Atmospheric Gravity Waves: Processes and Parameterization

ULRICH ACHATZ, M. JOAN ALEXANDER, ERICH BECKER, HYE-YEONG CHUN, ANDREAS DÖRNBRAK, LAURA HOLT, RIWAL PLOUGONVEN, INNA POLICHTCHOUK, KAORU SATO, ADITI SHESHADRI, CLAUDIA CHRISTINE STEPHAN, ANNELIZE VAN NIEKERK, AND CORWIN J. WRIGHT

a Goethe Universität Frankfurt, Frankfurt, Germany
b NorthWest Research Associates, Boulder, Colorado
c Yonsei University, Seoul, South Korea
d Institut für Physik der Atmosphäre, Deutsches Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen, Germany
e LMD/IPSL, Ecole Polytechnique, IP Paris, Sorbonne Université, Ecole Normale Supérieure, CNRS, Palaiseau, France
f European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom
g Department of Earth and Planetary Science, The University of Tokyo, Tokyo, Japan
h Department of Earth System Science, Stanford University, Stanford, California
i Max Planck Institute for Meteorology, Hamburg, Germany
j Centre for Space, Atmospheric and Oceanic Science, University of Bath, Bath, United Kingdom

(Manuscript received 25 May 2022, in final form 13 November 2023, accepted 13 November 2023)

ABSTRACT: Atmospheric predictability from subseasonal to seasonal time scales and climate variability are both influenced critically by gravity waves (GW). The quality of regional and global numerical models relies on thorough understanding of GW dynamics and its interplay with chemistry, precipitation, clouds, and climate across many scales. For the foreseeable future, GWs and many other relevant processes will remain partly unresolved, and models will continue to rely on parameterizations. Recent model intercomparisons and studies show that present-day GW parameterizations do not accurately represent GW processes. These shortcomings introduce uncertainties, among others, in predicting the effects of climate change on important modes of variability. However, the last decade has produced new data and advances in theoretical and numerical developments that promise to improve the situation. This review gives a survey of these developments, discusses the present status of GW parameterizations, and formulates recommendations on how to proceed from there.

KEYWORDS: Atmosphere; Gravity waves; Subgrid-scale processes; Climate models; Numerical weather prediction/forecasting

1. Motivation

Internal gravity waves (GW) are a basic wave mode of atmospheric motion that can occur wherever the atmosphere is stratified. They are a mesoscale phenomenon, with horizontal wavelengths ranging from below 1 km to a few thousand kilometers and vertical wavelengths ranging from below 100 m to several kilometers. Useful discussions of their dynamics and effects are given in many textbooks (e.g., Andrews et al. 1987; Sutherland 2010; Nappo 2012; Vallis 2017; Achatz 2022). They are emitted by various processes, for the most part from the troposphere. The most relevant of these source processes are thought to be flow over mountainous terrain, convection, and jets and fronts. Those may act simultaneously, and they are supplemented by various other instabilities and forcing mechanisms. Corresponding reviews have been given by Fritts and Alexander (2003), Kim et al. (2003), Alexander et al. (2010), Plougonven and Zhang (2014), and Sutherland et al. (2019). Upward-propagating GWs are filtered partially by the larger-scale flow. Their amplitudes typically increase, essentially due to energy conservation in an atmosphere where density decreases with altitude. They eventually induce static instabilities and shear instabilities leading to the onset of turbulence. The growth of turbulence draws on the GW energy. Hence it checks the amplitude growth of GWs, and the ensuing wave breaking is an important, though not the only, mechanism leading to a GW impact on the ambient flow.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Ulrich Achatz, achatz@iau.uni-frankfurt.de

DOI: 10.1175/JAS-D-23-0210.1

© 2024 American Meteorological Society. This published article is licensed under the terms of the default AMS reuse license. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
Many essential aspects of the atmospheric circulation and variability are critically affected or even controlled by GWs. There is a consensus that they have a leading-order influence on the mean circulation in the mesosphere and lower thermosphere (MLT) and that they also impact the tropospheric and stratospheric circulation. Beyond the atmospheric mean circulation, its seasonal cycle, and its climate variability on centennial time scales, GW impacts also extend to atmospheric variability and predictability from subseasonal to decadal time scales.

Hence it is essential that regional and global atmospheric models describe GWs and their effects with reliable accuracy. Due to the ongoing refinement of model resolutions, an increasing fraction of the GW spectrum can now be simulated explicitly. Yet, even the highest-resolution numerical weather predictions (NWP) still cannot accurately represent the full GW spectrum, so that they will depend on parameterizations for some time to come (Kruse et al. 2022; Polichtchouk et al. 2021; Sandu et al. 2019). The need for gravity wave parameterizations (GWP) is even more obvious in climate modeling. While centennial climate simulations with kilometer resolution in the horizontal, resolving a large part of the GW spectrum, might be possible within the next decade (Slingo et al. 2022), their computational costs will remain too high for routine and ensemble simulations. Moreover, the complexity of the dynamics described by kilometer-scale codes will defy systematic understanding unless a hierarchy of models with decreasing complexity is available (Held 2005). GWPs will remain essential in such a hierarchy, and their reliability can be considered to be a benchmark for conceptual understanding of GWs in the atmosphere.

GWPs as they are presently used in operational weather forecast and climate codes have a few common characteristics (e.g., Kim et al. 2003). The four essential aspects that are represented by them are GW emission in the troposphere, upward GW propagation, GW dissipation, and the GW impact on the resolved flow. 1) The emission process is usually subdivided between orographic emission, emission due to convection, and emission by jets and fronts. Separate schemes are typically in use for each of those. They all inject GWs at some, potentially variable, launch altitude in the troposphere. Those GWs have a phase-velocity spectrum that is either prescribed or may respond to the resolved flow and the parameterized convection. 2) GW propagation is treated as a boundary-value problem between the launch altitude and the model top (or the GW-breaking altitude), following linear theory. Given a vertical profile of winds and stratification (both assumed to be steady and horizontally invariant), one determines a vertical profile of GW amplitude. Important in this context are critical levels where GW horizontal phase velocity and ambient horizontal wind velocity agree. Within the above-mentioned assumptions GWs are not able to pass critical levels. Hence the latter filter the propagating GW phase-velocity spectrum. This is decisive for the direction of the momentum impact that GWs can exert once they dissipate. It explains, e.g., why GW breaking in the MLT leads to the climatological mean to a meridional circulation directed from the summer to the winter pole. 3) GW breaking is triggered by diagnosing static instability or shear instability. The former is invoked whenever a GW (or a GW spectrum) can locally lead to negative vertical derivatives of potential temperature. The latter is thought to occur when the GW field superposed on the resolved flow has vertical derivatives of potential temperature and winds locally leading to Richardson numbers below a critical value. Whatever the criterion used, whenever a potential GW instability is diagnosed, vertical GW-amplitude profiles are adjusted to remove the potential instability. 4) The GW impact on the resolved flow is represented in the horizontal momentum equations, by vertical derivatives of the GW vertical flux of horizontal momentum or the vertical GW pseudomomentum flux (often termed GW Eliassen–Palm flux). Within the above-mentioned assumptions these flux convergences can only change the resolved flow when GW dissipation occurs. The procedure sketched here is commonly adopted due in part to its simplicity and its computational efficiency. However, it also reflects a viewpoint on what the essential aspects of GW dynamics and the GW impact on the atmosphere are.

