Increased vertical resolution in the stratosphere reveals role of gravity waves after sudden stratospheric warmings

Wolfgang Wicker[1], Inna Polichtchouk[2], and Daniela I. V. Domeisen[1,3]

[1]Université de Lausanne, Lausanne, Switzerland
[2]European Centre for Medium-Range Weather Forecasts, Reading, UK
[3]ETH Zurich, Zurich, Switzerland

Correspondence: Wolfgang Wicker (wolfgang.wicker@unil.ch)

Abstract. Sudden stratospheric warmings (SSW) have a long-lasting effect within the stratosphere as well as impacts on the underlying troposphere. However, sub-seasonal forecasts of the winter polar stratosphere fail to use their full potential for predictability as they tend to underestimate the magnitude and persistence of these events already within the stratosphere. The origin of this underestimation is unknown. Here, we demonstrate that **the associated polar** stratospheric cold bias following SSW events in sub-seasonal hindcasts can be halved by increasing vertical model resolution, suggesting a potential sensitivity to gravity wave forcing. While the predictability of the planetary Rossby wave flux into the stratosphere at lead times longer than a week is limited, the existence of a critical layer for gravity waves with a small zonal phase speed caused by the disturbed polar vortex provides predictability to the upper stratosphere. Gravity wave breaking near that critical layer can, therefore, decelerate the zonal flow consistently with anomalous subsidence over the polar cap leading to warmer temperatures in the middle polar stratosphere. Since the spectrum of gravity waves involves vertical wavelengths of less than 4000 m, as estimated by wavelet analysis, a high vertical model resolution is needed to resolve the positive feedback between gravity wave forcing and the state of the polar vortex. Specifically, we find that at a spectral resolution of TCo639 (approximate horizontal grid-spacing of 18 km) at least 198 levels are needed to correctly resolve the spectrum of gravity waves in the ECMWF Integrated Forecasting System. Increasing vertical resolution in operational forecasts will help to mitigate stratospheric temperature biases and improve sub-seasonal predictions of the stratospheric polar vortex.

1   Introduction

Atmospheric waves entering the extra-tropical stratosphere from below comprise a wide range of spatial scales. The large-scale part of the spectrum is occupied by Rossby waves with zonal wavenumbers 1-3, often termed planetary waves. Depending on their phase speed and the zonal-mean zonal background wind, planetary waves can propagate from the troposphere into the stratosphere where they break, deposit easterly momentum, and slow down the polar vortex (e.g. Charney and Drazin, 1961; Domeisen et al., 2018). A particularly strong and sustained planetary wave flux can lead to a complete breakdown of the polar vortex, a reversal of the zonal-mean westerlies, and a rapid warming of the polar stratosphere which is termed a sudden stratospheric warming (SSW) (Matsuno, 1971; Baldwin et al., 2021). While some SSWs are associated with Rossby wave

reflection and have no long-term impact (Kodera et al., 2016), the majority of events decay slowly over the course of a month
25 or longer (Limpasuvan et al., 2004).

Gravity waves with horizontal wavelengths from $\sim$10 to $\sim$1000 km and frequencies between the Brunt-Väisälä frequency N and the inertial frequency f form the small-scale part of the spectrum (e.g. Fritts and Alexander, 2003). Typically, extra-tropical gravity waves are excited near the surface by flow over orography or in the upper troposphere by jet/front imbalances. **These waves commonly propagate via the stratosphere into the mesosphere, where their amplitudes grow until the waves** 30 **break. Depending on the phase speed of the waves and the velocity of the background wind, one can define a critical layer where the intrinsic frequency of the waves $\hat{\omega}$ would approach the inertial frequency f and the vertical wavelength would approach zero (e.g. Fritts and Alexander, 2003). If such a critical layer is present, gravity waves will break somewhere below that level and deposit more momentum already in the stratosphere.** Compared to planetary waves, gravity waves receive less attention in extra-tropical stratosphere studies. However, both observational and modeling studies 35 document increased gravity wave amplitudes at the edge of the polar vortex during minor stratospheric warmings concurrently with the peak of the planetary wave flux (e.g. Duck et al., 1998; Venkat Ratnam et al., 2004; Wang and Alexander, 2009; Yamashita et al., 2010; Dörnbrack et al., 2018; Polichtchouk and Scott, 2020). During major warmings, the polar vortex breaks down and the downward propagation of the zero-wind line prevents the propagation of stationary gravity waves into the upper stratosphere and mesosphere (e.g Wang and Alexander, 2009; Hindley et al., 2020). The absence of gravity wave breaking 40 in the mesosphere explains the mesospheric cooling during an SSW by a relaxation to radiative equilibrium (Holton, 1983). Moreover, during Southern Hemisphere spring-time polar vortex breakdown, gravity waves are found to contribute significantly to the polar vortex deceleration in ERA5 reanalysis (Gupta et al., 2021). While recent studies consider the role of gravity waves in pre-conditioning the vortex for planetary waves before the onset of vortex split events (Albers and Birner, 2014; Song et al., 2020; Kuchar et al., 2022), the present study focuses on the effect of gravity waves on the zonal-mean zonal momentum after 45 the central date of an SSW.

Due to their downward response (Baldwin and Dunkerton, 2001), SSWs hold great potential for sub-seasonal to seasonal prediction (Domeisen et al., 2020a), and an accurate simulation and representation of the winter stratosphere in weather and climate models is important for understanding and predicting tropospheric impacts of stratospheric variability. **While SSWs can be skillfully predicted up to two weeks ahead (e.g. Tripathi et al., 2015a; Domeisen et al., 2020b), their magnitude** 50 **is generally underestimated in sub-seasonal prediction models (e.g. Karpechko et al., 2018; Lawrence et al., 2022). This is probably at least in part due to an underestimation of the upward planetary wave flux (Wu et al., 2022). On the other hand, a connection to misrepresented gravity wave momentum fluxes is plausible as well. One potential model adjustment to improve sub-seasonal prediction is therefore an increase in vertical resolution.**

Generally, vertical resolution receives less attention than horizontal resolution in atmospheric modeling studies but it is 55 well established that vertical resolution needs to be set in consideration of its horizontal counterpart (e.g. Lindzen and Fox-Rabinovitz, 1989; Roeckner et al., 2006; Skamarock et al., 2019). The representation of Rossby waves requires a consistent aspect ratio of the vertical and horizontal grid spacing $\Delta z/\Delta L$ determined by the ratio of **the scale height** H to **the radius of deformation** L_R (Lindzen and Fox-Rabinovitz, 1989)

$$\frac{\Delta z}{\Delta L} \approx \frac{H}{L_R} = \frac{f}{N} \tag{1}$$

To resolve the regime of stratified turbulence at small horizontal scales, the vertical grid spacing depends primarily on the stratification and can be set following the ratio

$$\Delta z \approx \frac{U}{N} \tag{2}$$

of the horizontal velocity scale U to the buoyancy frequency N (Waite, 2016; Cullen, 2017). For the undisturbed polar vortex, this gives a vertical grid spacing of $\mathcal{O}(1\text{km})$ in the stratosphere. For a disrupted vortex with $U = 0\,\text{m}\,\text{s}^{-1}$, on the other hand, this relation would require infinite vertical resolution. The vertical wavelength of a gravity wave approaching its critical layer reduces to zero. According to Lindzen and Fox-Rabinovitz (1989), a requirement for infinite vertical resolution is only prevented by the presence of a complex phase speed, i.e. damping. Approximating the imaginary part of the phase speed using a linear damping rate σ_i, they postulate an aspect ratio of the vertical to the horizontal grid spacing of

$$\frac{\Delta z}{\Delta L} \approx \frac{\sigma_i}{N} \tag{3}$$

However, the damping rate σ_i can be very difficult to infer and dedicated **model experiments** are required to determine the correct settings of vertical resolution at a given horizontal resolution, as we do in the present study. Given the strong constraint for vertical resolution near a critical layer and the widely different horizontal scales (Lindzen and Fox-Rabinovitz, 1989), higher sensitivity can be expected for processes involving gravity waves than for planetary waves.

A detailed description of the model experiments is given in Section 2. In Section 3 we **quantify the** vortex state-dependent temperature bias, i.e. the underestimated SSW amplitude, in hindcasts from the S2S database and demonstrate how this bias can be mitigated **by increasing vertical model** resolution. The benefit of vertical resolution for gravity wave drag and potential energy is presented in Section 4. A discussion of these results follows in Section 5.

2   Data and methods

For this study, we conduct model experiments with the ECMWF Integrated Forecasting System (IFS CY47R3) with different vertical resolutions. The subject of these experiments is the simulation of the major SSW in boreal winter **2017/18** and its downward influence on the tropospheric circulation on sub-seasonal timescales. Most of these hindcasts are initialized on 8th February and run for 46 days. For robustness analysis, additional simulations are initialized on 17th January 2006 and 5th February 2010 that confirm our findings for a vortex displacement and another split SSW event as discussed in Appendix A. For each start date and each model configuration, we initialize an ensemble of 51 realizations. The output of individual realizations is used whenever we compute quadratic quantities such as variance, eddy-heat, or eddy-momentum fluxes. Ensemble-mean

anomalies from climatology are regarded as the predictable signal in the sense of probabilistic predictability as is commonly done for sub-seasonal prediction. Initial conditions are taken from the ERA5 reanalysis dataset (Hersbach et al., 2020), which is also the reference against which we evaluate the model experiments.

