Direct comparison between a non orographic gravity wave drag scheme and constant level balloons

Francois Lott, Raj Rani, Aurélien Podglajen, Francis Codron, Lionel Guez, Albert Hertzog, Riwal Plougonven

Direct comparison between a non orographic gravity wave drag scheme and constant level balloons

F. Lott[1], R. Rani[1], A. Podglajen[1], F. Codron[2], L. Guez[1]A. Hertzog[3], and R. Plougonven[4].

[1]Laboratoire de Météorologie Dynamique (LMD)/IPSL, PSL Research Institute, Ecole Normale Supérieure, Paris, France.
[2]LOCEAN/IPSL, Sorbonne Université/IRD/MNHN/CNRS, Paris, France.
[3]LMD/IPSL, Sorbone Université, Paris, France.
[4]LMD/IPSL, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France

Key Points:

- A non-orographic parameterization tuned to produce a realistic tropical quasi-biennial oscillation is used to predict in-situ observations
- Parameterized gravity waves needed in large-scale models have realistic amplitudes in the tropical lower stratosphere
- Day-to-day variations of the estimated gravity wave momentum fluxes correlate well with observations

Corresponding author: Francois Lott, `flott@lmd.ens.fr`

Abstract

The parameterization scheme that represents gravity waves due to convection in LMDz-6A, the atmospheric components of the IPSL coupled climate model (IPSLCM6), is directly compared to Strateole-2 balloon observations made in the lower tropical stratosphere from November 2019 to February 2020. The input meteorological fields necessary to run the parameterization offline are extracted from the ERA5 reanalysis and correspond to the instantaneous meteorological conditions found underneath the balloons. In general, we find a fair agreement between measurements of the momentum fluxes due to waves with periods less than 1 hr and the parameterization. The correlation of the daily values between the observations and the results of the parameterization is around 0.4, which is statistically elevated considering that we analyse around 600 days of data and surprisingly good considering that the parameterization has not been tuned: the scheme is just the standard one that helps producing a Quasi-Biennial Oscillation in the IPSLCM6 model. Online simulations also show that the measured values of momentum fluxes are well representative of the zonally and averaged values of momentum fluxes needed in LMDz-6A to simulate a QBO. The observations also show that longer waves with periods smaller than a day carry about twice as much flux as waves with periods smaller than an hour, which is a challenge since the low period waves that make the difference are potentially in the grey zone of most climate models.

Plain Language Summary

In most large-scale atmospheric models, gravity wave parameterizations are based on well understood but simplified theories and parameters which are keyed to reduce systematic errors on the planetary scale winds. In the equatorial regions, the most challenging errors concern the quasi biennial oscillation. Although it has never been verified directly, it is expected that the parameterizations tuned this way should transport a realistic amount of momentum flux in both the eastward and westward directions when compared to direct observations. Here we show that it is the case, to a certain extent, using constant-level balloon observations at 20 km altitude. The method consists in comparing directly, each day and at the location of the balloon the measured momentum fluxes and the estimation of a gravity wave parameterization using observed values of the large-scale meteorological conditions of wind, temperature and precipitation.

1 Introduction

It is well known that precipitation forces gravity waves (GWs) that propagate in the stratosphere (Fovell et al., 1992; Alexander et al., 2000; Lane & Moncrieff, 2008). These waves carry horizontal momentum vertically and interact with the large scale flow when they break. The horizontal scale of these waves can be quite short, much shorter than the horizontal scale of General Circulation Models (GCMs) so they need to be parameterized (Alexander & Dunkerton, 1999). Although there are other sources of gravity waves that need to be parameterized, like mountain waves (Palmer et al., 1986; Lott, 1999) and frontal waves (Charron & Manzini, 2002; Richter et al., 2010; de la Cámara & Lott, 2015), the convective GWs are believed to dominate largely in the tropics. In these regions, they contribute significantly to the forcing of the Quasi-Biennial Oscillation (QBO), a near 28-month oscillation of the zonal mean zonal winds that occurs in the lower part of the equatorial stratosphere (Baldwin et al., 2001). For these reasons, the parameterization of convective GWs is necessary for most GCMs to explicitly realize the QBO.

Although convective gravity wave parameterizations are now used in many models with success (Beres et al., 2005; Song & Chun, 2005; Lott & Guez, 2013; Bushell et al., 2015), their validation using direct in situ observations remains a challenge. There exist observations of GWs using global satellite observations (Geller et al., 2013) but the GWs identified this way still have quite large horizontal scales, and some important quantities like the Momentum Fluxes (MFs) are often deduced indirectly, for instance from temperature measurements using polarization relations (Alexander et al., 2010; Ern et al., 2014). For these two reasons, in situ observations are essential, and the most precise ones are provided by constant-level long-duration balloons, like those made in the Antarctic region during Strateole-Vorcore (Hertzog, 2007) and Concordiasi (Rabier et al., 2010), or in the deep tropics during PreConcordiasi (Jewtoukoff et al., 2013) and more recently Strateole 2 (Haase et al., 2018). Among many important results, these balloon observations have shown that the momentum flux entering in the stratosphere is extremely intermittent (Hertzog et al., 2012). This intermittency implies that the mean momentum flux is mostly transported by few large-amplitude waves that potentially break at lower altitudes than when the GW field is more uniform. This property, when reproduced by a parameterization (de la Cámara et al., 2014; Kang et al., 2017; Alexander et al., 2021), can help reduce systematic errors in the midlatitudes, for instance on the timing of the final warming in the Southern Hemisphere polar stratosphere (de la Cámara et al., 2016), or on the QBO (Lott et al., 2012). Balloon observations have also been used to characterize the dynamical filtering by the large scale winds (Plougonven et al., 2017), and to validate the average statistical properties of the GW momentum flux predicted offline using reanalysis data (Kang et al., 2017; Alexander et al., 2021).

However, to the best of our knowledge, the evaluation of parameterizations using balloon observations have remained quite indirect so far, with the common belief that the best a parameterization can do is to reproduce a realistic statistical behaviour (Jewtoukoff et al., 2015; Kang et al., 2017; Alexander et al., 2021). The fact that a paramerization could be used to simulate the observed momentum flux at a given time and place has never been tried, and there are few good reasons for that. For example, parameterizations are often based on simplified quasi-linear wave theory, they assume spectral distributions that are loosely constrained, and they ignore lateral propagation almost entirely (some attempt to include it can be found in Amemiya and Sato (2016)). Nevertheless, some factors could mitigate these weaknesses. One is that in most parameterizations the wave amplitude is systematically limited by a breaking criterion that encapsulates nonlinear effects. An other is that many parameterizations explicitly relate launched waves to sources, and there is constant effort to improve the realism of the convective ones (Liu et al., 2022). Also, observations systematically suggest that dynamical filtering by the large scale wind is extremely strong for upward propagating GWs (Plougonven et al., 2017), and this central property is represented in most GW parameterizations. For

all these reasons, it may well be that a GW parameterization keyed to the large scale conditions found at a given place and time gives MFs that can be directly compared to the MFs measured by a balloon at the same place.