It appears that the reliability of operational GWPs, formulated along those lines, is not sufficient. This is especially apparent in studies investigating the effects of climate change on the tropical quasi-biennial oscillation (QBO). Schirber et al. (2015) and Richter et al. (2020) show that climate model predictions of the QBO in a warming climate critically depend on the choice and configuration of the GWP. Figure 1 illustrates the difference in the climate change response in different models, and how different GWPs, implemented even in the same model, result in vastly different predictions of how the QBO changes in a warmer climate. It is especially noteworthy that the differences in the launch-level phase-speed spectrum between the GWPs in the bottom two rows are relatively small. This highlights the remarkable sensitivity of the QBO to small changes in GWPs. Similar issues apply to the effect of GWPs on the Brewer–Dobson circulation (BDC; see below). Present-day GWPs seem to be overtuned, e.g., they are often configured to obtain a realistic circulation in the middle atmosphere while exhibiting inconsistencies with observations at lower altitudes (e.g., de la Cámara et al. 2016). Closely linked is the problem that GWPs are typically not scale aware, i.e., they must be tuned anew for each model resolution, and that there is no satisfactory exchange between resolved and parameterized GWs (van Niekerk et al. 2020; van Niekerk and Vosper 2021).

These issues provide an incentive to reconsider GWPs. Questions to be asked are as follows: 1) Do we capture the GW sources correctly? 2) Do we handle GW propagation in an appropriate way? 3) Is the treatment of GW dissipation sufficiently accurate? 4) Is the GW impact on the resolved flow formulated well? All of these questions can only be addressed by careful comparison between GWPs and corresponding results from GW-resolving data. The latter have traditionally been taken from measurements. They are increasingly being supplemented by kilometer-scale global simulations, by even more highly resolved regional simulations, and by idealized simulations resolving even the turbulent wave breaking process. Such simulation data have the potential
advantage of providing the full information on all dynamical fields on all modeled scales. They can only be trusted, however, if verified against observational data, so that measurements nowadays have the twofold purpose of validating GWPs directly and of validating GW-process-resolving simulations. All of the questions 1–4 are to be amended by the question of 5) how theoretical and numerical developments can help us to close consistency gaps between GWPs and validation data.

The intention of this review is to help the reader acquiring a basis of knowledge helpful in the assessment, application, and development of GWPs. To this end it is to give a short list of GW effects that justify the development and use of GWPs. Furthermore it shall give an overview over available benchmarks provided by our knowledge of GW processes and effects, by atmospheric measurements, and by high-resolution simulations. Finally it also is to give an account of the status and of recent developments in the formulation of GWPs. The focus is to be on progress and innovations from the last decade, supplementing previous reviews. Within this framework we summarize GW effects in section 2, discuss observational and model data in sections 3 and 4, sketch the most critical processes affecting GWs in section 5, discuss the state of GWPs in section 6, and summarize and evaluate the whole in section 7.

Fig. 1. Sensitivity of the QBO to the formulation of the GWP. (a) In the left column the QBO as simulated by two climate models and as seen in analysis data (in the lowermost row). The center and right columns show the response to double and quadruple anthropogenic climate gas forcing. (b) The impact of two different GWPs in the same climate model on the QBO for the present climate (in the left column) and in climate change experiments (in the right column). Both the period and the amplitude of the QBO are modified in very different ways in the climate change simulations depending on the GWP. For further details on (a) and (b) see Richter et al. (2020) and Schirber et al. (2015), respectively.
2. Gravity wave effects in the atmosphere

Small-scale GWs decelerate or accelerate the large-scale flow by elastic forcing and momentum-flux convergence, and to a smaller degree also heat or cool it, by entropy-flux convergence and indirectly also by turbulent dissipation. Typically, they can influence the large-scale flow only either if they break and dissipate or if their amplitudes vary significantly (e.g., Andrews et al. 1987; Achatz 2022). We summarize here what has been learned about GW impacts in general, with a focus on new findings from the last decade. First the seasonal-mean circulation and its variability will be touched upon, moving from higher altitudes in the mesosphere, where GW effects are typically strongest, downward to the troposphere. Then GW effects on atmospheric variability in general will be discussed, again proceeding from high to low altitudes.

Beginning with the seasonal-mean circulation, the first incentive for the development of GWPs for climate models was the effect of GWs on the MLT, where they cause an interhemispheric circulation (the mesospheric counterpart of the BDC) directed from the summer pole to the winter pole, thereby cooling the upper summer mesosphere, heating the lower winter mesosphere, and reversing the jet-stream direction across the mesopause (Lindzen 1981; Matsuno 1982; Holton 1982, 1983; Garcia and Solomon 1985). From even the first attempts of GWP, these effects were well represented.

Recent studies indicate that GWs also contribute significantly to the BDC in the stratosphere, e.g., its latitudinal extension and turn-around latitude (Sato and Hirano 2019), with consequences for the exchange of momentum and constituents between the stratosphere and troposphere (McLandress and Shepherd 2009; Abalos et al. 2015). It even appears that global-model responses of the BDC to a warming climate cannot be understood without the role of GWPs (Butchart 2014). Yet, they also pose an issue to our understanding: The BDC response, and hence climate predictions in general, are subject to considerable uncertainties. Climate models consistently predict an acceleration of the BDC due to climate change. However, the predicted strength varies considerably among models, and the mechanisms driving these changes are still unclear (Eichinger et al. 2019). The limited observational evidence on the trend of the BDC even seems to suggest that it is decelerating (Engel et al. 2008; Abalos et al. 2021). All of this indicates considerable uncertainty in the GWPs presently in use.

Roughly in parallel to their implementation into climate models, GWPs have also been introduced into NWP. It had been noticed in Northern Hemisphere wintertime that orographically excited GWs induce in the upper troposphere and lower stratosphere (UTLS) a circulation that contributes to an increase of zonal-mean temperature and a deceleration of the stratospheric polar night jet (PNJ), even leaving an imprint on surface pressure (Palmer et al. 1986; McFarlane 1987; Lott and Miller 1997). Although GWPs are able to represent this basic effect, they still have problems capturing it sufficiently well (Sandu et al. 2019)—for example, because of a lack of adaptivity to grid resolution (van Niekerk and Vosper 2021).

