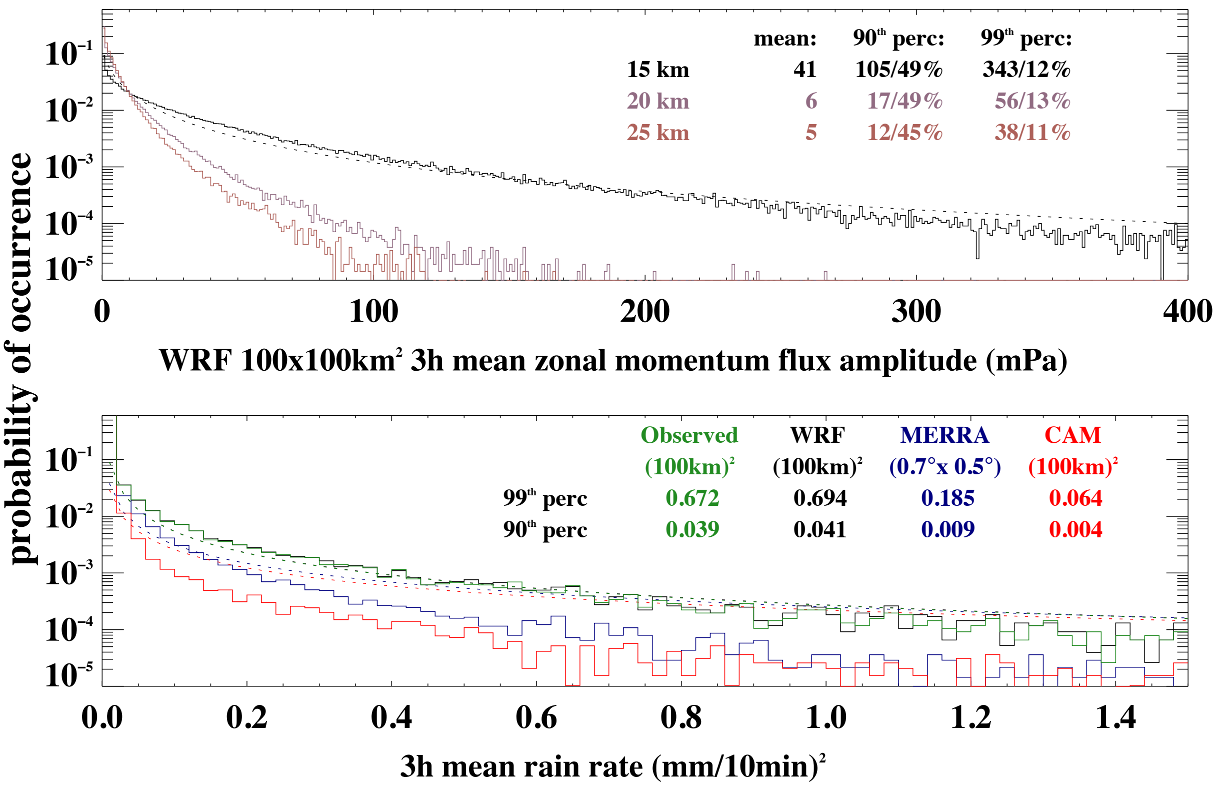


FIG. 5. Combining data from all simulations, the top panel shows probability density functions of absolute zonal momentum flux amplitudes averaged over 100 km×100 km and 3 h for different altitudes. The mean value, 90th and 99th percentiles as well as the percentages of flux associated with values larger than the percentiles are also indicated. The black dashed line is a lognormal distribution with the same mean and standard deviation as the distribution at 15 km. The bottom panel displays probability density functions of 3 h squared precipitation for the Stage IV data (green), the data used to force the WRF model (black), MERRA (blue) and CAM (red). The respective horizontal resolutions are indicated. Dashed lines are again lognormal distributions with the same mean and standard deviation as the data.