Based on the relative success of the offline calculations done in the past using reanalysis data (Jewtoukoff et al., 2015; Kang et al., 2017; Alexander et al., 2021), the purpose of this paper is to attempt such a direct comparison using the most recent observations. We will use for that the balloons of the first Strateole 2 campaign that flew in the lower tropical stratosphere between November 2019 and February 2020 (Corcos et al., 2021). For each of these flights and each time, we will identify the grid point in the ERA5 reanalysis (Hersbach et al., 2020) that is the nearest and used the vertical profiles of wind and temperature as well as the surface value of precipitation to emulate the Lott and Guez (2013)'s (LG13) parameterization of convective GWs. The plan of the paper is as follows. Section 2 describes the methodology used, section 3 analyzes in-depth the statistics and compares with online simulations. A summary with discussion on the method and some perspectives is provided in section 4.

2 Data and method

2.1 Parameterization of convective gravity waves

We take the LG13 parameterization of non-orographic gravity waves forced by convection that is operational in LMDz6-A (Hourdin et al., 2020) the atmospheric component of the IPSL Earth System model used to complete the CMIP6 experiments (IPSLCM6, Boucher et al. (2020)). This version of the parameterization is also used for the LMDz-6A experiments carried out in the frame of the QBO intercomparison project (QBOi) (Bushell et al., 2022; Holt et al., 2022). Among the salient aspects of the scheme, one is that it is multiwave and stochastic, the subgrid scale GW field (e.g., vertical wind disturbance w') being represented by stochastic Fourier series of monochromatic waves,

$$w'(x,y,t,z) = \sum_{n=1}^{\infty} C_n \hat{w}_n(z) e^{i(\mathbf{k_n}\cdot\mathbf{x}-\omega_n t)}, \tag{1}$$

where the intermittency parameters satisfy $\sum_{n=1}^{\infty} C_n^2 = 1$, and where $\mathbf{k}_n$ and ω_n are the horizontal wave vector and frequency respectively. To determine the wave amplitude the variance of the subgrid scale precipitation field, P', is assumed to compare in amplitude with the gridscale averaged precipitation P by writing

$$P' = \sum_{n=1}^{\infty} C_n P e^{i(\mathbf{k_n}\cdot\mathbf{x}-\omega_n t)}. \tag{2}$$

We then translate precipitation into diabatic heating which we distribute vertically over a fixed depth Δz in the troposphere. For each harmonic the heating produces a GW whose MF varies with the square of the precipitation P^2 times a tuning parameter G_{uw0} (see Equation 9, in LG13) and which is imposed at a fixed launching altitude z_l. Three factors then limit vertical propagation, (i) the presence of critical levels, (ii) a dissipative term controlled by a kinematic viscosity ν, and (iii) a criterion for saturation controlled by a saturation parameter S_c. All these effects are summarized in Equation 12 of LG13, but to illustrate how the background flow controls the outcome of the scheme, we recall the third criterion that saturation limits the amplitude of the momentum flux transported by each harmonic to values below

$$|\rho \hat{\mathbf{u}}_n \hat{w}_n^*| \leq \rho_r S_c^2 \frac{|\mathbf{k}_n \cdot (\mathbf{c}_n - \mathbf{U}(z))|^3}{N(z)} \frac{k_m}{|\mathbf{k}_n|^4}, \tag{3}$$

where the star stands for the complex conjugate, $\hat{\mathbf{u}}_n$ the harmonic of the horizontal wind vector disturbance, and $\mathbf{U}(z)$ and $N(z)$ the vertical profiles of horizontal wind vector and

buoyancy frequency respectively. Still in Equation 3 the reference density $\rho_r = 1\text{kgm}^{-3}$ and k_m is the the minimal horizontal wavelength that needs to be parameterized: its inverse scales the gridcell horizontal size. Compared to LG13, we have slightly rewritten the saturation criteria to make explicit that the saturated MF has small amplitude when the intrinsic phase speed amplitude in the wavevector direction $|\mathbf{k}_n \cdot (\mathbf{c}_n - \mathbf{U})|$ is small, the absolute horizontal phase speed being $\mathbf{c}_n = \mathbf{k}_n \omega_n / |\mathbf{k}_n|^2$. This mechanism is referred to as dynamical filtering in the following and is probably central in explaining the good correlations we describe next between the observed and parameterized MFs.

In practice, we make a distinction between the time scale of the life cycle of the waves Δt which we consider to be shorter than one day and the physical time step that separates two calls of the parameterization, and which is around $\delta t = 10$mn online. This distinction permits to launch few waves each time-step, typically $J = 8$, and to accumulate their effect over the day via an AR-1 process with decorrelation time of one day. On average and each day, the GW field is then made of $J\frac{\Delta t}{\delta t} \approx 1000$ harmonics, a number we judge sufficient for a realistic representation. In the offline comparison we will not use such a large number of waves, essentially because it makes little sense to interpolate ERA5 along 10mn intervals, but we will still launch 8 waves per hour to leave the scheme unchanged and average over the day. To test if the reduction in terms of number of harmonics has significant impact we have rather made statistics averaging the parameterization results over 9 adjacent gridpoints. In this case the number of harmonics involved becomes comparable to that used online and we did not find large qualitative differences. For completeness, we recall here the operational values of the different tuning parameters of the scheme used for CMIP6,

$$z_l = 5\text{km},\ \Delta z = 1\text{km},\ G_{uw0} = 0.23, S_c = 0.6,\ \rho_r = 1\text{kg/m}^3,\ k_m = 0.02\text{km}^{-1}. \quad (4)$$

The scheme selects randomly the horizontal wave number between $k_s < k < k_m$ using uniform distribution and select the intrinsic phase speed at the launch level z_s according to a Gaussian distribution with standard deviation C_M. The operational values for these parameters are,

$$k_s = 1\text{km}^{-1},\ C_M = 30\text{m/s},\ z_s = 5\text{km}. \quad (5)$$

It is important to emphasize that the scheme selects phase speeds rather than frequency, whereas the balloon data measure MFs as a function of intrinsic frequency. We therefore analysed the characteristic distribution of the intrinsic frequency of the parameterized waves that enter in the stratosphere and verified that more than 75% of the parameterized momentum fluxes are carried by harmonics with intrinsic period around and below 1hr (not shown). Note also that in its operational version, and to limit computational costs, only waves with horizontal wavenumber in the zonal direction are launched.

2.2 Offline parameterization runs

To activate the scheme in offline mode we will use ERA-5 hourly data of precipitation and 3-hourly data of winds, surface pressure and temperature at $1^o \times 1^o$ horizontal grid to mimic a large scale climate model resolution. Winds and temperature are then linearly interpolated on 1hr time step to be synchronised with precipitation. In the vertical we use data at 67 model levels, taking one every two ERA5 levels, to speed up calculations but also to mimic the vertical resolution we have in the LMDz-6A GCM and which is slightly below 1km (in ERA5 and around 20km the vertical resolution is around 500m when all the 137 levels are considered).