Effects of GWs on atmospheric variability, in the extratropics most often accompanied by those of Rossby waves, are numerous, and often intertwined mutually. A summary overview of the possible interactions between GWs, Rossby waves, and mean flow is given by Fig. 2. To begin with the mesosphere, model studies using a GWP indicate that variable GW forcing contributes there to leading order to an observed subseasonal variability of the summer-to-winter pole circulation (Becker and Fritts 2006; Gumbel and Karlsson 2011; Karlsson et al. 2007; Körnich and Becker 2010; Karlsson and Becker 2016). However, it is a sign of the still significant uncertainties in GWPs that the leading role of GWs is questioned not only by observational studies (France et al. 2018; Lieberman et al. 2021) but also by other model studies using a GWP as well (Smith et al. 2020, 2022; Yasui et al. 2021).

Numerous further variability modes are observed in the tropics, that is, the mesospheric QBO, the mesospheric semiannual oscillations, and mesospheric intraseasonal oscillations, that are thought to be driven to a large part by GWs. Recent reanalysis and satellite observations at least partly capture these phenomena (Smith et al. 2017; Kawatani et al. 2020; Koshin et al. 2022), but they have not been used sufficiently to test GWPs.

Moreover, GW forcing can establish potential vorticity distributions leading, via baroclinic instability, to the growth of Rossby waves that contribute to extratropical mesospheric variability. This is also seen in GW-permitting global simulations (Watanabe et al. 2009), observations (Matthias and Ern 2018), and in global simulations using a GWP (Sato et al. 2018). A complex interplay with Rossby waves in modifying the mean field is diagnosed, in which GW propagation is directly and indirectly modulated by Rossby waves (Sato and Nomoto 2015; Okui et al. 2021).

On daily and shorter time scales, solar tides modulate GW propagation, but the hence modulated GW forcing also affects the tides (Senf and Achatz 2011). Finally, secondary GWs forced by GW dissipation (Vadas et al. 2003, 2018) or by transient GW forcing (Tabaei and Akylas 2007; van den Bremer and Sutherland 2018; Wilhelm et al. 2018) are often very large in scale. Hence, secondary GWs are also part of the larger-scale flow. Modern GWPs, as discussed in section 6, have been designed to represent such effects, but their corresponding efficacy remains to be tested.

In the stratosphere, GW effects on the variability can be expressed in terms of modulating global atmospheric teleconnections. Since the latter play a major role in the predictability of surface weather and climate (e.g., Scaife et al. 2022; Baldwin et al. 2019), an accurate representation of GW effects is needed both for NWP and climate modeling. Figure 3 illustrates some of the troposphere–stratosphere teleconnections and the interacting GWs from various sources.

In the tropics (Fig. 3, label 1), GWs excited by deep convection are the dominant drivers of the QBO (e.g., Kawatani et al. 2010; Kim and Chun 2015; Scaife et al. 2000). The QBO influences the troposphere directly by modulating tropical tropopause temperature, wind, and deep convection associated with the Madden–Julian oscillation (MJO) (e.g., Martin et al. 2021) or indirectly by modulating the strength of the subtropical
jet and planetary-scale waves that affect the Arctic polar vortex (Anstey and Shepherd 2014, Fig. 3, labels 2 and 3). As outlined above, operational GWPs do not represent the GW impact on the QBO in a sufficiently reliable manner (Fig. 1, Schirber et al. 2015; Richter et al. 2020).

In the Northern Hemisphere high latitudes (Fig. 3, label 2), GWs excited by orography, jets, and fronts can break in the stratosphere, thereby impacting the strength of the Arctic polar vortex and the frequency of sudden stratospheric warming events (SSW), as diagnosed from a GWP in a high-resolution atmospheric model (Polichtchouk et al. 2018a). SSWs are sometimes followed by the formation of an elevated stratospheric warming event at $z = 80$ km, as seen in observations, in global simulations using a GWP, and in global GW-permitting simulations (Manney et al. 2008; Chandran et al. 2013; Limpasuvan et al. 2016; Stephan et al. 2020; Okui et al. 2021). The GW impact on SSW is relevant for tropospheric predictability because there is a well-documented link (Baldwin and Dunkerton 2001) between the strength of the polar vortex and the Arctic Oscillation (AO), including the North Atlantic Oscillation (NAO). Similarly, in the Southern Hemisphere (Fig. 3, label 4), GWs from sources that are debated (e.g., Garcia et al. 2017; Hindley et al. 2019) contribute to the timing of the springtime polar vortex breakdown (e.g., Garcia et al. 2017; Polichtchouk et al. 2018b; Gupta et al. 2021). This has an influence on the seasonal evolution of the tropospheric eddy-driven jet and the associated temperature and precipitation anomalies (e.g., Polichtchouk et al. 2018a; Lim et al. 2018;
While most of the listed processes are generated by tropical cyclones can enhance the intensity of the lat-
and Str"acker 1990). Especially, convective GWs gen-
are observational indications that ice clouds in the
Lat. scale (SGS) orographic GWs on rainfall requires a parameteri-
Byrne et al. 2019). Also affecting the latitude of the eddy-
driven jets in both hemispheres is the GW-driven deceleration of the subtropical jets in the lowermost extratropical strato-
sphere (e.g., van Niekerk et al. 2020; Chen and Zurita-Gotor 2008; Fig. 3, label 3). Again, this is a result that so far has been obtained using GWPs and that would merit validation by observations and high-resolution global models.

Directly in the troposphere GWs promote the growth of shallow convection into banded clouds (Shige and Satomura 2001), provide a mechanism for the longevity of convective systems (Rotunno et al. 1988; Lin and Chun 1991; Stechmann and Majda 2009), or initiate severe convection (Uccellini 1975). The effects of GWs on convective organization range from scales of a few kilometers (Clark et al. 1986; Balaji and Clark 1988; Lane and Clark 2002) to mesoscales (e.g., Ruppert et al. 2022). Convection generates discrete wave modes that modulate their mesoscale environment through wave-induced ascent and descent, respectively (Mapes 1993; Stephan et al. 2016), which may then influence convective systems in both the tropics (Mapes et al. 2003; Lane and Zhang 2011; Stephan et al. 2021) and midlatitudes (Koch and Siedlarz 1999; Ferretti et al. 1988; Schneider 1990). Especially, convective GWs generated by tropical cyclones can enhance the intensity of the latter through GW divergence in the UTLS (Kim and Chun 2011; Kim et al. 2014). While most of the listed processes are resolved by weather forecast models, the impact of subgrid-scale (SGS) orographic GWs on rainfall requires a parameterization, addressed, for example, by Smith et al. (2015). Last, there are observational indications that ice clouds in the UTLS are significantly influenced by GW dynamics (Kärcher and Ström 2003; Kim et al. 2016; Bramberger et al. 2022). Corresponding parameterizations have been proposed by Joos et al. (2008) and Weimer et al. (2023). This is of relevance for climate modeling because such clouds still represent an important uncertainty in the radiation balance of the atmosphere.