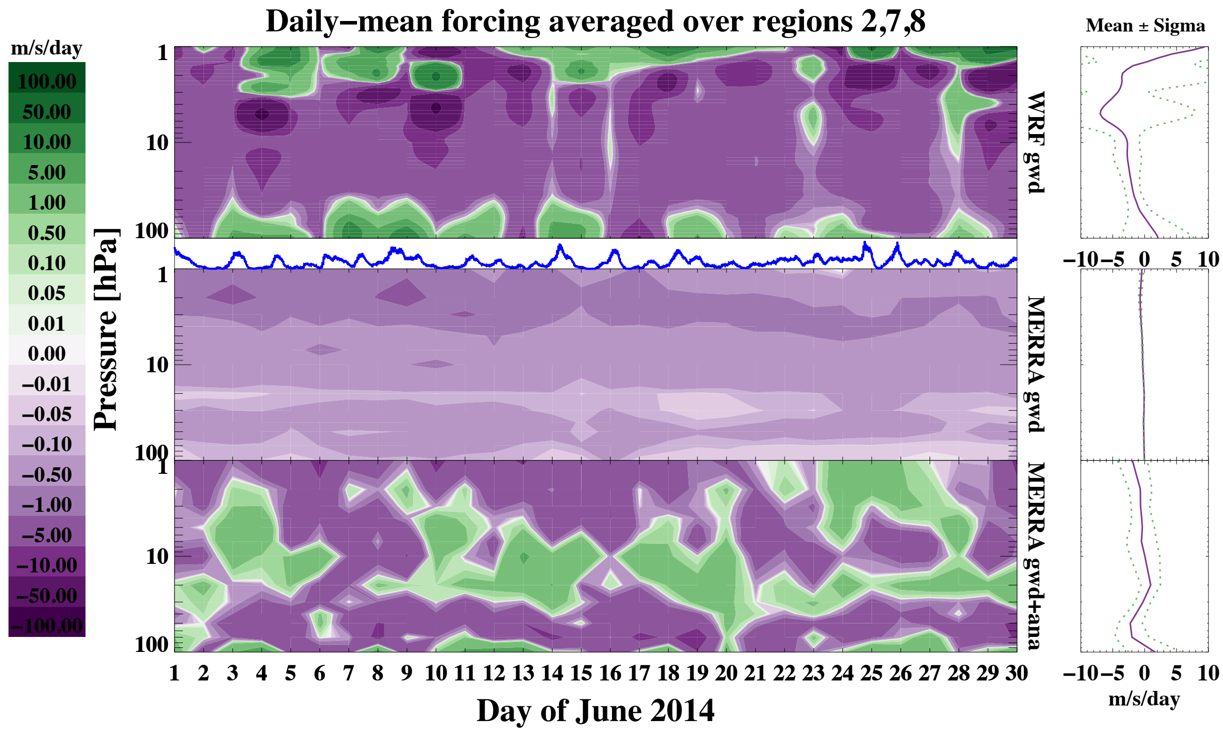


Fig. 6. Daily mean zonal wind tendencies for simulated GW drag (top), MERRA GW drag (middle) and MERRA GW drag plus analysis increments. The blue line below the top panel shows the time evolution of precipitation and the panels on the right show the monthly mean forcing (solid purple line) plus/minus one standard deviation (dotted green lines).