2.3 Strateole 2 balloon observations

The in situ observations we use are from the 8 constant level balloon flights which flew between 18.5 and 20km altitude for about 2-3 months during the Nov. 2019-February

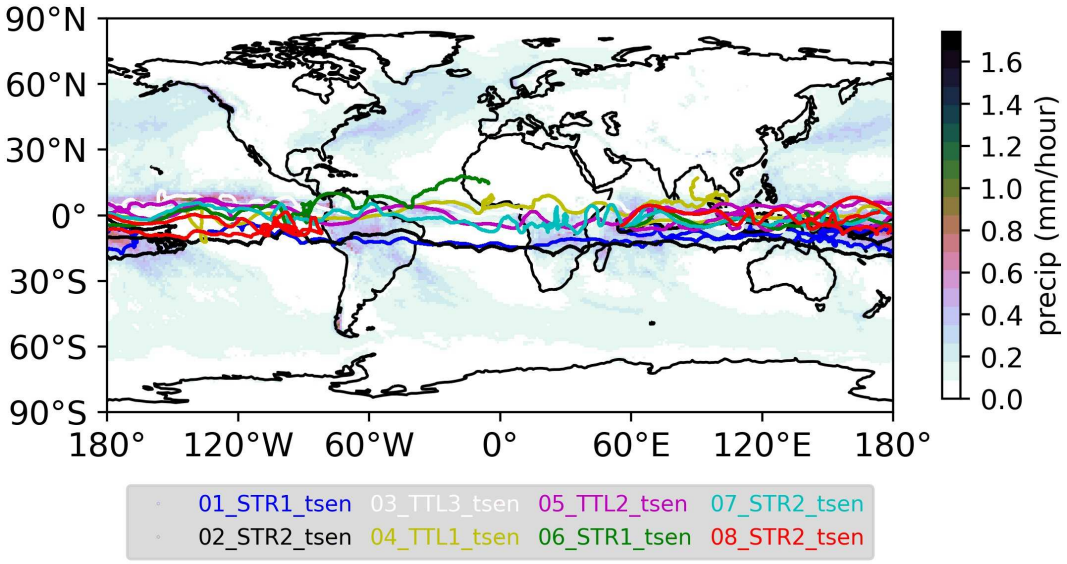

Figure 1. Strateole 2 balloon trajectories taking place between November 2019 and February 2020. Shading presents the precipitation field from ERA5 averaged over the period.

2020 periods of Strateole-2 (Corcos et al., 2021). Their trajectories are shown in Figure 1, superimposed upon the averaged precipitation. In the MFs calculated from observations Corcos et al. (2021) distinguish the waves with short periods (1hr-15mn) from the waves with period up to one day (1d-15mn). They also distinguish the eastward waves giving positive MF in the zonal direction from the westward waves giving negative MF, and from the MF amplitudes including all the directions of propagation. It is coincidental that the flights took place during the 2nd documented QBO disruption (Anstey et al., 2021), but the fact that the measurements are below the altitude at which the disruption manifests makes us believe that our comparison between gravity wave MFs over the period is not much affected by the disruption (beyond the fact that the disruption potentially affects the large scale winds, which is something that translates well in the parameterization).

In the following we will compare the momentum fluxes derived from the balloon data, emphasize the intrinsic frequencies that the scheme represents (the intrinsic periods below $1hr$) and consider the ERA5 data at the points that are the nearest from the balloon. The prediction is then made every hour and averaged over the day, again because this is the time scale needed for our scheme to sample realistically a GW field, and also because it takes around a day for a balloon flight to cover about a model grid-scale. We will discuss sensitivities to these choices in the first paragraph of the conclusion.

2.4 Online simulations

An important aspect of our work is that we use an operational scheme in offline mode without further tuning, a methodology that potentially impacts the amplitude of the MFs when compared to "free" model runs where dynamics and GWs evolve jointly. To measure the differences we will therefore compare offline and online calculations over around 3 QBO cycles (8 years). More precisely, we will make global estimations of the MFs and drag predicted offline using 8 years of ERA5 6hourly data (2013-2020). We will compare them to the MFs and drag issued from the LMDz-6A atmospheric model running over the same period and at its medium resolution (144x143 regular longitude-latitude grid, 80 vertical levels with top at 1Pa). We first consider a "free run" , only forced with

the observed sea surface temperatures and sea-ice from the CMIP database, and the ozone climatology from the ACC/SPARC ozone database. This free run have the same settings of the parameterizations of orographic GWs (Lott, 1999), convective GWs (LG13), and GWs due to fronts and jet imbalances (de la Cámara & Lott, 2015) as those used for CMIP6 (Hourdin et al., 2020). To make a smooth transition from the offline estimations with ERA5 to the free run done with LMDz-6A, we will also present a simulation with LMDZ-6A where the fields of horizontal wind and temperature are nudged toward ERA5 every 6hr with a relaxation constant of 1hr^{-1}.

3 Results

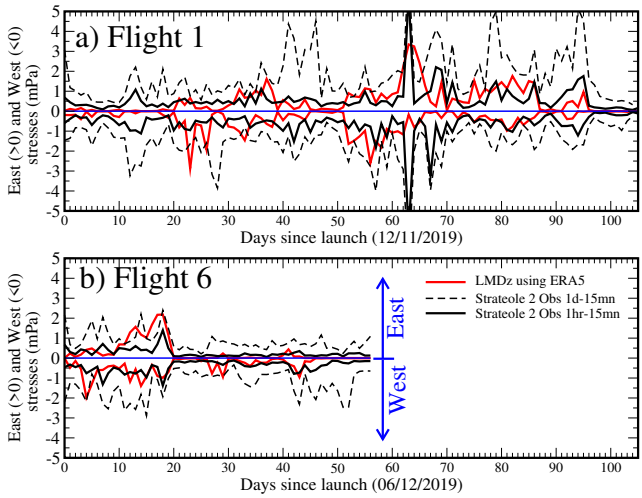

Figure 2. Comparison between daily averaged values of the eastward and westward MFs measured by the balloons and estimated by the GW scheme at the balloon location and altitude. Red curves are for the GW scheme predictions using ERA5, black curves are for the observed MFs due to the 15mn-1hr GWs, the thin dashed black curves are for MFs due to the 15mn-1day GWs.

3.1 Offline estimate of the observed values

Figure 2 shows time series of momentum fluxes measured during two balloon flights and the corresponding offline estimates. For clarity we present results for the Eastward and westward MF only, we will return more briefly to the accumulated MF and to the MF amplitude later. Note nevertheless that the operational GW scheme using only two directions of propagation, the accumulated MF is the sum of the eastward and westward MFs, the MF amplitude being their difference. Overall one sees in the two panels that the amplitudes of the momentum fluxes corresponding to the 1hr-15mn periods in the measurements compare well to the parameterized amplitudes in both the eastward and westward directions, the eastward and westward fluxes being of comparable amplitude but of opposite sign, as expected. For both, the observed momentum fluxes related to the 15mn-1day waves are substantially larger. In general and for flight 1 in Figure 2a, one sees that the parameterized fluxes are sometime small in amplitudes and in both directions (between days 10 and 20), something that rarely happens in the observations. One also sees a tendency for the observed and estimated values to become larger jointly, like for instance the Eastward fluxes fluxes between days 60 and 95, before becoming small jointly afterward. This contrast between periods with larger and smaller MFs is even more pronounced in the flight 6 shown in Figure 2b. In it, one sees that the eastward and west-