3. Constraints from observations of atmospheric gravity waves

GW observations provide vital benchmarks both for GWPs and for wave-permitting high-resolution simulations. Yet, observational studies of GWs are inherently difficult to carry out, due to the observational filter problem (e.g., Alexander 1998; Preusse et al. 2002; Alexander et al. 2010). Briefly, no single observing instrument can see all spatiotemporal scales of GWs at a given time and location. Diverse observations across a range of averaging and sampling scales are required. In the last two decades, however, significant progress has been made in covering almost all parts of the GW spectrum by, inter alia, combining different sensors (e.g., Alexander 2015; Wright et al. 2016a,b; Vincent and Alexander 2020; Ern et al. 2021). This section is to give a short survey of available measurements and their potential contributions to the required constraints.

Global distributions of GW properties are only available via the view from space provided by satellites. By now the combination of specialist retrievals (Hoffmann and Alexander 2009) and increased computer power has led to global studies able to use 3D satellite temperature fields from nadir-sensing radiometers such as AIRS (Ern et al. 2017; Wright et al. 2017; Hindley et al. 2020) and IASI (Hoffmann et al. 2014). These instruments, with their wide fields of view allowing near-global daily coverage at extratropical latitudes, have also provided near-real-time data, supporting both operational planning of GW measurement campaigns (Eckermann et al. 2019) and rapid analysis of extreme GW activity associated with events such as sudden stratospheric warmings (e.g., Dörnbrack et al. 2018) and the 2022 Hunga Tonga eruption (Wright et al. 2022). The AIRS record now spans over two decades, and climatological stratospheric GW momentum fluxes (GWMF) have been derived by Hindley et al. (2020). Figure 4 shows 18-yr averages of stratospheric wintertime directional GWMFs in both hemispheres. Limb sounders, such as SABER and MLS, and other high-vertical resolution satellite instruments such as the COSMIC and COSMIC-2 constellations have also proven important for characterizing key features of the statistical behavior of GWs. Relevant among those are on the one hand global climatological distributions of GW temperature amplitudes and GWMFs (e.g., Preusse et al. 2009; Hindley et al. 2020). On the other hand analyses of these satellite data have also provided maps of Gini coefficients of GWMFs (Wright et al. 2013, 2017; Ern et al. 2022). Those inform us about the intermittency of GWMFs, that is, the contribution of comparatively rare events with strong GWMFs. Such results pose important tests to high-resolution models and GWPs.

While satellites provide global information, their spatial accuracy is limited by their comparatively wide field of view. More accurate local information with better vertical resolution is given by ground-based or in situ measurements. In
recent years, the capabilities of middle atmosphere observa-
tions of temperatures, but increasingly also of winds, by
ground-based lidars have gradually improved. Longer obser-
vation periods due to advanced technologies were achieved:
for example, there are lidars capable of measuring day and
night (Baumgarten et al. 2017) and lidars operating autono-
mously at remote locations (Kaißer and Kaißer 2021). The
vertical coverage of middle atmosphere temperature measure-
ments extends by now from 20- to 70-km altitude (Strelnikova
et al. 2021) and even to approximately 100-km altitude (Alpers
et al. 2004; Reichert et al. 2021). All groups are now able to
study the seasonality of GW activity in the middle atmosphere
with reliable confi-
dence. The annual cycle peaks outside of the
tropics in the respective winter months. Furthermore, as illus-
trated in Fig. 5, hodograph analyses of simultaneous wind and
temperature measurements indicate a substantial contribution
of downward-propagating GWs to the total spectrum that
could be explained by wave reflection and by emission pro-
cesses in the middle atmosphere (Strelnikova et al. 2020). It
still is a challenge to wave-resolving models and GWPs to rep-
resent the measured seasonally dependent vertical profiles of
wind and temperature variations. Moreover, none of the
GWPs operational in global models represent downward-
propagating GWs. Finally, as the ground-based lidar data are
not assimilated in the forecast cycle of NWP centers, they con-
stitute a valuable and independent source for validating the
GW-permitting predictions by the latter (e.g., Le Pichon et al.
2015; Ehard et al. 2018; Strelnikova et al. 2021; Gisinger et al.
2022). For example, ECMWF’s IFS captures the seasonal cycle
of GW activity well, but underestimates amplitudes at alti-
tudes higher than about 40–50 km due to explicit and numeri-
cal damping. It also appears that a considerable part of the
spectrum is not resolved even by using resolutions as fine as
about 1 km (Polichtchouk et al. 2021; section 4 below).
Nonetheless, the often remarkable agreement of wave structures in the IFS simulations and in the lidar observations (e.g., Fig. 6, showing an example of good alignment between lidar backscatter and IFS temperatures) indicates that finer resolution and increasing realism of operational NWP model outputs do offer a valuable quantitative source for mesoscale flow components, which were hitherto not accessible globally (e.g., Dörnbrack et al. 2017; Kaißer et al. 2020) and that can be used as benchmarks for GWPs.

Supplementing lidar, mesosphere–stratosphere–troposphere (MST) radars continuously observe vertical profiles of three-dimensional winds in the troposphere, lower stratosphere, and mesosphere with high accuracy, vertical and temporal resolution, and can capture GWs in frequency and wavenumber space different from satellites. Recent efforts have led to the establishment of a sparse but global radar network that includes the Arctic and Antarctic (Latteck et al. 2012; Sato et al. 2014). Although MST radar observations in the mesosphere are only possible during daylight hours when atmospheric ionization is strong, continuous observations over several dozen days in the polar regions in summer provide GWMFs over a wide frequency range (Sato et al. 2017). In addition, meteor radar observations have been widely used to measure horizontal winds and atmospheric waves in the MLT inter alia to determine propagation conditions for GWs (e.g., Hocking et al. 1997; Jacobi et al. 2007; McCormack et al. 2017; de Wit et al. 2017; Stober et al. 2018). Hence, radar data as well provide benchmarks that GW modeling will have to pass.