The bulk of our simulations is run at a spectral horizontal resolution of TCo639, which corresponds approximately to 18 km grid-spacing, and a vertical resolution of either 91 or 198 levels in hybrid η-coordinates with a sponge layer starting at 1 hPa. The vertical grid-spacing for the different model configurations is illustrated in Figure 1 depicting high vertical resolution near the surface and a gradually coarsening vertical grid in the stratosphere. Sensitivity tests for a comparison with operational forecasts at ECMWF reveal that while 137 vertical levels seem sufficient at a spectral resolution of TCo319 (36 km grid-spacing), at least 198 vertical levels are required at TCo639 (see Section 5 and Appendix B).

Changing resolution will presumably affect the magnitude of resolved gravity wave **drag**. However, at a horizontal resolution of approximately 18 km the stratospheric gravity wave spectrum is only partly resolved and sub-grid scale gravity waves need to be parameterized. Note that **an intercomparison of various** ECMWF IFS simulations suggests that the gravity wave spectrum is not fully resolved even at a 4 km horizontal grid-spacing (Polichtchouk et al., 2022a, b). The settings for parameterized gravity wave drag are unchanged between the different vertical resolution configurations. To verify whether changes in gravity wave drag are caused directly by changes in the vertical resolution as opposed to changes in the stratospheric zonal-mean basic state, we performed nudged simulations following the SNAPSI protocol (Hitchcock et al., 2022). Specifically, zonal-mean temperature, vorticity, and divergence above 90 hPa are relaxed towards the 'observed' time series of ERA5 on a time-scale of 6h. Consequently, there are no significant differences in the zonal-mean zonal wind and polar-cap temperature anomaly between the different model configurations in the nudged simulations.

The vertical eddy-momentum flux convergence on pressure levels $\left[-\partial \overline{u'\omega'} \,/\, \partial p \right]$, where u and ω are zonal and vertical (pressure) velocities and p is pressure, is used as the diagnostic for the zonal resolved gravity wave forcing. We follow the methodology of Polichtchouk et al. (2022b) and represent primes as all spherical harmonics with total wavenumbers between 21 and the truncation limit. Selecting only these harmonics explicitly eliminates the contributions by planetary and synoptic-scale waves. The very same horizontal filtering is used when diagnosing the vertical power spectra of potential temperature anomalies as is explained in the following. The bar denotes an average over the largest resolved wavelength.

In order to estimate vertical wavenumbers m and vertical wavelengths $2\pi/m$ we use log-pressure coordinates where the following wave ansatz can be made (Fritts and Alexander, 2003)

$$\frac{\theta'}{\overline{\theta}} = \widetilde{\theta} \cdot \exp\left[imz + \frac{z}{2H}\right] \tag{4}$$

The scale height H is set to 7000 m, θ is potential temperature, and $\widetilde{\theta} = \widetilde{\theta}(x,y,t)$ is the amplitude of a **wave mode** at a point in (x,y,t). To fully use the vertical resolution of the model, we interpolate potential temperature data from native model levels to a constant grid spacing of 100 m (marked by the vertical line in Fig. 1) instead of using standard pressure level output. Then we apply a wavelet analysis, i.e. we convolve profiles of normalized anomalies $\left[\exp\left(-z/2H\right)\left(\theta'/\overline{\theta}\right)\right]$ with a Morlet wavelet (e.g. Torrence and Compo, 1998) to compute spectral power and estimate the dominant vertical wavenumber at each altitude. The
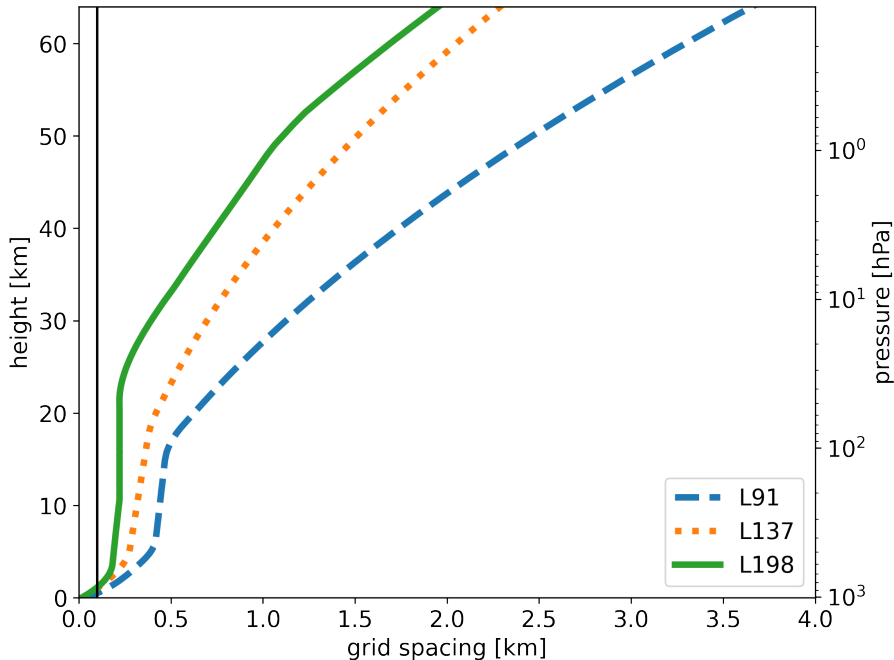
Figure 1. Vertical grid spacing in log-pressure coordinates with a scale height of 7000 m for three vertical grid configurations with standard surface pressure. The vertical line marks a constant grid spacing of 100 m.

spectral power of $\left[\exp\left(-z/2H\right)\left(\theta'/\overline{\theta}\right)\right]$ is proportional to potential energy per unit volume. To remove artifacts introduced by the lower boundary of the model domain, we apply a Hanning window tapering to the vertical column of normalized perturbations.

To confirm the state dependence of stratospheric temperature biases on the strength of the polar vortex and to demonstrate the relevance of our simulations, we analyze the extended winter (NDJFM) IFS hindcasts provided by ECMWF to the S2S database (Vitart et al., 2017). Specifically, we estimate composites of polar cap temperature biases at 50 hPa using zonal-mean zonal wind at 10 hPa and 60 °N in reanalysis during initialization as the criterion to select initialization dates and construct the composites (Tripathi et al., 2015b). The thresholds of $40\,\mathrm{m\,s^{-1}}$ for a strong vortex state and $5\,\mathrm{m\,s^{-1}}$ for a weak vortex state are chosen following Domeisen et al. (2020a) to ensure sufficiently large sample sizes. This sample of S2S hindcasts comprises more than 21,000 ensembles members for the winters 1999-2020 including 4191 members initialized during strong vortex states and 2761 members during weak vortex states. The significance of selecting the composite based on the strength of the vortex during initialization is tested using a Monte Carlo technique by comparing the observed difference with the distribution of hindcast biases for randomly selected start dates.

3 Increased vertical resolution strengthens SSW temperature anomaly in the polar stratosphere

Over the polar cap, the ECMWF S2S hindcasts develop a mid-to-lower stratospheric cold bias of more than 2 K at 50hPa with increasing lead time (dashed grey line in Fig. 2a), which is in line with the findings of Lawrence et al. (2022) for high-top models. Hindcasts with weak vortex initial conditions develop a stronger than average cold bias (dotted blue line in Fig. 2a). Specifically, the difference between the weak vortex composite mean and the climatological bias reaches its maximum of about -0.5 K after 20 days into the forecast before the influence of the initial conditions becomes less important and an increasing number of composite members reach a neutral or strong vortex state (Fig. 2b). This aggravated cold bias corresponds to an underestimated magnitude of weak vortex events. Similarly, Lawrence et al. (2022) find that at a lead time of one week, the wind change associated with an SSW is underestimated by approximately 5 $\mathrm{m\,s^{-1}}$. In contrast to weak vortex events, the strong vortex composite mean is not significantly different from the climatological bias for the first two weeks of the hindcast until horizontal model resolution is reduced and the temperature bias diminishes by 0.25 K compared to the average (Fig. 2, see Appendix B for a discussion of the model resolution in S2S hindcasts). A comparison with the Monte Carlo distribution in Figure 2b shows that the observed difference lies outside of the 99.9% significance which confirms that the null hypothesis can be rejected and that the cold bias is significantly aggravated following weak vortex conditions. The existence of a critical layer in the upper stratosphere during weak vortex states and underestimated resolved gravity wave forcing due to insufficient vertical resolution offer one potential explanation for the vortex state dependent bias in the S2S hindcasts.