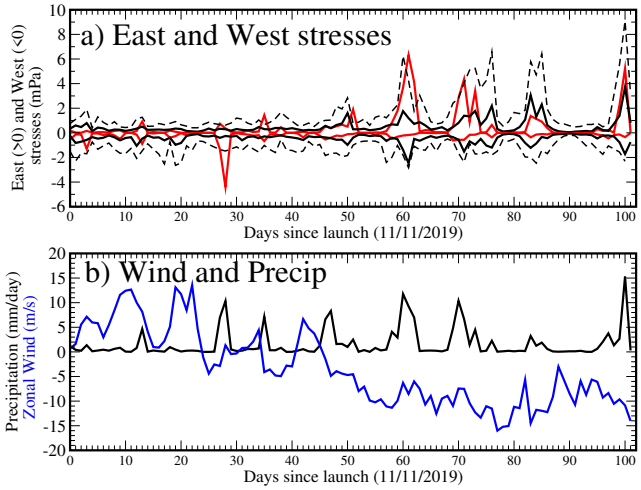

Figure 3. a) Same as Figure 2 but for Strateole Flight 2. b) ERA5 precipitations and zonal wind at the flight altitude.

ward MFs are quite large in amplitude before day 20, and are much smaller afterward. If we look at the trajectory in Figure 1 one sees that at its end, Flight 6 moves from the equatorial regions toward the subtropics and ends up over Sahel, i.e. a region of low precipitation.

To give a better sense of what can cause the resemblances and differences between the observed MFs and their estimation, we plot in Figure 3a the eastward and westward MFs for flight 2, and in Figure 3b the precipitation and the (ERA5) zonal wind at the flight altitude. This is an interesting period since the zonal wind during this flight changes direction. Without a surprise, one sees in Figure 3 that the estimated flux peaks when the precipitation is large. One also sees that the MFs peaks are more pronounced in the direction opposed to the zonal wind consistent with the fact that waves with large amplitude intrinsic phase speed can carry more momentum than waves with small intrinsic phase speed (by dynamical filtering, see the first numerator in Equation 3). To a certain extent, the relation with intense precipitation can be seen in the observations, mainly in the eastward direction after day 40. Dynamical filtering is also active for the measured fluxes, the observed westward fluxes being small compared to the eastward flux when the zonal wind becomes negative after day 50. Again, when the precipitation is small the simulated MFs are often small, whereas the observed ones always have non-zero backgrounds.

The fact that the parameterization estimates fluxes of about the right amplitude is summarized in Figure 4, where the average of the fluxes over the 8 entire flights are shown. It confirms systematically that the offline estimations are quite good on average and in the zonal direction, for the eastward and westward components again, but also on the accumulated flux (i.e the sum of the two and where the contributions from eastward and westward propagating GWs largely oppose each other). In terms of stress amplitudes one sees that the observations give larger value on average, but this is due to the fact that in Corcos et al. (2021) the amplitudes take into account the meridional components of the stress which are not included in the parameterization tested here. In the panel are also shown the correlations between the balloon averaged values of the stresses, they are often quite significant, despite the fact that only 8 flights are used.

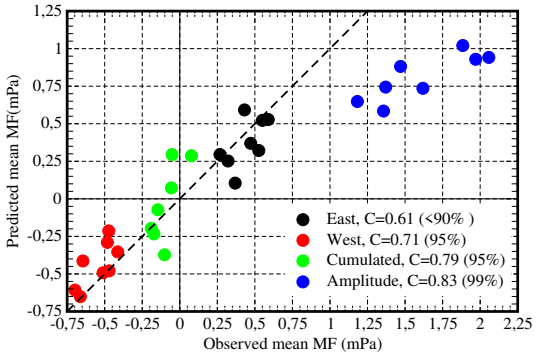

Figure 4. Scatter plot of the momentum fluxes measured by the balloon versus parametrized. Eight values are averages for the eight balloon flights. Significant test are for the correlation between the values, significance is estimated via a Pearson test with 6 degrees of freedom.

The curves in Figure 2 and Figure 3 also suggest that observations and offline estimations evolve quite similarly day after day, both measured and parameterized MFs being sensitive to precipitation and dynamical filtering. To test more systematically this relationship, we next calculate the correlations between measured and estimated MFs and for each flight (Table 3.1). To test the significance, we measure the number of Degrees of Freedom (DoF) present in each dataset, and calculate for that the decorrelation time scale, which we take as the lag in day beyond which the lag-autocorrelation of the series falls below 0.2. As this time-lag varies from one series to the other, we give explicitly in column 5, the number of DoF, which is the duration of the flight divided by the decorrelation time scale. Note that for their decorrelation time, we consider for simplicity that evaluated with daily averaged observations, but found that it is not much different from that evaluated with the offline estimates (not shown). In each case, we find positive correlations, they are often significant in the Eastward direction and for the amplitude, the estimated westward fluxes presenting more errors. Even weaker correlations occur for the accumulated stress, which most certainly reflects that in the accumulated stresses, large quantities of opposite sign balance one another, the resulting balance being more difficult to predict.

As already mentioned, a defect of the scheme is it lacks background wave activity in the absence of precipitation. This means that momentum fluxes are underestimated in many circumstances, despite the fact that the amplitudes are realistic when considering long term averages. To analyse better this difference and its potential consequences, the Figure 5 presents PDFs of the distributions of the momentum fluxes considering all the daily data. For the PDFs (solid line), one sees that the balloons almost systematically measure fluxes with amplitude between 0.1mPa and 10mPa (see Figure 5a), whereas in the parameterization there are many more contributions from the smaller amplitude momentum fluxes (solid red), not mentioning that the zero values are excluded from PDFs when plotted versus the logarithm of MF amplitudes. To test if this difference in MF amplitude distribution has consequences, the dashed lines represent the contribution of a given MF value to the mean stress (which is just the PDF multiplied by the MF value itself). For the amplitudes, the values which actually contribute lie between 0.1mPa and 10mPa in both the observations and the offline estimations. The fact that the small amplitude waves are more frequent in the estimations is also true for the westward and eastward components of the stress (Figures. 5b and 5c respectively). For both, nevertheless, the contribution to the average stress is due to larger amplitude waves in the estimations than in the observations, as indicated by the shifts towards larger values of the MFs between the black dotted curve and the red dotted curve in Figures 5b and 5c.

Table 1. Correlation coefficient (24 hours averaged) between Strateole -2 Balloon flight (1hour15 min waves) and offline estimation (Notations for Significance Level: 99% : bold black with underline; 95%: bold black; 90% : solid; below 90%: solid italic). The significance are attributed following a Pearson test with degrees of freedom measured as the number of day divide by the decorrelation time of the series.