Moreover, instrumented aircraft continue to be a valuable source for the determination of the vertical flux of horizontal momentum. A central result of the DEEPWAVE campaign (Fritts et al. 2016) is the spectral breadth and the different shapes of horizontal and vertical wind spectra measured in the lower stratosphere (Smith and Kruse 2017). Dividing the airborne observations into a longwave and a shortwave part, with horizontal wavelength shorter and longer than 60 km, respectively, revealed that only one-third of the total momentum flux is carried by the longer waves. Hence a significant portion of GWs are not explicitly resolved by most models and need to be described by GWPs. Nonetheless it is often the longer waves that control where and when the shorter waves dissipate and deposit momentum. This modulational and filtering process is hard to capture in the nontransient approach taken by presently operational GWPs, where equilibrium profiles of GW amplitudes are determined on an assumedly steady resolved flow. During DEEPWAVE, a unique combination of upward-pointing remote sensing instruments (Rayleigh and sodium lidars and an Advanced Mesosphere Temperature Mapper) extended the range of the airborne observations into the middle atmosphere (e.g., Bossert et al. 2015; Eckermann et al. 2016; Taylor et al. 2019; Pautet et al. 2019; Dörnbrack et al. 2022). The majority of measurements in the middle atmosphere report large values of vertical momentum flux carried by the shorter waves (<100 km). Similar measurements were conducted recently over the southern Andes (Rapp et al. 2021) in the SOUTHTRAC campaign.
Finally, in addition to high-resolution radiosonde profiles (e.g., Geller and Gong 2010; Okui and Sato 2020), observations from superpressure balloons, deployed in both commercial contexts (Lindgren et al. 2020) and scientific campaigns, have provided valuable in situ quantification of GWS in the lower stratosphere. Their quasi-Lagrangian behavior allows mapping measured fluctuations to frequencies in the GW spectrum and to quantify momentum fluxes and induced temperature fluctuations (Podglajen et al. 2016; Schoebert et al. 2017; Haase et al. 2018; Vincent and Alexander 2020; Bramberger et al. 2022). These measurements are considered to be particularly accurate, hence an especially important benchmark to GWPs, and they have pioneered the investigation of GWMF intermittency. Most recently they have highlighted the role of convection in the tropics (Corcos et al. 2021; Köhler et al. 2023) and the peak of GW-fluctuation amplitudes at the inertial frequency (Conway et al. 2019; Podglajen et al. 2016).

4. Gravity wave–permitting and other process-resolving simulations

The last decade has seen many global simulations with horizontal grid spacing in the range of 1–10 km. Model simulations at these resolutions are in the “gray zone” between fully resolving the relevant small-scale processes and parameterizing them. Even at these GW-permitting horizontal resolutions, not all of the GW spectrum is resolved. Already the “effective resolution” of a model—for example, about 8 km for a model with 1-km horizontal resolution—prevents the accurate representation of smaller-scale waves. Even so, high-resolution simulations of GWS have displayed an unprecedented degree of realism both in regional models (Grimsdell et al. 2010; Orr et al. 2015; Vosper 2015; Stephan and Alexander 2015; Stephan et al. 2016; Hindley et al. 2021; Kruse et al. 2022) and global models (Holt et al. 2017; Müller et al. 2018; Stephan et al. 2019b, a; Polichtchouk et al. 2021, 2023; Okui et al. 2021; Wei et al. 2022; Sato et al. 2023). More than ever, GW-permitting simulations provide the opportunity to estimate the full spatially and temporally varying GWMFs, and to verify the output of GWPs.

A significant challenge lies, however, in validating the GW-permitting simulations against the measurements sketched in section 3. Estimating the GWMFs and, especially, the resulting forcing from observations is still an unsettled issue (Alexander and Sato 2015), and different wave analysis methods give very different estimates (Geller et al. 2013). Moreover, there is no standard method of extracting GWS in models and comparing them with GW observations. Kruse and Smith (2015) point to a possible approach. Another that might help circumventing some of the issues is “measuring” GWs in the model, that is, applying exactly the same analysis methods to model data as to observational data (e.g., Okui et al. 2023).

Another issue is that simulated GWs and corresponding fluxes continue to be very sensitive to the chosen model configuration. The often good agreement of single wave events between measurements and high-resolution simulations (see section 3) indicates that wave propagation tends to be simulated well, but horizontal and vertical resolution still have a strong influence on the total GW statistics. In general, an increase in horizontal resolution leads to an increase in GW forcing in global models (Holt et al. 2016, 2020; Polichtchouk et al. 2023), but they appear to be even more sensitive to an increase in vertical resolution (Vosper 2015; Watanabe et al. 2015; Holt et al. 2016, 2020). The importance of vertical resolution for the representation of tropical waves has been known for some time (e.g., Boville and Randel 1992; Hamilton et al. 1999; Giorgetta et al. 2002; Richter et al. 2014; Anstey et al. 2016), but it seems that it is also important for improving temperature biases and SSW in the winter hemisphere (Sato et al. 2012; Wicker et al. 2023). There is some evidence from the ECMWF model (with deep-convection parameterization switched off) that the amplitude of the resolved GW forcing in the tropics converges as horizontal grid spacing decreases from 10 to 1 km (Polichtchouk et al. 2021), but the scales contributing to that forcing do not converge. Agreement with simulations from other models would be needed to draw stronger conclusions. Figure 7 shows that outside of the tropics there is still no evidence of convergence: as horizontal grid spacing decreases, GW forcing increases over both nonorographic and orographic regions (Polichtchouk et al. 2023). The same is found in simulations by limited-area models over various...
orographic hotspots (van Niekerk et al. 2020; van Niekerk and Vosper 2021; Vosper et al. 2020; Kruse et al. 2022). This suggests that even at 1 km horizontal grid spacing a portion of the GW spectrum is still unresolved. Because GWs close to the grid scale can produce significant fluxes, even the numerical scheme can affect the simulation results (Yao and Jablonowski 2015). This still missing resolution convergence presents an important caveat to the validation of GWPs against corresponding simulations. It appears that it is not only caused by the underrepresentation of the GW field itself but rather also by the unreliable representation of SGS processes generating or affecting GWs. There is a contribution from SGS orography, but also SGS vertical mixing (Malardel and Wedi 2016), moisture (Wei and Zhang 2014), and convection (Kim et al. 2007; Preusse et al. 2014; Stephan et al. 2019a; Polichtchouk et al. 2021) play a role.

As an important example, even at kilometer-scale horizontal meshes, individual convective cells are not fully resolved. The treatment of convection in GW-permitting models is not straightforward, and models often use a combination of explicit and parameterized convection. However, resolved GWs respond differently to explicit and parameterized convection.

As also illustrated in Fig. 8, zonal mean GW momentum fluxes in the tropics and subtropics are drastically lower in simulations with parameterized convection than in simulations with explicit convection (Müller et al. 2018; Stephan et al. 2019a; Polichtchouk et al. 2021).

Another SGS process of relevance for GW life cycles in high-resolution models is the interaction between GWs and turbulence. This affects global GW-permitting models encompassing the full MLT where GWs break at the latest in their life cycle (e.g., Borchert et al. 2019; Stephan et al. 2020; Becker et al. 2022). It even matters, however, for GW simulations by lower-top models with kilometer-scale resolution: They show a clear dependence on the divergence damping employed (Holt et al. 2016). Moreover, boundary layer turbulence is known to affect the emission of GWs from flow over...
mountainous terrain (e.g., Lott and Miller 1997). Large-eddy simulations (LES) and direct numerical simulations (DNS) can provide useful data for the advancement of turbulence parameterizations in interaction with GWs, but open issues remain. By validation of LES against DNS of breaking GWs, Remmler et al. (2015) show that a dynamic Smagorinsky approach (Germano et al. 1991; Lilly 1992) is well able to parameterize, at resolutions not too coarse, the turbulence arising there. However, using such a parameterization, Fritts et al. (2022a) find considerable deviations between mountain-wave simulations with either 500-m or 1- and 2-km grid spacing. Validations against and/or the advancement based on such LES, of the turbulence parameterizations used in global GW-permitting models would be desirable. DNS of breaking GWs or of turbulence arising in corresponding Kelvin–Helmholtz instabilities, now possible at extremely fine resolutions (Fritts et al. 2022b,c), could also provide useful data for the further advancement of LES and for the description of GW breaking in GWPs, but such an application still seems to be in its infancy.