The analysis of a case study for the 2018 boreal winter SSW can be fruitful for a better understanding of the mechanisms that might cause biases in the large ensemble of hindcasts. The central date of an SSW in reanalysis is typically preceded and followed by a downward propagation of positive polar cap temperature anomalies on the order of 15 K to 20 K from the stratopause to the lower stratosphere, as shown in Figure 3a for the 2018 event. The recovery of the polar vortex is marked by the downward propagation of negative temperature anomalies from the mesosphere to the mid-stratosphere in the weeks following the event. In the lower stratosphere, on the other hand, positive anomalies of about 5 K are sustained for at least six weeks after the 2018 event. While the general development of the SSW is reproduced in sub-seasonal hindcasts, the magnitude of sustained positive temperature anomalies is underestimated by 2 K to 4 K (Fig. 3b). Increasing vertical resolution from 91 levels to 198 levels reduces this negative bias significantly and prolongs the stratospheric warm temperature anomalies in the ensemble mean (Fig. 3c). By comparing these results with simulations of SSW events in 2006 and 2010 (Fig. A1, A2) it is confirmed that the strengthening and the extension of the stratospheric warming signal in time by increased vertical resolution is robust amongst different SSW events. While the stratospheric signal shows a strong resemblance between different SSW events, the timing and character of the downward response can be very different. So, despite the importance of the stratospheric signal for tropospheric predictability following an SSW (e.g. Domeisen et al., 2020a; Kautz* et al., 2020), **no robust improvement of surface predictability by increased vertical resolution can be inferred from the simulation of these three SSW events.**

An evident hypothesis to explain the reduction of the cold bias might be that the planetary wave flux into the stratosphere is improved by increased vertical resolution. The meridional eddy-heat flux at 100 hPa shown in Figure 3d is used as a proxy for the vertical Rossby wave **propagation** in the lower stratosphere. Between Feb 08 and Feb 12, shortly after initialization
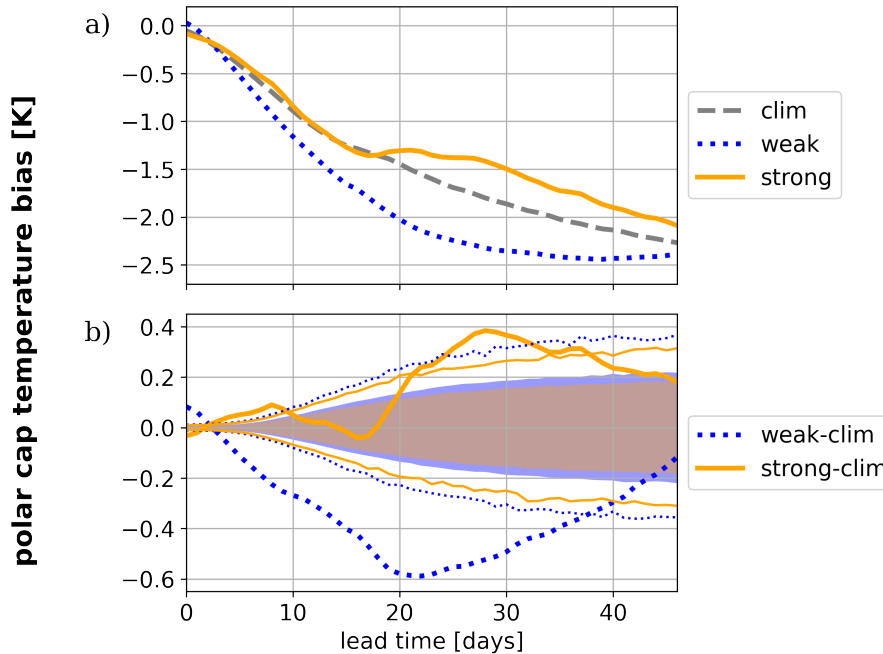
Figure 2. Polar cap temperature bias (hindcasts - reanalysis) at 50 hPa horizontally averaged between 60 °N and 90 °N of ECMWF S2S hindcasts plotted against lead time. Panel a) shows as a dashed gray line the average over all hindcasts ("clim"), as a dotted blue line the composite mean of hindcasts initialized during a weak vortex state ("weak"), and as a solid orange line the composite mean of hindcasts initialized during a strong vortex state ("strong"). Panel b) compares the difference between the "weak" composite bias (thick dashed blue) or the "strong" composite bias (thick solid orange) and the climatological bias with the 2.5% and 97.5% (shading) and the 0.05% and 99.95% (thin lines) percentiles of the Monte Carlo distribution of randomly selected composite means. The percentiles depend on the size of random samples and are shown in blue and orange for the "weak" and "strong" composites, respectively.

of the hindcast, both reanalysis and the ensemble simulations exhibit a pronounced peak that is three times larger than the climatological average. Capturing this peak is crucial for predicting the onset of an SSW. Around Feb 20, reanalysis shows a second peak in planetary wave flux, which both hindcasts fail to predict and which likely contributes to the forecast error in the stratosphere. After the initial peak, the ensemble mean eddy-heat flux shows little difference from the climatological average (Fig. 3d), indicating limited predictability. In particular, there is no significant difference in heat flux between the model configurations with high and low vertical resolution. Therefore, we conclude that planetary waves do not play the predominant role in maintaining positive temperature anomalies in the later stage of the SSW and explaining the ensemble-mean temperature difference between the different model configurations. This leads to the hypothesis that the improvement in the temperature bias is rather due to the sensitivity of gravity waves to the vertical resolution. Hence, the remainder of this study focuses on the representation of gravity wave forcing in the model hindcasts.
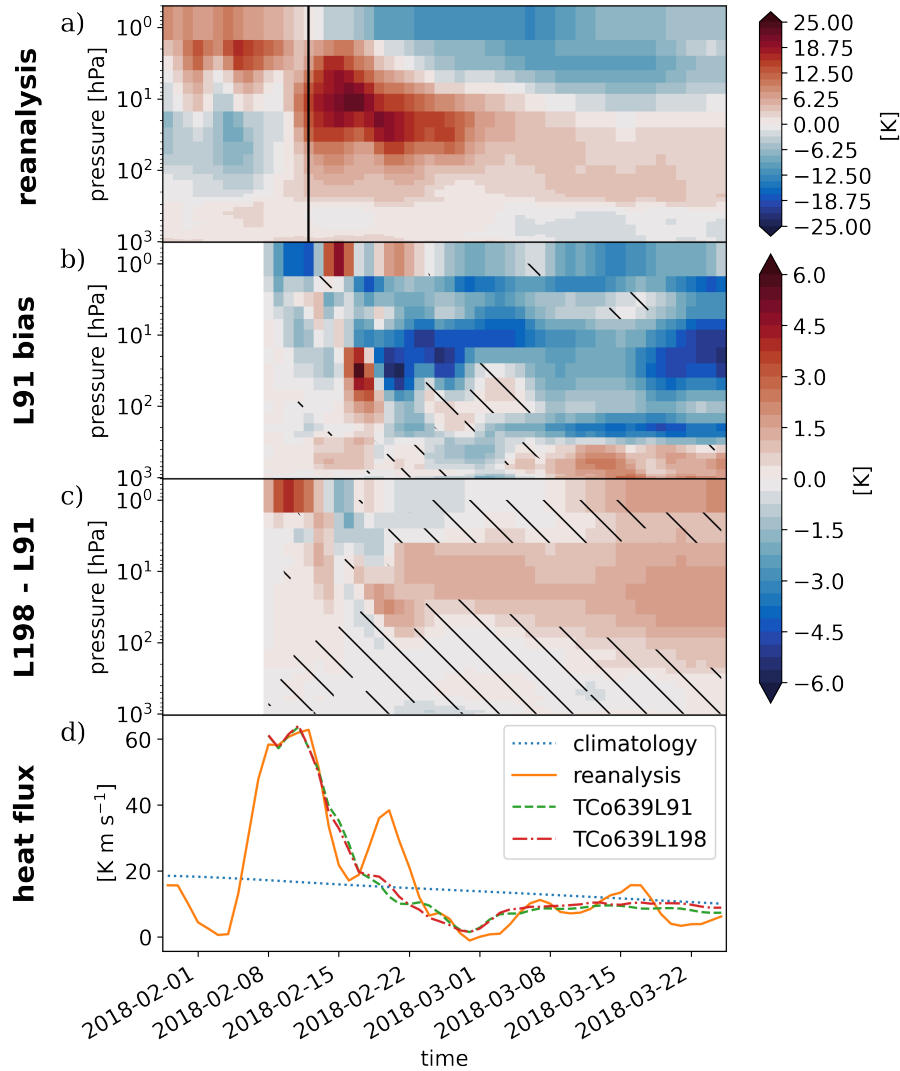
Figure 3. Polar cap temperature anomalies averaged between 60 °N and 90 °N during the 2018 SSW event. The central date of the event is marked by the reversal of zonal mean westerlies at 10 hPa and 60 °N indicated by the vertical line. (a) The development of temperature anomalies [K] in reanalysis, (b) the ensemble mean bias of TCo639L91 hindcasts initialized on Feb 08 compared to reanalysis, and (c) the improvement of the TCo639L198 hindcast ensemble mean compared to the TCo639L91 hindcasts. Hatching indicates areas where ensemble-mean differences are not significantly different from zero at a 95% confidence level estimated by a one-sample (b) or two-sample t-test (c). Panel d) shows the ensemble-mean meridional eddy heat flux by zonal wavenumbers 1 to 3 at 100 hPa averaged between **45** °N and **70** °N as a measure of the planetary wave flux in the lower stratosphere for climatology (dashed blue), the event in reanalysis (orange), and the two model configurations (green and red).