Flight	Altitude	Launch	End	Duration DOF	Cumulated	Amplitude	East	West
01_STR1	20.7	12/11/2019	28/02/2020	107/53	0.23	**0.28**	**0.46**	*0.07*
02_STR2	20.2	11/11/2019	23/02/2020	103/51	*0.21*	**0.62**	**0.62**	*0.05*
03_TTL3	19.0	18/11/2019	28/02/2020	101/33	**0.49**	**0.42**	**0.49**	**0.43**
04_TTL1	18.8	27/11/2019	02/02/2020	67/22	**0.41**	**0.55**	**0.55**	**0.53**
05_TTL2	18.9	05/12/2019	23/02/2020	79/19	0.36	*0.29*	0.36	*0.24*
06_STR1	20.5	06/12/2019	01/02/2020	57/10	*0.39*	**0.67**	**0.71**	**0.59**
07_STR2	20.2	06/12/2019	28/02/2020	83/16	*0.01*	*0.09*	*0.08*	*0.06*
08_STR2	20.2	07/12/2019	22/02/2020	77/12	*0.18*	**0.7**	**0.66**	*0.37*
ALL	x	x	x	670/170	**0.30**	**0.41**	**0.51**	**0.29**

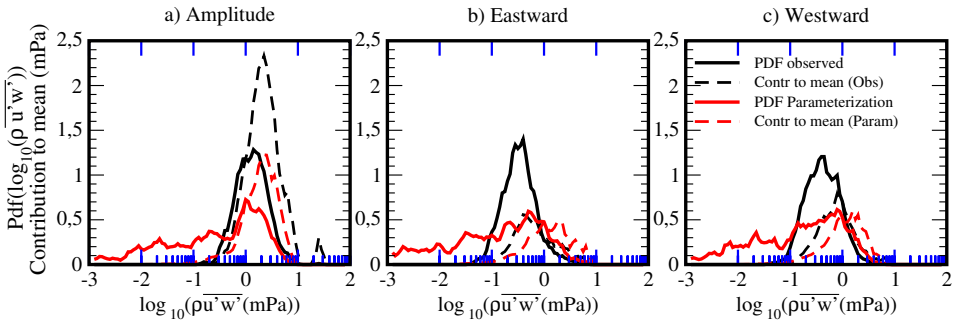

Figure 5. PDFs of daily values of momentum flux distribution (solid lines). The PDFs are calculated from histograms of 670 MFs daily value within intervals of $\Delta\left(\log_{10}\rho\overline{u'w'}(\text{mPa})\right) = 0.05$, thereafter smoothed by a 5 point non-recursive filter with weight $(0.1, 0.2, 0.4, 0.2, 0.1)$. For the contribution of the waves to the MF (dashed lines), the PDF values are multiplied by the MF values $\rho\overline{u'w'}$ (in mPa). Measured values are in black, estimations using ERA5 data and the LG13 GWs parameterization are in red.

3.2 Global prediction and comparison with GCMs results

To appreciate whether the offline GW drag estimations using ERA5 are representative of the GWs MFs that a GCM requires to simulate a QBO, Figure 6a) presents time-altitude sections of the equatorial zonal winds and GWD predicted by the scheme globally and in offline mode between 2013-2020. In it we see that the gravity wave drag is negative (positive) where the zonal mean zonal wind vertical shear is negative (positive) consistent with the fact that it contributes to the descent of the QBO. We also note that the amplitudes vary between ±0.5m/s/day, a range characteristic of the parameterized GW tendency used in GCMs that produce a quasi-biennial oscillation (Butchart et al., 2018). The figure also indicates with a green rectangle the region and period during which the balloons operated.

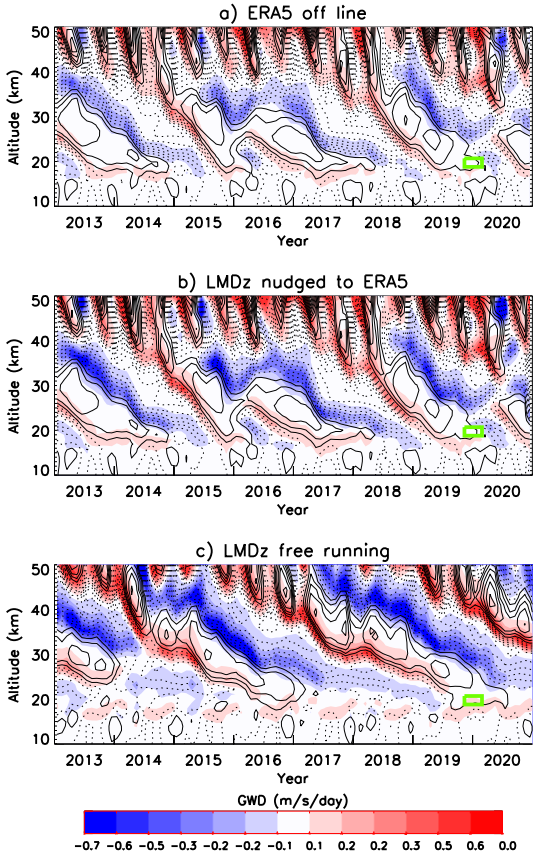

Figure 6. Time vertical sections of the zonal mean zonal wind (CI=10m/s, negative values dashed) and non orographic gravity wave drag zonal tendency (color) averaged over the equatorial band (5^{o}S-5^{o}S). Input data and GWD tendency are from a) ERA5 reanalysis and offline GWD scheme; b) LMDz-6A nudged to ERA5 and online GWD scheme; c) LMDz-6A free run an online GWD scheme. The green box indicates schematically the altitude and time ranges of the Strateole-2 flights considered in this study.

To check that comparable relations between GW drag and zonal wind shear occur in the LMDz-6A model, Figure 6b) first shows result from the nudged simulation. In it we see a strong resemblance in the GWD compared to the offline GWD in Figure 6a, the amplitudes being nevertheless substantially larger (around 25%) in the model. If we then look at the free run in Figure 6c) one sees again quite realistic relations between wind shears and drag, the GW drag is again substantially larger than that predicted using ERA5, despite the fact that the QBO period in LMDz-6A shown here is quite long (about 3 years here, LMDz-6A being a GCM with a long QBO (Bushell et al., 2022)). Note that this longer period could be reduced by enhancing the GW amplitude via for instance the tuning parameter G_{uw0} in Equation 4. Our experience with LMDz-6A is that the increase in G_{uw0} needed is of about few percents, a value that should not affect significantly the calculations done here with the operationnal scheme (not shown but for a systematic discussion about tuning parameters and the QBO see Garfinkel et al. (2022)).

To address the differences in MFs globally, Figure 7a) shows the zonal mean of the eastward and westward MFs averaged over the equatorial band at 20km (i.e. about the balloon flights altitude) in the three runs, and the corresponding averaged precipitations (Figure 7b). The amplitude of the eastward and westward GW stresses are about 25%

and systematically larger in the free run (long dash) than in the offline test (solid), the nudged simulation being in between (dotted). Also interesting, the average of the stresses amplitude in the offline calculations are near ±0.5mPa, which is quite close to the average value of the amplitude of the stresses estimated locally and measured during the balloon flights (see y values of the red and black dots in Figure 4).

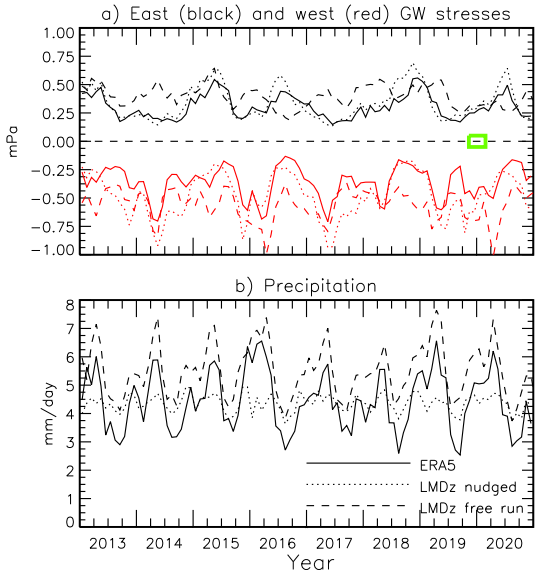

Figure 7. Time series of the z = 20km zonal mean zonal non-orographic GW stresses and of the zonal mean precipitation averaged over the equatorial band (5^{o}S-5^{o}S). Same input datas as in Figure 6.