In summary, DNS, LES, regional and global GW-permitting simulations have increased our understanding of GWs. Caveats apply to the missing convergence of GW-permitting global simulations. Further improvements will require progress in the representation of GW-affecting SGS processes by reliable parameterizations.

5. New lessons on processes affecting gravity waves

Observations and simulations have been vital support to studies on processes affecting the GW field. The question is which of those contribute to a degree that they should be better taken into account in GWPs. Many investigations have focused on sources. GW source spectra and amplitudes remain poorly constrained and are presently tuned in climate models (Alexander et al. 2010). Improved knowledge of the atmospheric GW field has however also contributed to the identification of other discrepancies between the life cycle of real waves and their parameterized counterparts. The present section highlights the most relevant for GWPs.

To begin with, while a first approximation of atmospheric sources can be obtained from a linear model (e.g., convective waves proportional to the diabatic heating in convective clouds), process studies (Chun et al. 2008) emphasize that emitted GW momentum fluxes have a nonlinear relation to the forcings. As also discussed in section 6, poor representation of intermittency of parameterized sources further contributes to errors in parameterized GW strengths and the GW intermittency that is well known from observations (Hertzog et al. 2012; Plougonven et al. 2013; Wright et al. 2013; Minamihara et al. 2020).

Second, while present GWPs assume that each GW is emitted just by a single source mechanism, GW emission often results from various processes together. Case studies from several campaigns have repeatedly emphasized the complexity of the observed wave field and the combination of several forcing processes, e.g., orography and imbalance (Ehard et al. 2017; Dörnbrack et al. 2018; Krisch et al. 2020; Geldenhuys et al. 2021; Strube et al. 2021) or jet/front dynamics and convection (Wei and Zhang 2014). In the latter case, nonlinearities profoundly affect the waves emitted from the jet/front source. For example, Fig. 9 shows that dry cyclogenesis emits considerably less GWs than if in interaction with moisture. Fundamental understanding of the generation of waves from balanced dynamics (Plougonven and Zhang 2014) has seen progress coming from theoretical, numerical, and experimental approaches (Yasuda et al. 2015; Rodda et al. 2019; Rodda and Harlander 2020), yet its atmospheric manifestation will involve interaction with other processes, for example, convection and precipitation (Plougonven et al. 2015; Holt et al. 2017). Moreover, convective GWs generated by tropical cyclones, visible in satellite observations, radiosondes, and numerical simulations (Hoffmann et al. 2018; Chane Ming et al. 2022; Kruse et al. 2022), are also subject to further investigations using ray-tracing models for the three-dimensional propagation of GW packets demonstrating a strong impact on the resulting momentum fluxes (Senf and Achatz 2011; see Fig. 10). This implies that interactions between the middle atmospheric flow and GWs are substantially influenced by lateral propagation, notably in the case of strongly perturbed flows such as SSW (Song et al. 2020; Stephan et al. 2020; Gupta et al. 2021). Present-day GWPs also do not take into account some nonhydrostatic propagation effects. Those allow, for example, short GWs to tunnel through the PNJ into the lower mesosphere where their breaking leads to significant momentum deposition (e.g., Mixa et al. 2021).

A fourth discrepancy is the restriction of sources to the troposphere: parameterizations launch GWs from the surface (for orographic waves) or the upper troposphere (for orographic waves). Yet it has been known for a long time that local instabilities (e.g., shear instabilities) could generate GWs (e.g., Scinocca and Ford 2000, and references therein), albeit weakly. Recent studies have revisited this generation, with current computing power pushed to resolve complex, multiscale interactions involved in such generation of waves or in the regeneration of waves (hence called secondary) when substantial turbulence results from the breaking of a primary wave (Fritts et al. 2015, 2018; Dong et al. 2020). Results from high-resolution global modeling suggest that secondary generation could play a significant role in the circulation at mesospheric heights (Becker and Vadas 2018). Other relevant stratospheric source processes could be GW generation from planetary-wave critical layers (Polichtchouk and Scott 2020), imbalances from the PNJ (Dörnbrack et al. 2018) or transient
GW packets (Tabaei and Akylas 2007; van den Bremer and Sutherland 2018; Wilhelm et al. 2018). The nonlinearities, the wide range of scales involved, and the remoteness of the atmospheric regions make investigations challenging, but a confirmation of the importance of such distributed sources could have profound implications for parameterizations. The possibility of distributed sources changes the range of phase speeds possibly occurring at higher altitudes (de Wit et al. 2017). The exponential growth of amplitudes with height implies that weak stratospheric sources might matter for the mesosphere, but Medvedev et al. (2023) conclude in their analysis to the contrary.

The discrepancies between the simplified descriptions of GWs in GWPs and their complex manifestation in the real atmosphere highlight the challenges in developing and constraining GWPs. In this context it seems also worthwhile to give attention to the spectral properties of the GW field. These have received renewed interest, both observationally as a result of novel approaches to separate the GW component from the total horizontal-wind energy spectrum (Callies et al. 2014; Lindborg 2015; and top panels of Fig. 11) and numerically as a result of higher resolutions (Stephan et al. 2022; and bottom panels of Fig. 11). As discussed by Bierdel et al. (2016) and Morfa and Stephan (2023), energy spectra can give clues on the relevance and nature of GWs on the mesoscale. In general, an indication of the physical realism of both GW-permitting models and GWPs is given by their ability to represent the corresponding observed spectra.

6. Gravity wave parameterizations: Status and possible solutions to shortcomings

The development of GWPs within the first three decades since their introduction into climate models (Lindzen 1981; Holton 1982, 1983; Garcia and Solomon 1985; McFarlane 1987) and into NWP (Palmer et al. 1986) has been discussed by Kim et al. (2003), and Alexander et al. (2010), and more recent assessments have also been given by Plougonven et al. (2020) and Kruse et al. (2023). In comparing present operational schemes with those from the early days, it appears that the most significant modifications have been applied to the GW source: the impact of flow blocking on the emission of orographic GWs (Lott and Miller 1997) is as established in GWPs as the horizontal dependence and variability of non-orographic sources (e.g., Chun and Baik 1998; Beres et al. 2005; Richter et al. 2010). In the following we list the most important developments having occurred in the last decade, partially in response to the issues listed in section 5, and lessons learned therefrom. A final discussion of main issues and recommendations for the future will follow in section 7.