4   Stratospheric drag by gravity waves with small vertical wavelengths

Having established a sensitivity of the large-scale flow measured in terms of polar cap temperature anomalies to changing vertical model resolution, we are now looking for a mechanism to explain this sensitivity. Small-scale gravity waves can affect the large-scale flow via non-linear processes such as wave breaking. The force on the zonal-mean flow exerted by the resolved part of the gravity wave spectrum can be diagnosed from the velocity output of the model in the form of the vertical eddy-momentum flux convergence (see Section 2). The force exerted by subgrid-scale gravity waves is output separately as accumulated momentum tendencies for orographic and non-orographic waves. The sum of those three respective forces is shown in Figure 4**a** where positive values mean an eastward acceleration of the zonal wind.

For an undisturbed vortex, gravity waves propagate into the mesosphere before they dissipate due to the reduced density (not shown). The disrupted vortex during the SSW, on the other hand, is further decelerated by gravity waves that dissipate in the mid-stratosphere below 5 hPa (Fig. 4**a**). Above the center of upper-stratospheric easterlies during the major warming, the momentum tendency has positive values of more than $2 \, \mathrm{m \, s^{-1} \, day^{-1}}$, which corresponds to an acceleration of the disrupted vortex in the lower mesosphere. After the decay of upper-stratospheric easterlies around Feb 28, the mesosphere slowly returns to values of about -1 $\mathrm{m \, s^{-1} \, day^{-1}}$, indicative of gravity waves with **low negative** phase speed reaching the mesosphere again **(the vertical profile of the resolved gravity wave momentum flux $(-g)^{-1}\overline{u'\omega'}$ is shown in Figure 4b, where g denotes the gravitational acceleration and negative values indicate an upward flux of easterly momentum)**. As noted in Section 1, the existence of a **breaking layer** near the zero-wind line during SSWs has been reported in observational and modeling studies (e.g. Wang and Alexander, 2009; Yamashita et al., 2010). In the present hindcasts, however, westward drag in the mid-stratosphere is sustained after the return of the vortex **from easterly to weak westerly winds**.

As explained in Section 1 gravity waves can be expected to be sensitive to vertical resolution. To quantify this sensitivity, we investigate the difference in ensemble-mean gravity wave drag between the model configurations with high and low vertical resolution (Fig. 5). To verify that the expected change in resolved gravity wave drag with higher resolution is not compensated for by a change in parameterized momentum tendencies, the wave drag difference is split into its three components. The deceleration of the upper-stratospheric flow by resolved gravity wave drag in the simulations with higher resolution is continuously stronger **than with low resolution** by up to $0.2 \, \mathrm{m \, s^{-1} \, day^{-1}}$, which is consistent with increased temperature anomalies in the mid-stratosphere compared to the configuration with low resolution (Fig. 5a). This difference in resolved gravity wave drag is not offset by parameterized waves. On the contrary, **parameterized** non-orographic gravity wave drag adds to the enhanced deceleration of the zonal-mean flow in the upper stratosphere (Fig. 5b), while the **parameterized** orographic gravity wave drag shows a difference only in the lower stratosphere (Fig.5c). The combined difference in Figure 5**a-c** represents almost 100% of full gravity wave drag at 10 hPa in the hindcasts with 198 vertical levels seen in Figure 4**a**. **Note that given the noisy nature of the troposphere, no statistically significant difference in resolved gravity wave momentum flux is found below 70hPa. Hence, we conclude that the sensitivity in wave drag is caused by the gravity wave propagation characteristics in the stratosphere, and not by source processes in the troposphere.** Also note that these results for the SSW in February 2018 are confirmed by an analysis of additional SSW events, which show a very similar behaviour in terms of absolute values and the
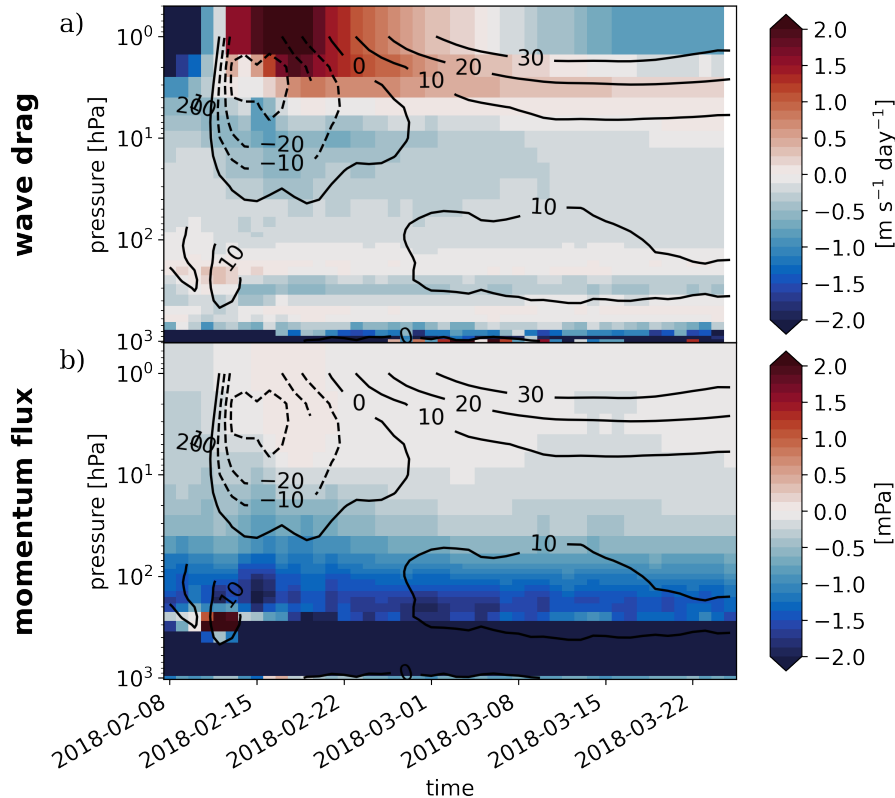
Figure 4. TCo639L198 ensemble mean of (a) full gravity wave drag (resolved + orographic (parameterized) + non-orographic (parameterized)) **and (b) resolved gravity wave momentum flux** horizontally averaged between **45 °N** and **70 °N** with contours of zonal-mean zonal wind $[\mathrm{m\,s^{-1}}]$ at 60 °N.

difference between model configurations (Fig. A3). Sensitivity tests with a different set of horizontal and vertical resolution are presented in Appendix B.

An important question when inferring whether gravity waves are responsible for the reduced temperature bias with increased vertical resolution is whether changes in gravity wave drag result directly from changes in resolution or whether they are a consequence of a modified background state. To assess this question, the diagnostics of Figure 5 have been repeated for the simulations where the mid- and upper-stratospheric zonal mean state is nudged to reanalysis (Fig. 6). In these nudged simulations, the differences between the two model configurations inevitably result from increased vertical resolution since the stratospheric background state is unchanged. It is found that the sensitivity of resolved and parameterized wave drag does not depend on a difference in the zonal-mean background state, which supports the above hypothesis.

To better understand how the resolved gravity wave momentum flux depends on vertical resolution we aim to estimate the dominant vertical length scale resolved by the model. Therefore, we compute wavelet spectra of gravity wave potential energy
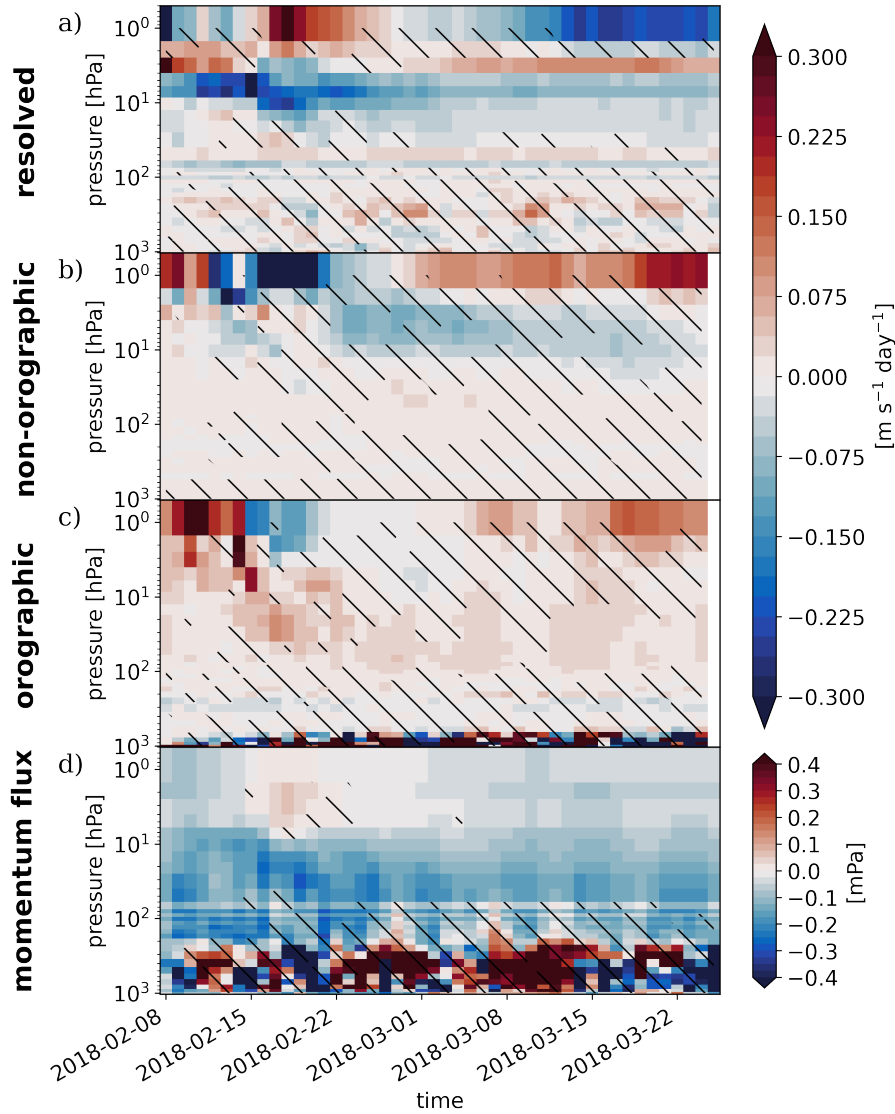
Figure 5. Difference in ensemble-mean gravity wave drag between the TCo639L198 and TCo639L91 free running hindcasts split into the (a) resolved, (b) **parameterized** non-orographic, (c) **parameterized** orographic components, **and (d) the difference in ensemble-mean resolved gravity wave momentum flux between the two sets of hindcasts**. Wave drag **and momentum flux are** horizontally averaged between **45** °N and **70** °N. Hatching indicates areas where the improvement with higher vertical resolution is not significantly different from zero at a 95% confidence level estimated by a two-sample t-test.