As said in the introduction, the version of the parameterization used is operational in the atmospheric component of the IPSLCM6 model, and we have tried to change it as little as possible, which forces us to make choices. One is that we only call the parameterization every hour in offline mode, using interpolated data from ERA5, rather than every 10mn in the model. Another is that LMDz-6A has a different grid yielding interpolation errors that could make the behaviour of the parameterization very different between ERA5 and LMDz-6A. Despite these differences it is remarkable that the errors are not outrageous, they also have a cause that is quite identifiable. In Figure 7b) one sees that in the free run, LMDz-6A overestimates by about 15% precipitation compared to ERA5, as in the scheme the sources use square precipitation (see Equation 9 in LG13), a 25% difference in the MFs is therefore not a surprise.

Interestingly also, the MFs in the nudged runs (dotted lines in Figure 7a) follow quite well the offline predictions. This is quite informative because in these nudged runs, the precipitation quite fails in representing an annual cycle (see the dotted line in Figure 7b), which is probably due a mismatch between the nudged fields of temperature and winds and the model representation of diabatic processes. Although we did not analyse further the precipitation errors in these nudged runs, the fact that in them the MFs in Figure 4a) follows quite well the offline predictions is an indirect evidence that dynamical filtering plays a central role. Regarding that in Figures. 2-3 and in Table 3.1 the good correlations between predictions and balloon data seems in part related to the precipitation, it may be that globally such precise relationships are not necessarily needed, once about the right amount of precipitation is predicted.

4 Conclusion

The main result of this paper is that a state of the art parameterization of GWs due to convection reproduces reasonably well the momentum flux due to the high-frequency waves (periods between 15mn and 1hr) deduced from in situ measurements made onboard constant-level balloons. The parameterization represents well the eastward and westward values of the stress and their variations from day to day. We have made sensitivity tests, considering averages over 3hrs or 6hrs instead of a day, or/and averages over neighbouring points to increase the number of harmonics in the offline predictions and found little qualitative differences. For instance, averaging balloon data and predictions over shorter period, say 3hrs, result in much more noisy and decorrelated series, the correlation between observations and measurements is lower but the DoF increase so the level of significance stays about the same as when using daily data as in Table 3.1. We have the impression that the best relations between observations and predictions are always for periods around a day and above. Note that this does not contradict our understanding of what a parameterization should do or a single balloon flight sample. In fact, a parameterization like LG13 needs successive iterations to evaluate the large number of harmonics needed to represent realistically a GW field. Quite similarly, a balloon that progresses at a speed around 10m/s takes about 3 hours to travel through a 1^{o} long model gridcell, and this is just one transect. An ergocidity argument could be used to justify that averaging over a few 3-hours transects to cover a gridcell is equivalent to averaging the balloon data over 1day. This being said, we cannot exclude that better could be done with an other parameterization when considering shorter time scale. We nevertheless suspect that that the gain will not be to large in terms of correlation significance for instance, simply because the signals we handle are in essence extremely noisy. The daily averages we have done, on top of making some sense in the GW context help to increase the signal/noise ratio.

Another important aspect of our work is that these results demonstrate that the GWs parameterization used in a large scale model to simulate a QBO parameterize MFs directly comparable with in situ observations. Although the measurements are extremely local, we verify that the average value they give is representative of the global values needed by a large-scale model to produce a QBOs. This is an important result in our opinion and for two reasons.

The first is that according to a common belief, there are discrepancies of a factor larger than 2 between the MFs parameterized in models and the global observations (Geller et al., 2013), at least in the mid-latitudes. In the equatorial regions, and using the same data as here, Corcos et al. (2021) gave bulk arguments to justify that the MF carried by the 15mn-1day waves is about what a model requires to generate a QBO. The change in region and the higher resolution of the observations could explain that the observations now give larger but more realistic MFs, but we refine the results here and suggest that the contributions from the 15mn-1hr is sufficient. Of course this result should be refined, it may well be that LMDz-6A needs larger GW drag to decrease its QBO period or increase its intensity at lower altitudes. This is ongoing work, with a priority to include a background of GWs in LG13, and to optimize the scheme parameters using the available data. We also plan to test against balloon data the schemes used by the models that have contributed to QBOi. To be more complete quantitatively in terms of MFs, it is noticeable that we have not checked the contribution of the waves with period slower than 1day that should be explicitly resolved in the model, but we suspect it is quite small simply because the LMDz-6A spatial resolution is quite coarse (for an evaluation of the large scale waves in LMDZ-6A see Maury and Lott (2014) and Holt et al. (2022)).

The second is that balloon measurements are extremely rare. Showing that they are representative of what occurs over much longer periods and over many different places suggests that they could be used, in conjunction with other products to provide much

larger datasets where GWs momentum fluxes and large scale conditions are combined. Among the datasets to consider, and on top of the satellite data (Ern et al., 2014; Alexander et al., 2021), the convection permitting global models look promising (Stephan et al., 2019), the gravity waves in the high resolution global models becoming more and more realistic (Sato et al., 1999; Shibuya & Sato, 2019). The huge datasets produced by these global models will become necessary if the tuning of GWs parameterization necessitate data assimilation techniques (Tandeo et al., 2015) and for the development of machine learning based parameterizations of GWs (Matsuoka et al., 2020; Chantry et al., 2021; Espinosa et al., 2022). For these, it seems crucial that physically based techniques can be validated against in situ data before shifting to machine learning techniques using synthetic data. We therefore plan to extend the analysis to the Loon LLC superpressure balloon data (Lindgren et al., 2020) which covers extratropical regions as well as tropical. It will permit to test, and maybe calibrate better, the orographic and frontal GWs parameterizations used in LMDz-6A (Lott, 1999; de la Cámara & Lott, 2015).

5 Open Research

Balloon data presented in (Haase et al., 2018) can be extracted from the STRATEOLE 2 dedicated web site: https://webstr2.ipsl.polytechnique.fr

ERA5 reanalysis data are described in (Hersbach et al., 2020) and can be extracted from the COPERNICUS access hub: https://scihub.copernicus.eu/

The LMDz-6A GCM used for CMIP6 project is described in (Hourdin et al., 2020), it can be directly installed from the dedicated webpage: https://lmdz.lmd.jussieu.fr/utilisateurs/installation-lmdz

Acknowledgments
This work was supported by the VESRI Schmidt Future project DataWave.