Numerous studies and developments have continued to address the formulation of the GW source. With regard to the
orographic source it has been shown (van Niekerk and Vosper 2021) that enhancing the description of SGS orography by local spectral decomposition substantially improves the ability of the orographic SGS scheme to adjust to modifications in model resolution, and it enables it to account for partial critical level filtering (Teixeira 2014; van Niekerk et al. 2023), thereby improving various biases. The efficient matrix method suggested by Smith and Kruse (2018) for the calculation of the mountain drag allows these generalizations to be taken into account.

The classic approach to account for GWs emitted from convection is to determine them predominantly from heating profiles provided by the convection parameterization (Chun and Baik 1998; Beres et al. 2005; Richter et al. 2010; Choi and Chun 2011). These profiles can have various biases, but their characteristics, especially their depth, are indispensable information for the determination of the phase-speed spectrum of the emitted GWs. For simplified approaches total precipitation and precipitation rate have been suggested as proxies for convective activity and hence GW amplitude, while keeping the phase-speed spectrum fixed. (Lott and Guez 2013; Bushell et al. 2015). Following Eckermann (2011) and Lott et al. (2012a) it has also been proposed to describe the emission by SGS clouds and convection as a stochastic process (de la Cámara et al. 2014), with the purpose of both enhancing computational efficiency and contributing to GWMF intermittency. Arguably the simulated intermittency has a positive impact on the stratospheric circulation (de la Cámara et al. 2016): occasional, intense waves break at lower altitudes, thereby forcing the stratosphere. This seems to contribute to alleviating a problem that climate models have in reproducing the springtime stratospheric polar vortex breakdown.

Most challenging among the classically considered source processes seems to be the emission of GWs by jets and fronts (Plougonven and Zhang 2014). Based on theoretical work on GWs due to tropospheric potential vorticity anomalies (Lott et al. 2010, 2012b), de la Cámara and Lott (2015) have formulated a stochastic parameterization of this emission process that leads to a GW momentum-flux intermittency as observed in the stratosphere (Fig. 12). However, the deterministic GWP of Bushell et al. (2015), with GW emission linked to precipitation, yields an analogous result. Hence a good representation of GWMF intermittency seems to rely less on stochasticity of the source approach than on a nonlinear relation between the forcing flow and the GW source. Beyond that, the scale and intensity of SGS heating/precipitation has also been shown to strongly influence GW intermittency (Alexander et al. 2021).

Because GW emission by balanced flows might also be at work in the stratosphere (Sato and Yoshiki 2008; Dörnbrack et al. 2018; Polichtchouk and Scott 2020) the approach of de la Cámara and Lott (2015) has been extended by Ribstein et al. (2022) to also allow for spontaneous GW emission in the middle atmosphere, bearing also potential for the description of the emission of secondary GWs from GW breaking (e.g., Vadas et al. 2018; Becker et al. 2022).

Next to sources, GW propagation and the accompanying transient interaction between GWs and the resolved flow has seen increased attention. As outlined in section 1, presently operational GWP s neglect GW impacts due to GW transience, arising from the gradual upward propagation of GWs from a time-dependent source. Muraschko et al. (2015) and Bölöni et al. (2016) have developed a Lagrangian numerical scheme that allows the simulation of this process. Application in a global climate model indicates a significant impact of GW 

![Fig. 10. Time and zonal mean zonal GW forcing (m s⁻¹ day⁻¹) from offline ray tracing through a model atmosphere with solar tides in (a) a general setup (full), (b) a setup noREF without lateral propagation, and (c) a TS experiment in which GW transience and lateral propagation have both been neglected. For further details see Senf and Achatz (2011).](image-url)
transience on the intermittency of GWMFs (Bölöni et al. 2021; Kim et al. 2021). In the tropics transience reduces the GW intermittency due to the variability of the convective source, whereas in the extratropics GW propagation through transient resolved winds causes and enhances intermittency (see Fig. 13).

Another simplifying assumption of traditional GWPs is the neglect of horizontal GW propagation. Ray-tracing simulations of GWs propagating through prescribed climatological flow fields (Senf and Achatz 2011; Kalisch et al. 2014) show a significant impact of horizontal propagation on the distribution of the seasonal and zonal mean GW forcing. Preliminary results from a fully coupled prognostic treatment of GWs in a global climate model indicate that the horizontal distribution of GWMFs in midlatitudes is affected considerably by horizontal propagation (Völker et al. 2023), as is the QBO (Kim et al. 2023). It should be mentioned that corresponding model formulations also expand on traditional GWPs by taking into account horizontal momentum and entropy fluxes, as well as an elastic term in the momentum equation. The corresponding theory has been outlined by Achatz et al. (2017, 2023). As has been shown by Wei et al. (2019), especially the interaction between low-frequency GWs and nonbalanced extratropical flow is potentially affected negatively by the classical approach of only using vertical pseudomomentum fluxes. With increasing model resolution, an ever-increasing part of the resolved flow will be unbalanced GWs. Hence with time this issue might gain in relevance.

Whether nonlinear wave–wave interactions (e.g., Eden et al. 2019; Savva et al. 2021; Völker et al. 2021) should be...
considered in GWPs is an open question to date. Worth noting is also that understanding and description, in GWPs, of the interaction between GWs and turbulence due to wave breaking has not progressed significantly from classic formulations. Those use saturation thresholds based on static stability (Lindzen 1981) or vertical shear instability (e.g., Fritts and Alexander 2003, and references therein), although this is known to be rather inaccurate (e.g., Achatz 2007). The GW-breaking threshold is still one of the most frequently used tuning parameters of GWPs.

A supplement to physics-based parameterization has emerged in the form of machine learning (ML). In principle it allows the learning of a parameterization, or of parts of it, from skillfully defined control data, either from models or observations. This development is still somewhat in its adolescent stages. Matsuoka et al. (2020) demonstrated that a convolutional neural network, when trained on a high-resolution reanalysis product, could represent the GW structure locally over Hokkaido, Japan. Artificial neural networks were used for the first time as a replacement for the conventional GWP in a seasonal forecasting system by Chantry et al. (2021) and in an intermediate complexity climate model by Espinosa et al. (2022). The latter study demonstrated the ability of the ML scheme to generalize out of sample, to a phase of the QBO, and promingly, to a CO₂ concentration that it had not encountered in its training phase. Technical issues regarding the efficient implementation of ML methods in operational models remain. Other open issues of ML-based parameterizations relate to a potential lack of interpretability of their responses as well as their unclear generalizability to situations perturbed away from the training samples, for example, in climate change simulations. A potential application of ML less affected by such challenges is the calibration of parameters in conventional GWPs, where Bayesian methods such as ensemble Kalman inversion seem promising (Mansfield and Sheshadri 2022). In all of its applications ML critically depends on the quality of the training data, be they from measurements or models; that is, it will remain directly affected by all the issues listed in sections 3 and 4.