from filtered temperature data as explained in Section 2. Note that gravity wave potential energy E_{pot} is connected to the absolute value of the vertical flux of horizontal momentum F according to the following identity (Ern et al., 2004)
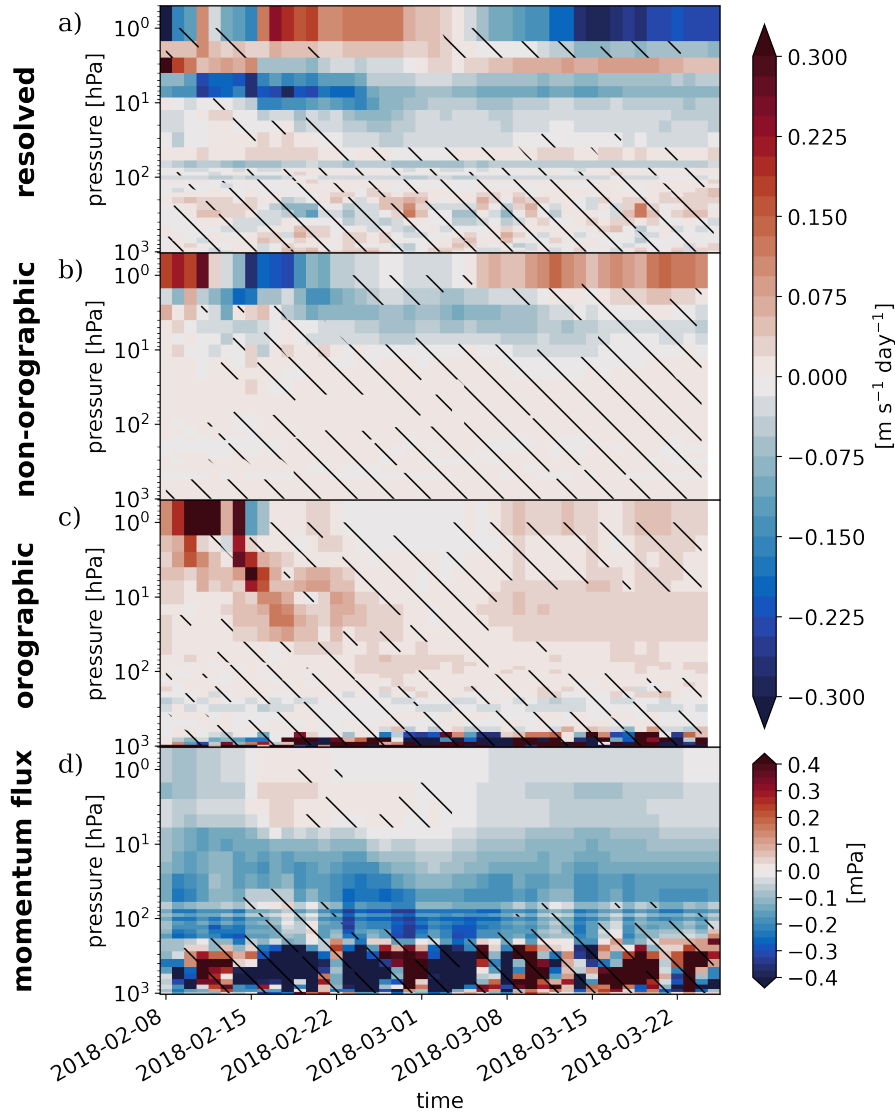
Figure 6. Same as Figure 5 but for the nudged simulations.

$$F = \frac{k_h}{m} E_{pot}. \tag{5}$$

where k_h and m are the total horizontal and the vertical wavenumber. The model configuration with 198 vertical levels shows

225   peak potential energy at a height of approximately 40 hPa for vertical wavelengths of about 6000 m, while the energy in the model version with 91 vertical levels peaks already at 150 hPa and 8000 m wavelength (Fig. 7a, b). The overall scale of vertical wavelengths is roughly in line with observational estimates (e.g. Sato, 1994; Preusse et al., 2002). The global minimum

of spectral power in the mesosphere above 40 km height can readily be explained by the gradually increasing vertical grid spacing as we go higher up into the middle atmosphere (Fig. 1), the influence of the sponge layer at the model top, and the existence of a critical layer in the stratosphere. In the lower stratosphere there is a sign of a local minimum, potentially indicative of a reflecting or refracting layer directly above the tropopause. A similar minimum can be observed in the gravity wave **drag** at around 100 hPa (Fig. 4a). However, the informative value of spectral power below 100 hPa in Figure 7 is reduced for two reasons. First, there are boundary effects caused by the convolution of the vertical column with a Morlet wavelet marked by the cone of influence (hatched area). Second, the potential temperature anomalies are tapered by multiplication with a Hanning window prior to the wavelet transform to reduce spectral leakage at the boundaries which marks a trade-off between removing artifacts and damping the signal. These limitations induced by the computation of the wavelet analysis are not easily overcome.

The difference of potential energy spectra between model configurations with high and low vertical resolution (Fig. 7c) shows a substantial increase of power with high resolution for vertical wavelengths of 4000 m and less in the mid- to upper stratosphere. The magnitude of this increase corresponds to more than 50% of the peak spectral energy in the configuration with 91 model levels. Given the small spread between individual ensemble members of the same model configuration, the difference of potential energy spectra is highly significant almost everywhere. The wavelength for the strongest increase in power corresponds nicely to the difference in vertical grid spacing (Fig. 1), indicating that the difference in gravity wave drag seen in Figures 5, 6 does indeed result from gravity waves with small vertical wavelengths.

5 Conclusions

Previous studies (e.g. Karpechko et al., 2018; Lawrence et al., 2022) and a composite analysis in Section 3 reveal that the magnitude of stratospheric polar-cap temperature anomalies **following** an SSW is commonly underestimated in sub-seasonal prediction models. In order to better understand this underestimation, this study investigates the means of improving the prediction by increasing vertical model resolution. At a vertical resolution of 91 levels, the ensemble-mean error in targeted simulations of the SSW in 2018 reaches up to 4 K compared to reanalysis. By increasing vertical resolution from 91 to 198 vertical levels, the error is reduced by 50%, prolonging positive temperature anomalies in the polar stratosphere following SSW events. We here show that the most likely explanation for the strengthened and more persistent stratospheric warming signal is an improved representation of gravity wave dynamics with a higher number of vertical levels.