References

Alexander, M. J., Beres, J. H., & Pfister, L. (2000). Tropical stratospheric gravity wave activity and relationships to clouds. *Journal of Geophysical Research: Atmospheres*, *105*(D17), 22299-22309. doi: https://doi.org/10.1029/2000JD900326

Alexander, M. J., & Dunkerton, T. J. (1999). A Spectral Parameterization of Mean-Flow Forcing due to Breaking Gravity Waves. *J. Atmos. Sci.*, *56*(24), 4167–4182. doi: 10.1175/1520-0469(1999)056⟨4167:ASPOMF⟩2.0.CO;2

Alexander, M. J., Geller, M., McLandress, C., Polavarapu, S., Preusse, P., Sassi, F., ... Watanabe, S. (2010). Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models. *Q. J. R. Meteorol. Soc.*, *136*, 1103-1124. doi: https://doi.org/10.1002/qj.637

Alexander, M. J., Liu, C. C., Bacmeister, J., Bramberger, M., Hertzog, A., & Richter, J. H. (2021). Observational validation of parameterized gravity waves from tropical convection in the whole atmosphere community climate model. *Journal of Geophysical Research: Atmospheres*, *126*(7), e2020JD033954. (e2020JD033954 2020JD033954) doi: https://doi.org/10.1029/2020JD033954

Amemiya, A., & Sato, K. (2016). A new gravity wave parameterization including three-dimensional propagation. *Journal of the Meteorological Society of Japan. Ser. II*, *94*(3), 237-256. doi: 10.2151/jmsj.2016-013

Anstey, J. A., Banyard, T. P., Butchart, N., Coy, L., Newman, P. A., Osprey, S., & Wright, C. J. (2021). Prospect of increased disruption to the QBO in a changing climate. *Geophysical Research Letters*, *48*(15), e2021GL093058. doi:

https://doi.org/10.1029/2021GL093058

Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., ... Takahashi, M. (2001). The quasi-biennial oscillation. *Rev. Geophys.*, *39*(2), 179-229. doi: 10.1029/1999RG00007

Beres, J. H., Garcia, R. R., Boville, B. A., & Sassi, F. (2005). Implementation of a gravity wave source spectrum parameterization dependent on the properties of convection in the whole atmosphere community climate model (waccm). *Journal of Geophysical Research: Atmospheres*, *110*(D10). doi: https://doi.org/10.1029/2004JD005504

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., ... Vuichard, N. (2020). Presentation and evaluation of the ipsl-cm6a-lr climate model. *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS002010. doi: https://doi.org/10.1029/2019MS002010

Bushell, A. C., Anstey, J. A., Butchart, N., Kawatani, Y., Osprey, S. M., Richter, J. H., ... Yukimoto, S. (2022). Evaluation of the quasi-biennial oscillation in global climate models for the SPARC-QBO initiative. *Quarterly Journal of the Royal Meteorological Society*, *148*(744), 1459-1489. doi: https://doi.org/10.1002/qj.3765

Bushell, A. C., Butchart, N., Derbyshire, S. H., Jackson, D. R., Shutts, G. J., Vosper, S. B., & Webster, S. (2015). Parameterized gravity wave momentum fluxes from sources related to convection and large-scale precipitation processes in a global atmosphere model. *Journal of the Atmospheric Sciences*, *72*(11), 4349–4371. doi: https://doi.org/10.1175/JAS-D-15-0022.1

Butchart, N., Anstey, J. A., Hamilton, K., Osprey, S., McLandress, C., Bushell, A. C., ... Yukimoto, S. (2018). Overview of experiment design and comparison of models participating in phase 1 of the sparc quasi-biennial oscillation initiative ("qboi"). *Geoscientific Model Development*, *11*(3), 1009–1032. doi: 10.5194/gmd-11-1009-2018

Chantry, M., Hatfield, S., Dueben, P., Polichtchouk, I., & Palmer, T. (2021). Machine learning emulation of gravity wave drag in numerical weather forecasting. *Journal of Advances in Modeling Earth Systems*, *13*(7), e2021MS002477. doi: https://doi.org/10.1029/2021MS002477

Charron, M., & Manzini, E. (2002). Gravity waves from fronts: Parameterization and middle atmosphere response in a general circulation model. *Journal of the Atmospheric Sciences*, *59*(5), 923 - 941. doi: 10.1175/1520-0469(2002)059⟨0923:GWFFPA⟩2.0.CO;2

Corcos, M., Hertzog, A., Plougonven, R., & Podglajen, A. (2021). Observation of gravity waves at the tropical tropopause using superpressure balloons. *Journal of Geophysical Research: Atmospheres*, *126*(15), e2021JD035165. doi: https://doi.org/10.1029/2021JD035165

de la Cámara, A., & Lott, F. (2015). A parameterization of gravity waves emitted by fronts and jets. *Geophys. Res. Lett.*, *42*(6), 2071-2078. doi: 10.1002/2015GL063298

de la Cámara, A., Lott, F., & Hertzog, A. (2014). Intermittency in a stochastic parameterization of nonorographic gravity waves. *J. Geophys. Res.: Atmospheres*, *119*(21), 11905-11919. doi: 10.1002/2014JD022002

de la Cámara, A., Lott, F., Jewtoukoff, V., Plougonven, R., & Hertzog, A. (2016). On the gravity wave forcing during the southern stratospheric final warming in LMDZ. *J. Atmos. Sci.*, *73*(8), 3213-3226. doi: https://doi.org/10.1175/JAS-D-15-0377.1

Ern, M., Ploeger, F., Preusse, P., Gille, J., Gray, L. J., Kalisch, S., ... Riese, M. (2014). Interaction of gravity waves with the QBO: A satellite perspective. *Journal of Geophysical Research: Atmospheres*, *119*, 2329 - 2355. doi: https://doi.org/10.1002/2013JD020731

Espinosa, Z. I., Sheshadri, A., Cain, G. R., Gerber, E. P., & DallaSanta, K. J.

(2022). Machine learning gravity wave parameterization generalizes to capture the QBO and response to increased co2. *Geophysical Research Letters*, *49*(8), e2022GL098174. doi: https://doi.org/10.1029/2022GL098174

Fovell, R., Durran, D., & Holton, J. R. (1992). Numerical simulations of convectively generated stratospheric gravity waves. *Journal of Atmospheric Sciences*, *49*(16), 1427 - 1442. doi: 10.1175/1520-0469(1992)049⟨1427:NSOCGS⟩2.0.CO;2

Garfinkel, C. I., Gerber, E. P., Shamir, O., Rao, J., Jucker, M., White, I., & Paldor, N. (2022). A QBO cookbook: Sensitivity of the quasi-biennial oscillation to resolution, resolved waves, and parameterized gravity waves. *Journal of Advances in Modeling Earth Systems*, *14*(3), e2021MS002568. doi: https://doi.org/10.1029/2021MS002568

Geller, M. A., Alexander, M. J., Love, P. T., Bacmeister, J., Ern, M., Hertzog, A., ... Zhou, T. (2013). A comparison between gravity wave momentum fluxes in observations and climate models. *J. Atmos. Sci.*, *26*(17).