7. Discussion of status, prospects, and needs for the future

Even with global-model resolutions increasing beyond the GW-permitting grayscale that we see now, GWPs will remain essential, both for practical and for conceptual purposes. Climate modeling and seasonal predictions by model ensembles will keep using codes where a significant part of the GW spectrum will have to be parameterized. However, even should it ever be possible to explicitly resolve all the relevant GWs and their multitude of influences, GWPs will still be useful in a model hierarchy because they enable understanding of processes and mechanisms in the simulations.

Present limitations of GWPs define major goals that should be pursued in future developments. Modeling centers using a suite of models, ranging from horizontally kilometer scale for weather predictions or singular climate projections to considerably coarser models for ensemble simulations, would benefit from seamless formulations in which GWPs adjust correctly to varying model resolutions. This scale adaptivity is not common now but should be a major goal because it would allow modeling centers to adjust model resolutions without retuning the GWP ad hoc to obtain realistic winds and temperatures. It would also ease substantially the interpretation of model results at different resolutions, especially from climate simulations.

The uncertainty of the responses of the QBO and the BDC in climate change simulations is at least partly due to uncertainties in GWPs. This leads to corresponding issues in predictions of teleconnections related to the QBO and BDC, some of which affect surface climate and weather. Together with the known inconsistencies between observed and parameterized GW fluxes this identifies GWPs as a major source of uncertainty in climate change projections. Hence a second major goal should be the robustness and realistic response of GWPs to perturbations of the climate state.

Both goals can only be achieved by enhancing the physical and mathematical basis of GWPs. Ad hoc tuning might often remain unavoidable, but it should only be accepted as an interim resort. Instead, GW-permitting or GW-resolving models might be an essential data source for the development and validation of improved GWPs, provided they use parameterizations of
FIG. 13. Illustration of the effect of GW transience on the intermittency of GW momentum fluxes: (b) Plots showing the westward GW momentum flux due to convection (indicated by red) at a tropical location from a simulation either taking transience into account (TR) or assuming steady state (ST). Transience reduces the intermittency due to the variability of the source. This is seen also in (a) the Gini coefficient, a measure of the strength of intermittency, shown for three different altitudes, as labeled. Note also that in the extratropics intermittency is larger in the transient simulation, indicating the enhancement of intermittency by the propagation of GWs through transient resolved winds. For further details see Kim et al. (2021).
SGS processes, relevant both for GW sources and GW dissipation, that are formulated in a robust manner themselves. This seems to call for increased efforts for observations and SGS-process-resolving idealized modeling of interactions between GWs and such processes, e.g., turbulence, convection, or flow over mountainous terrain. Likewise, measurements of global and local GW statistics will remain an important benchmark for the realism of global GW-permitting models.

The use of data from measurements and GW-permitting simulations for validation benchmarks is not straightforward. GWMFs from observations are highly derived quantities, e.g., often only using temperature information. The observational filter of the instruments also can indicate misleading variations of GW fluxes because the refraction of GWs in wave-number space can move them into and out of the observable scale range. Hence it seems most advisable to first simulate measurements in data from GW-permitting models and thereby make sure that the latter are in reasonable agreement with the observations. Once this is ascertained, GW-permitting models can provide benchmarks for GW fluxes obtained from GWPs.

Two classes of GWP approaches might emerge eventually: highly efficient nontransient schemes with at best very simple descriptions of the effects of lateral propagation will be at the one end, and transient GW models with full horizontal propagation at the other. The latter are more expensive than the first, but they are still considerably more efficient than global GW-resolving codes. They might therefore at least be a research tool guiding the improvement of the more efficient highly simplified schemes. It will have to be seen how much those can be formulated so that the goals of scale awareness and robustness under climate variability can be achieved. An intermediate approach might be provided by the aid of ML that might teach us how to better formulate single-column, steady-state schemes, how to formulate more general GW models in a more efficient manner, or how to carry through uncertainties in parameterization to model output. This way, GW research will remain a field in which a skillful combination of observations, modeling, and theory, including mathematics, promises the best gains.

Acknowledgments. Author Achatz thanks the German Research Foundation (DFG) for partial support through the Research Unit “Multiscale Dynamics of Gravity Waves” (MS-GWaves; Grants AC 71/8-2, AC 71/9-2, and AC 71/12-2), CRC 181 “Energy transfers in Atmosphere an Ocean” (Project Number 274762653, Projects W01 “Gravity-wave parameterization for the atmosphere” and S02 “Improved Parameterizations and Numerics in Climate Models”), and CRC 301 “TPChange” (Project-ID 428312742, Projects B06 “Impact of small-scale dynamics on UTLS transport and mixing” and B07 “Impact of cirrus clouds on tropopause structure”). He also thanks the German Federal Ministry of Education and Research (BMBF) for partial support through the program Role of the Middle Atmosphere in Climate (ROMIC II: QUBICC) and through Grant 01LG1905B. Author Alexander was supported by the U.S. National Science Foundation under Awards 1642644 and 2004512. Author Chun was supported by a National Research Foundation of Korea grant funded by the South Korea government (2021R1A2C100710212). Author Dornbrack has been partly funded by the Federal Ministry for Education and Research under Grants 01 LG 1907 (Project WASCLIM) and through Grant 01 LG 1907 (Project WASCLIM) in the frame of the Role of the Middle Atmosphere in Climate (ROMIC) program. Further support by DFG under Grants GW-TP/DO 1020/9-1 and PACOG/RA 14006-1 in the frame of the DFG-Research Group MS-GWAVES is also acknowledged. Author Holt was supported by the NASA Aura Science Team (NASA Grant 80NSSC20K0950). Author Sato acknowledges support by JSPS KAKENHI Grant 22H00169. Author Plougonven is supported by ANR Project BOOST3R (ANR-17-CE01-0016-01) and by the Strateole-2 project, which is sponsored by CNES, CNRS/INSU, and NSF. Author Sheshadri acknowledges support from the NSF through Grant OAC-2004492. Author Stephen was funded by the Minerva Fast Track Programme of the Max Planck Society. Author Wright acknowledges support from a Royal Society University Research Fellowship (URF/R221023) and by NERC Grants NE/S00985X/1 and NE/V01837X/1. Achatz, Sheshadri, and Stephen are furthermore grateful for support by Eric and Wendy Schmidt through the Schmidt Futures VESRI “DataWave” project.

Data availability statement. Except for Fig. 7, no datasets were generated or analyzed during the current study. With regard to Fig. 7, the raw model output on the native grids amounts to few hundred terabytes, and thus it is not possible to make all data available for more than a few steps. However, postprocessed output used to make the figures will be retained and made available to those who request it. The analysis codes can be made available upon request.

REFERENCES


van den Bremer, T. S., and B. R. Sutherland, 2018: The wave-induced flow of internal gravity wave packets with arbitrary