The present simulations are initialized four days prior to the onset of the SSW on February 12th, 2018. The model accurately predicts the strong flux of planetary waves into the stratosphere that caused the reversal of the prevalent westerlies. The associated peak in planetary wave drag is one to two orders of magnitude larger than the persistent deceleration of the zonal-mean flow that is induced by resolved and parameterized gravity waves. However, the planetary wave flux in the weeks following the SSW has only limited predictability and no significant difference between the model configurations with different vertical resolution was found. In comparison to planetary waves, the ensemble spread of small-scale gravity waves is considerably smaller. The small variance in time and across ensemble members suggests that the sub-seasonal prediction skill is not significantly hampered by a deterministic limit for predicting gravity wave source processes (e.g. Fritts and Alexander, 2003).
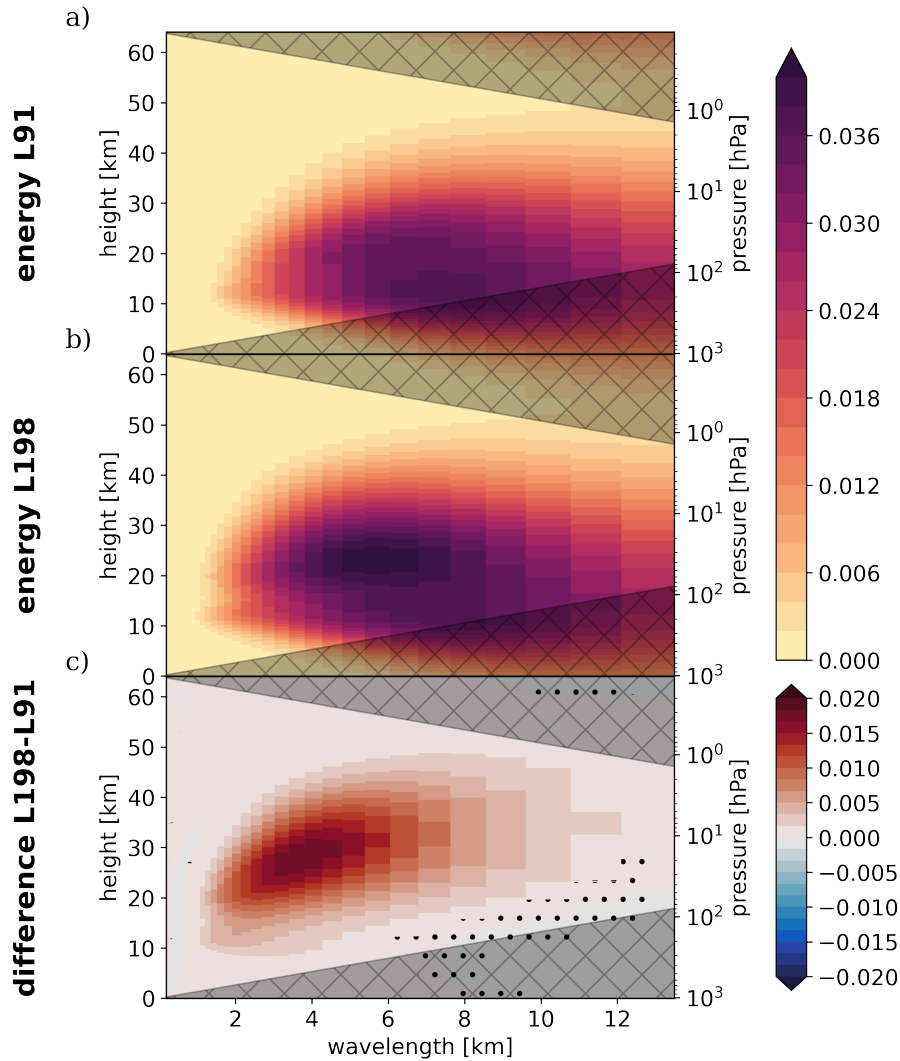
Figure 7. Ensemble-mean non-dimensional potential energy wavelet spectrum horizontally averaged between **45** °N and **70** °N for the period 2018-02-22 to 2018-03-22 in the TCo639L91 (a) and TCo639L198 (b) nudged simulations and the ensemble-mean difference TCo639L198 - TCo639L91 (c). The hatched area indicates the theoretical cone of influence and stippling in the lower panel indicates where ensemble-mean energies are not significantly different estimated by a parametric bootstrap.

The existence of a critical layer for near-stationary gravity waves is thus a good predictor for upper-stratospheric wave drag on sub-seasonal time scales. Gravity wave breaking near the critical layer constitutes a positive feedback between wave drag and the zonal-mean state of the vortex. To make full use of the critical layer as a predictor and to benefit from the aforementioned positive feedback, increased vertical resolution is necessary **to allow resolved gravity waves to effectively decelerate the zonal mean zonal wind.** Specifically, it is found that a significant part of gravity wave potential energy corresponds to waves

with as small vertical wavelength. **At low vertical resolution these waves will dissipate at lower altitudes than in a perfect model, whereas high vertical resolution enables the propagation of resolved gravity waves closer to the respective critical layer, which leads to the deposition of easterly momentum at higher altitudes with lower mean density exerting a stronger drag on the mean flow.**

270   The need for a high vertical resolution to resolve gravity wave dynamics agrees well with previous modeling studies. Waite (2016) and Skamarock et al. (2019) find model convergence for a vertical grid spacing of 200 m at a horizontal resolution of 11.8 km and 15 km, respectively, which corresponds roughly to an aspect ratio of f/N. A similar conclusion is drawn in a study about the sensitivity of the Integrated Forecasting System (IFS) to vertical resolution (Polichtchouk et al., 2019). While a vertical resolution set in accordance with the theoretical requirements to resolve stratified turbulence (see Section 1) performs

275 well in the free troposphere, a correct representation of gravity wave breaking in the stratosphere requires a higher resolution than estimated for stratified turbulence (Cullen, 2017). Hence, stratospheric temperature biases in our model with a spectral resolution of TCo639 might benefit from an even further increase in vertical resolution compared to L198 (see Appendix B).

     An open question that remains is to quantify the exact contribution of gravity wave drag to the observed reduction of the cold bias. There are two reasons that complicate such an estimate. First, the exact amount of wave drag required to maintain

280 long-lasting positive temperature anomalies in the polar stratosphere against dissipation depends on the radiative damping time scale and the 'depth' of the warming (Hitchcock and Shepherd, 2013; Hitchcock et al., 2013). Second, the stationarity assumption inherent to the downward control principle (Haynes et al., 1991), which relates wave drag to adiabatic warming, is violated by the transient nature of the experiments. Consequently, the sensitivity of polar cap temperature anomalies to vertical resolution might result from a reduced numerical error instead of adiabatic warming. While this study is concerned with the

285 polar stratosphere during weak vortex conditions, the climatological cold bias in the absence of a critical layer is shown to be sensitive to vertical resolution, especially in the tropical stratosphere (Polichtchouk et al., 2019). That sensitivity is associated with discretization errors in vertical advection and an unphysical $2\Delta z$ mode in the temperature equation. But even though some details of the mechanism remain unclear, the improvement of the stratospheric forecast with increased vertical resolution is evident.

290 **Appendix A**

This study presents results based on simulations of the SSW event in February 2018. The findings are robust and translate well to a greater number of SSW events as demonstrated by repeating the diagnostics for simulations of the events in January 2006 and February 2010. Figures A1 and A2 show the development of polar cap temperature anomaly for reanalysis and the hindcast ensemble mean as well as the ensemble-mean eddy heat flux at 100 hPa as a proxy for stratospheric planetary waves.

295 The strong resemblance amongst different events in terms of gravity waves is shown by the ensemble-mean gravity wave drag in Figures A3a and A3b. The difference in resolved and parameterised wave drag between the model configurations with low and high vertical resolution is likewise comparable (Fig. A3c-h).
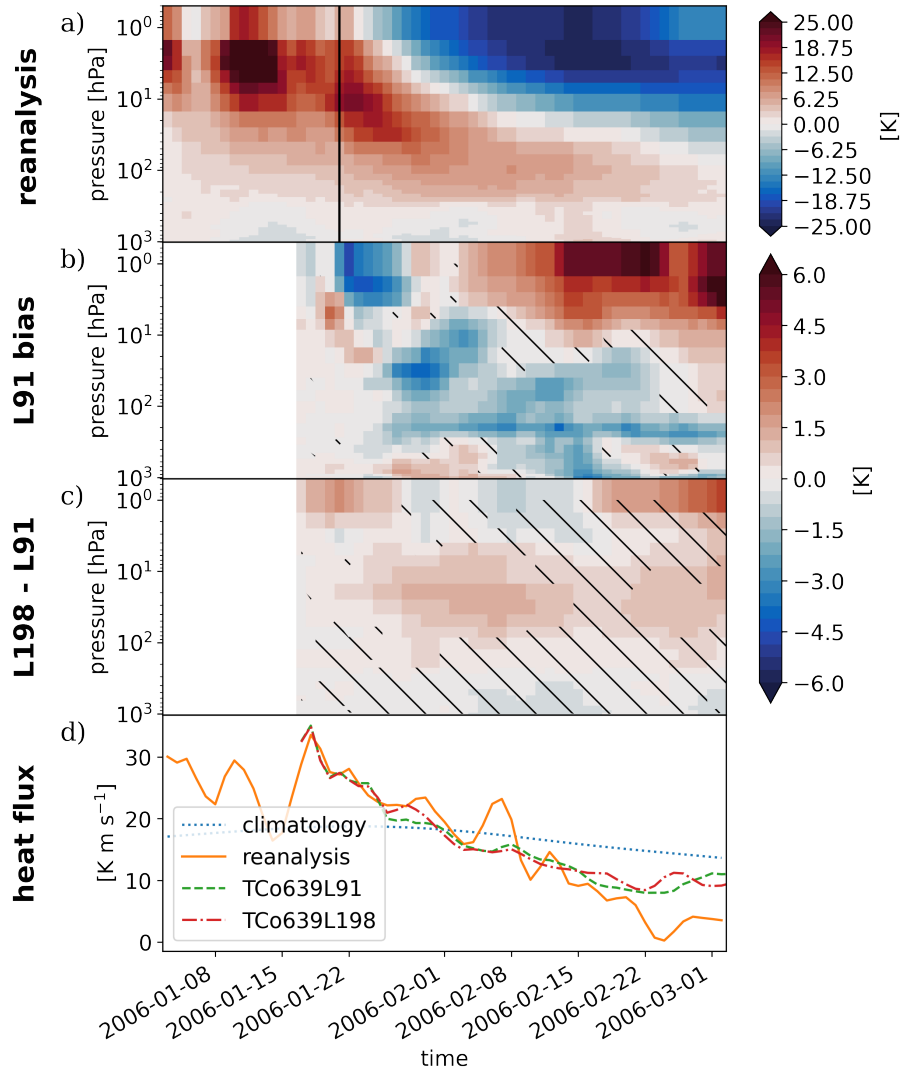
Figure A1. Same as Figure 3 but for the SSW event in 2006.