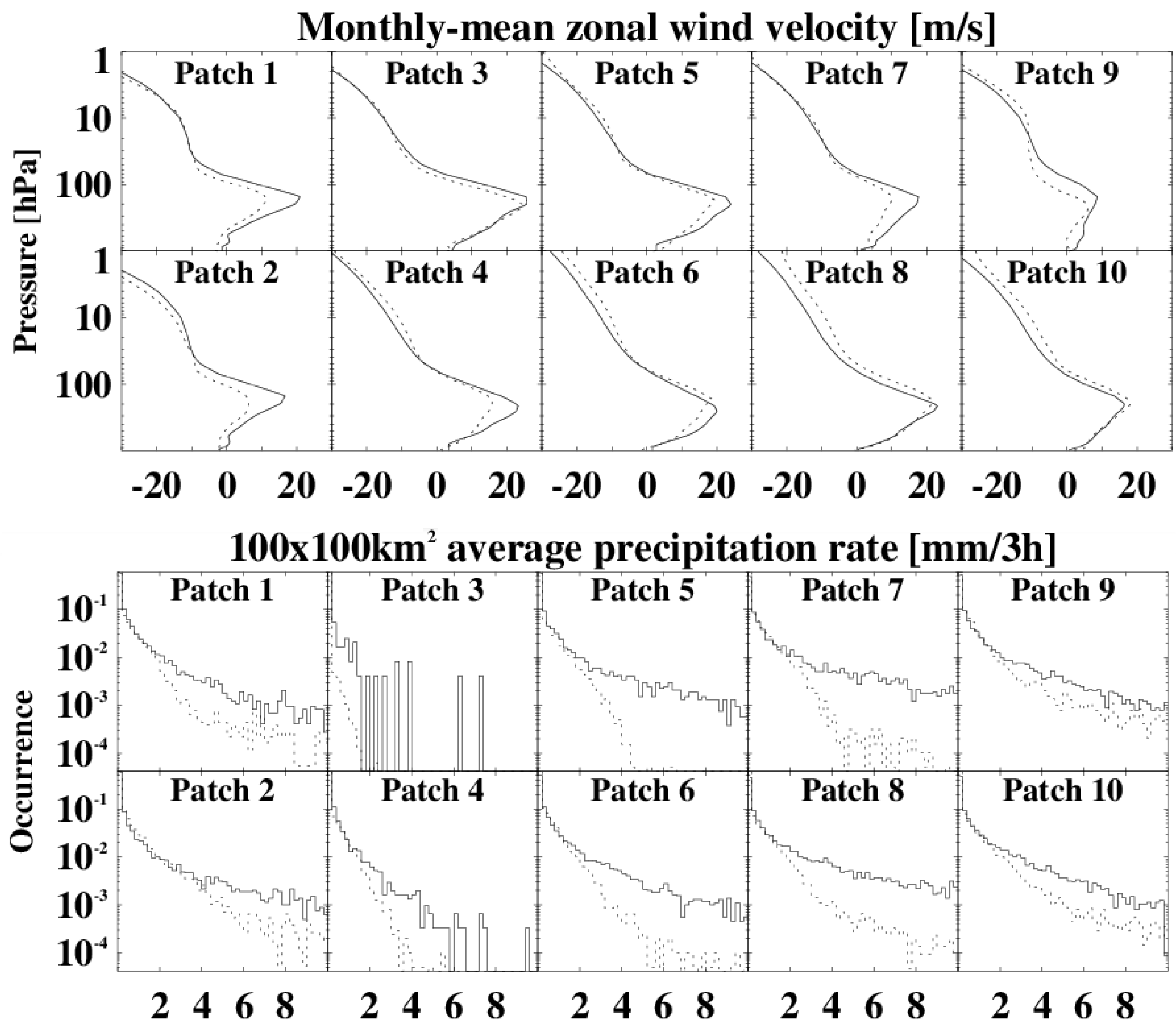


FIG. 7. For each domain June monthly-mean zonal wind and monthly mean 100 km $\times$ 100 km average precipitation strength distributions are shown. WRF data are solid lines and the corresponding values from the best matching CAM5 year are dashed lines.

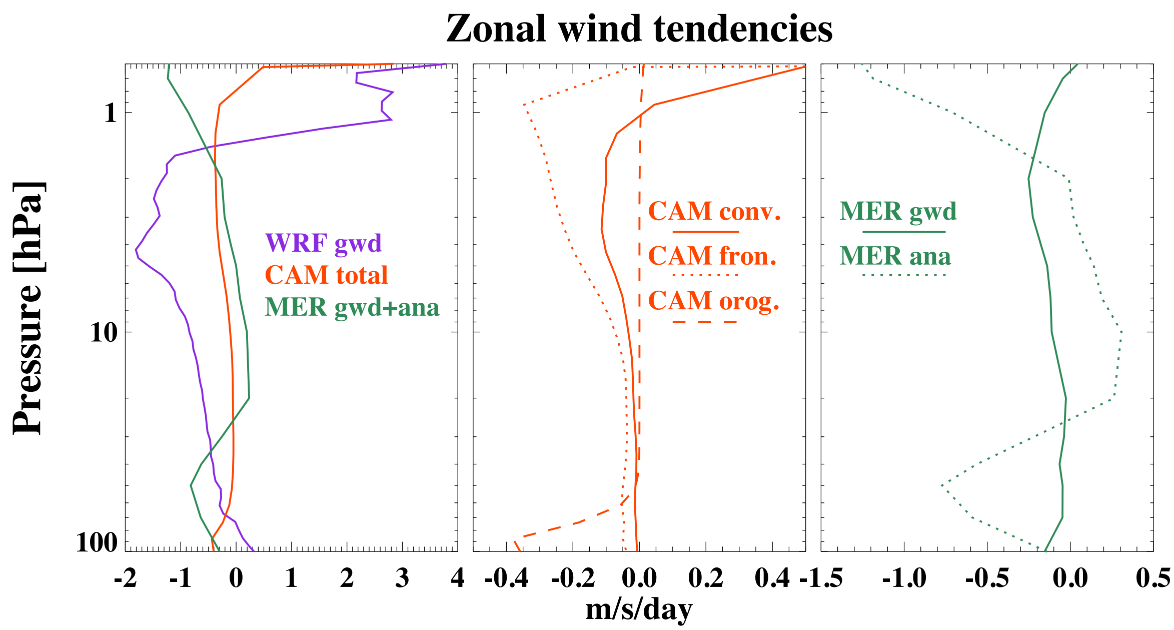


FIG. 8. Left panel: Zonal wind tendencies averaged over June 2014 and the area covered by the ten model domains for the WRF simulations (purple), all CAM5 GW drag schemes combined (orange) and MERRA GW drag plus analysis corrections (green). Middle panel: convective, frontal and orographic CAM5 tendencies. Right panel: MERRA GW drag and analysis increments. The values for the CAM5 model are composed of different years of simulations as described in the text.