Haase, J. S., Alexander, M. J., Hertzog, A., Kalnajs, L. E., Deshler, T., Davis, S. M., ... Venel, S. (2018). Around the world in 84 days [Dataset]. *Eos*, *99*. doi: https://doi.org/10.1029/2018EO091907

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Hornyi, A., Muoz-Sabater, J., ... Thpaut, J.-N. (2020). The ERA5 global reanalysis [Dataset]. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049. doi: https://doi.org/10.1002/qj.3803

Hertzog, A. (2007). The stratole-vorcore long-duration balloon experiment: A personal perspective. *Space Research Today*, *169*, 43-48. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1752929807800478` doi: https://doi.org/10.1016/S1752-9298(07)80047-8

Hertzog, A., Alexander, M. J., & Plougonven, R. (2012). On the Intermittency of Gravity Wave Momentum Flux in the Stratosphere. *Journal of the Atmospheric Sciences*(11), 3433–3448. doi: 10.1175/JAS-D-12-09.1

Holt, L. A., Lott, F., Garcia, R. R., Kiladis, G. N., Cheng, Y.-M., Anstey, J. A., ... Yukimoto, S. (2022). An evaluation of tropical waves and wave forcing of the QBO in the QBOi models. *Quarterly Journal of the Royal Meteorological Society*, *148*(744), 1541-1567. doi: https://doi.org/10.1002/qj.3827

Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., ... Ghattas, J. (2020). LMDZ6A: The atmospheric component of the ipsl climate model with improved and better tuned physics [Software]. *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS001892. doi: https://doi.org/10.1029/2019MS001892

Jewtoukoff, V., Hertzog, A., Plougonven, R., de la Cámara, A., & Lott, F. (2015). Comparison of gravity waves in the southern hemisphere derived from balloon observations and the ecmwf analyses. *J. Atmos. Sci.*, *72*(9). doi: DOI:10.1175/JAS-D-14-0324.1

Jewtoukoff, V., Plougonven, R., & Hertzog, A. (2013). Gravity waves generated by deep tropical convection: Estimates from balloon observations and mesoscale simulations. *Journal of Geophysical Research: Atmospheres*, *118*(17), 9690-9707. doi: https://doi.org/10.1002/jgrd.50781

Kang, M.-J., Chun, H.-Y., & Kim, Y.-H. (2017). Momentum flux of convective gravity waves derived from an offline gravity wave parameterization. part i: Spatiotemporal variations at source level. *Journal of the Atmospheric Sciences*, *74*(10), 3167 - 3189. doi: 10.1175/JAS-D-17-0053.1

Lane, T. P., & Moncrieff, M. W. (2008). Stratospheric gravity waves generated by multiscale tropical convection. *J. Atmos. Sci.*, *65*, 2598–2614. doi: DOI:10.1175/2007JAS2601.1

Lindgren, E. A., Sheshadri, A., Podglajen, A., & Carver, R. W. (2020). Seasonal and latitudinal variability of the gravity wave spectrum in the lower

stratosphere. *Journal of Geophysical Research: Atmospheres*, *125*(18), e2020JD032850. doi: https://doi.org/10.1029/2020JD032850

Liu, C., Alexander, J., Richter, J., & Bacmeister, J. (2022). Using trmm latent heat as a source to estimate convection induced gravity wave momentum flux in the lower stratosphere. *Journal of Geophysical Research: Atmospheres*, *127*(1), e2021JD035785. (e2021JD035785 2021JD035785) doi: https://doi.org/10.1029/2021JD035785

Lott, F. (1999). Alleviation of stationary biases in a GCM through a mountain drag parameterization scheme and a simple representation of mountain lift forces. *Monthly Weather Review*, *127*(5), 788 - 801. doi: 10.1175/1520-0493(1999)127⟨0788:AOSBIA⟩2.0.CO;2

Lott, F., & Guez, L. (2013). A stochastic parameterization of the gravity waves due to convection and its impact on the equatorial stratosphere. *J. Geophys. Res.*, *118*(16), 8897-8909. doi: 10.1002/jgrd.50705

Lott, F., Guez, L., & Maury, P. (2012). A stochastic parameterization of non-orographic gravity waves: Formalism and impact on the equatorial stratosphere. *Geophys. Res. Lett.*, *39*(6), L06807. doi: 10.1029/2012GL051001

Matsuoka, D., Watanabe, S., Sato, K., Kawazoe, S., Yu, W., & Easterbrook, S. (2020). Application of deep learning to estimate atmospheric gravity wave parameters in reanalysis data sets. *Geophysical Research Letters*, *47*(19), e2020GL089436. doi: https://doi.org/10.1029/2020GL089436

Maury, P., & Lott, F. (2014). On the presence of equatorial waves in the lower stratosphere of a general circulation model. *Atmospheric Chemistry and Physics*, *14*(4), 1869–1880. Retrieved from `https://acp.copernicus.org/articles/14/1869/2014/` doi: 10.5194/acp-14-1869-2014

Palmer, T. N., Shutts, G. J., & Swinbank, R. (1986). Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization. *Quarterly Journal of the Royal Meteorological Society*, *112*(474), 1001-1039. doi: 10.1002/qj.49711247406

Plougonven, R., Jewtoukoff, V., de la Cámara, A., Lott, F., & Hertzog, A. (2017). On the relation between gravity waves and wind speed in the lower stratosphere over the southern ocean. *J. Atmos. Sci.*, *74*(4), 1075-1093. doi: 10.1175/JAS-D-16-0096.1

Rabier, F., Bouchard, A., Brun, E., Doerenbecher, A., Guedj, S., Guidard, V., ... Steinle, P. (2010, January). The Concordiasi Project in Antarctica. *Bulletin of the American Meteorological Society*, *91*(1), 69-86. Retrieved from `https://hal-insu.archives-ouvertes.fr/insu-00562459` doi: 10.1175/2009BAMS2764.1

Richter, J. H., Sassi, F., & Garcia, R. R. (2010). Toward a physically based gravity wave source parameterization in a general circulation model. *Journal of the Atmospheric Sciences*, *67*(1), 136 - 156. doi: 10.1175/2009JAS3112.1

Sato, K., Kumakura, T., & Takahashi, M. (1999). Gravity waves appearing in a high-resolution GCM simulation. *Journal of the Atmospheric Sciences*, *56*(8), 1005 - 1018. doi: https://doi.org/10.1175/1520-0469(1999)056⟨1005:GWAIAH⟩2.0.CO;2

Shibuya, R., & Sato, K. (2019, 03). A study of the dynamical characteristics of inertiagravity waves in the antarctic mesosphere combining the pansy radar and a non-hydrostatic general circulation model. *Atmospheric Chemistry and Physics*, *19*, 3395-3415. doi: 10.5194/acp-19-3395-2019

Song, I.-S., & Chun, H.-Y. (2005). Momentum flux spectrum of convectively forced internal gravity waves and its application to gravity wave drag parameterization. part i: Theory. *J. Atmos. Sci.*, *62*(1), 107-124. doi: https://doi.org/10.1175/JAS-3363.1

Stephan, C. C., Strube, C., Klocke, D., Ern, M., Hoffmann, L., Preusse, P., &

Schmidt, H. (2019). Intercomparison of gravity waves in global convection-permitting models. *Journal of the Atmospheric Sciences*, *76*(9), 2739 - 2759. doi: 10.1175/JAS-D-19-0040.1

Tandeo, P., Pulido, M., & Lott, F. (2015). Offline parameter estimation using enkf and maximum likelihood error covariance estimates: Application to a subgrid-scale orography parametrization. *Quarterly Journal of the Royal Meteorological Society*, *141*(687), 383-395. doi: https://doi.org/10.1002/qj.2357