Appendix B

Operational sub-seasonal to seasonal forecasts issued by the ECMWF use 137 vertical levels and a horizontal resolution of
300   TCo639 up to day 15 and a reduced horizontal resolution of TCo319 for the remainder of the forecast (ECMWF, 2022). The
present study reveals an improved representation of stratospheric gravity waves and a reduction of the polar cap temperature
bias in the TCo639L198 hindcasts compared to a model configuration with the same horizontal resolution but only 91 vertical
levels. To compare these results with the resolution of operational forecasts, we conducted sensitivity tests with a horizontal
resolution of TCo319 and 137 vertical levels, respectively. They show that at lower horizontal resolution resolved gravity
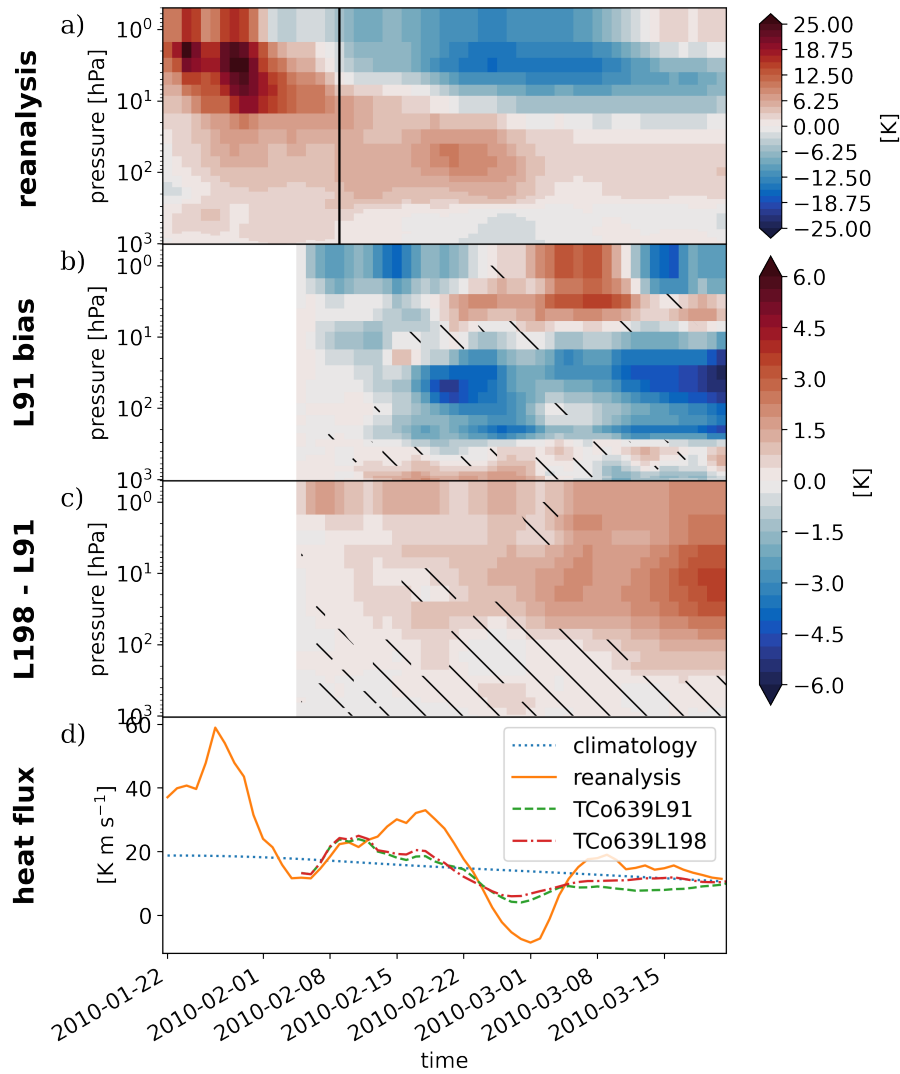
Figure A2. Same as Figure 3 but for the SSW event in 2010.

305   wave drag is reduced by 50% depending on altitude (Fig. B1). Furthermore, with TCo319 the improvement compared to 91 vertical levels is achieved to large degree already with 137 instead of 198 levels (Fig. B1). This is in contrast to the TCo639 simulations where increasing vertical resolution from L137 to L198 makes a significant difference (Fig. B2) especially in the upper stratosphere.
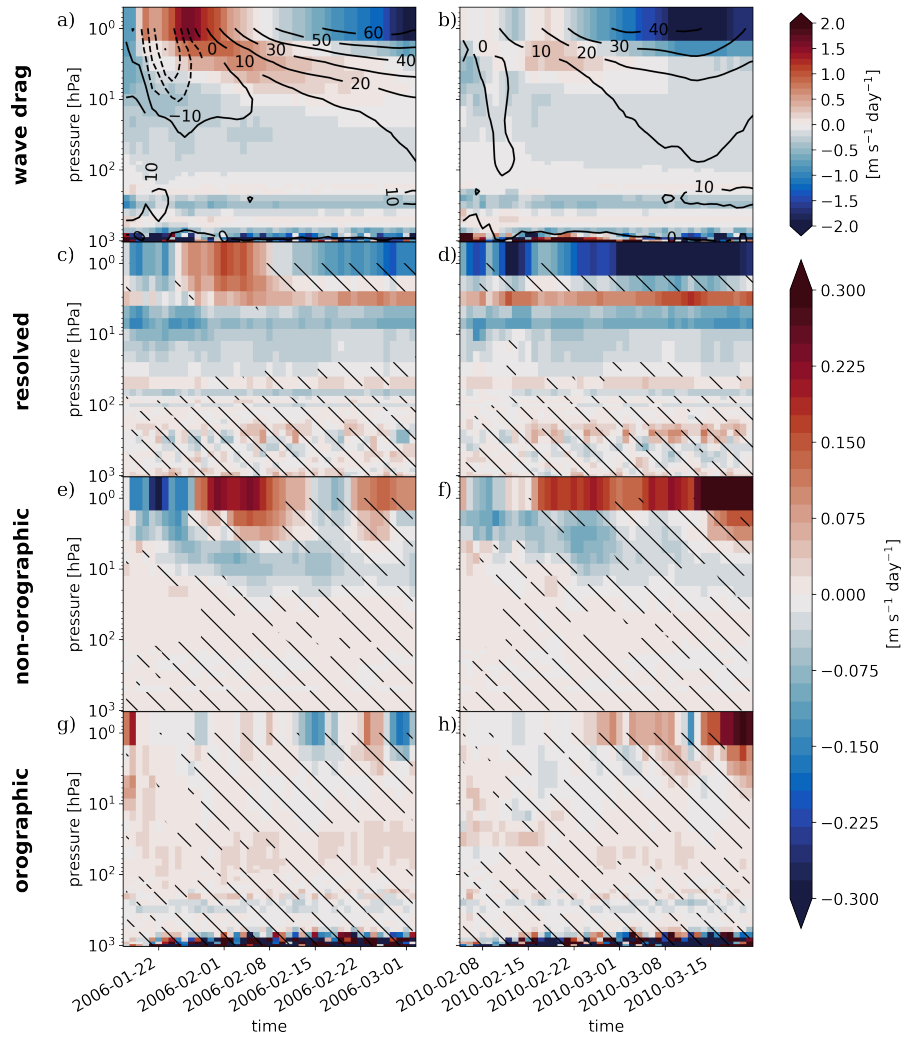
Figure A3. Ensemble-mean full gravity wave drag and zonal-mean zonal wind as in Figure B1 for the 2006 (a) and 2010 (b) SSW events. Differences in ensemble-mean gravity wave drag between the TCo639L198 and TCo639L91 free running hindcasts as in Figure 5 for 2006 (c, e, g) and 2010 (d, f, h) SSW events.

Author contributions. All authors jointly designed the study. I.P. performed the model runs and resolved gravity wave flux calculation. W.W. performed the data analysis, made the figures, and wrote the manuscript. All authors contributed to interpreting the results and editing the manuscript.

Competing interests. The authors declare no competing interests.
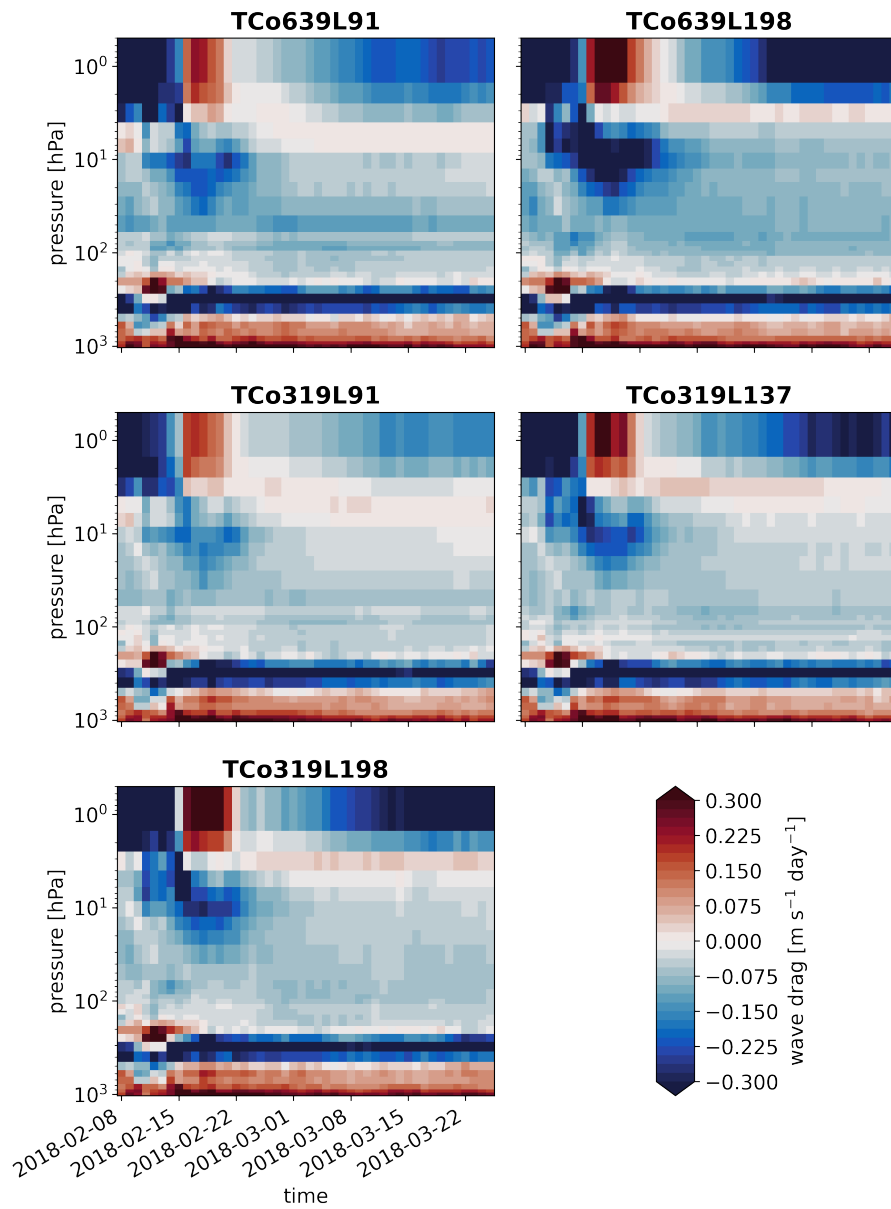
Figure B1. Ensemble-mean resolved gravity wave drag horizontally averaged between 45°N and 70°N for different model configurations initialized on February 08, 2018.
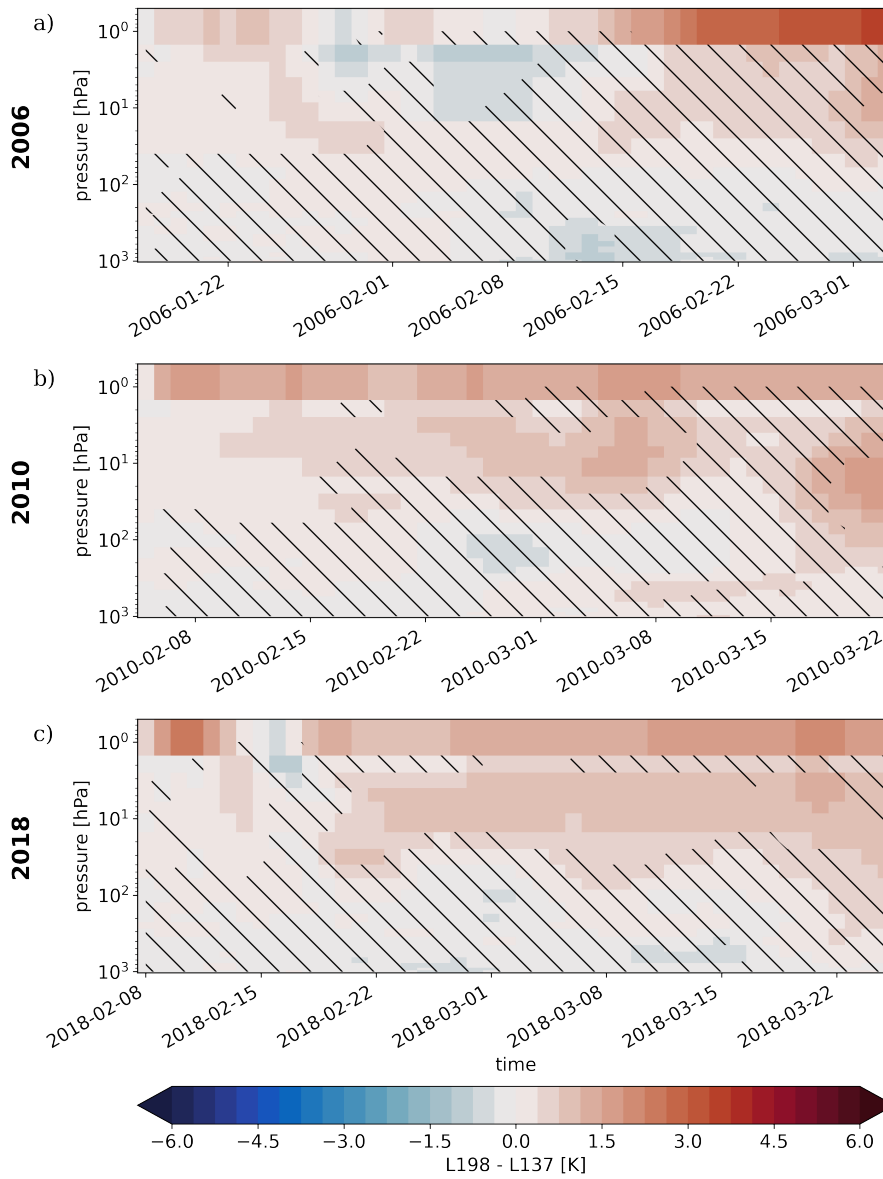
Figure B2. Differences of the TCo639L198 hindcast ensemble-mean polar cap temperature anomalies during the 2006, the 2010, and the 2018 SSW compared to the TCo639L137 hindcasts.

Acknowledgements. We thank Richard J. Greatbatch for insightful discussions and Tim Stockdale for providing the zonal-mean nudging code and for providing the 198L vertical resolution configuration. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 847456). Support from the Swiss National Science Foundation through project PP00P2_198896 to D.D. is gratefully acknowledged.

315

References

Albers, J. R. and Birner, T.: Vortex preconditioning due to planetary and gravity waves prior to sudden stratospheric warmings, Journal of the Atmospheric Sciences, 71, 4028–4054, 2014.

Baldwin, M. P. and Dunkerton, T. J.: Stratospheric harbingers of anomalous weather regimes, Science, 294, 581–584, 2001.

Baldwin, M. P., Ayarzagüena, B., Birner, T., Butchart, N., Butler, A. H., Charlton-Perez, A. J., Domeisen, D. I., Garfinkel, C. I., Garny, H., Gerber, E. P., et al.: Sudden stratospheric warmings, Reviews of Geophysics, 59, e2020RG000 708, 2021.

Charney, J. and Drazin, P.: Propagation of planetary-scale disturbances from the lower into the upper atmosphere, J. Geophys. Res, 66, 83–109, 1961.

Cullen, M.: The impact of high vertical resolution in the Met Office Unified Model, Quarterly Journal of the Royal Meteorological Society, 143, 278–287, 2017.

Domeisen, D. I., Martius, O., and Jiménez-Esteve, B.: Rossby wave propagation into the Northern Hemisphere stratosphere: The role of zonal phase speed, Geophysical Research Letters, 45, 2064–2071, 2018.

Domeisen, D. I., Butler, A. H., Charlton-Perez, A. J., Ayarzagüena, B., Baldwin, M. P., Dunn-Sigouin, E., Furtado, J. C., Garfinkel, C. I., Hitchcock, P., Karpechko, A. Y., et al.: The role of the stratosphere in subseasonal to seasonal prediction: 2. Predictability arising from stratosphere-troposphere coupling, Journal of Geophysical Research: Atmospheres, 125, e2019JD030 923, 2020a.

Domeisen, D. I., Butler, A. H., Charlton-Perez, A. J., Ayarzagüena, B., Baldwin, M. P., Dunn-Sigouin, E., Furtado, J. C., Garfinkel, C. I., Hitchcock, P., Karpechko, A. Y., Kim, H., Knight, J., Lang, A. L., Lim, E.-P., Marshall, A., Roff, G., Schwartz, C., Simpson, I. R., Son, S.-W., and Taguchi, M.: The Role of the Stratosphere in Subseasonal to Seasonal Prediction: 1. Predictability of the Stratosphere, Journal of Geophysical Research: Atmospheres, 125, e2019JD030 920, 2020b.

Dörnbrack, A., Gisinger, S., Kaifler, N., Portele, T. C., Bramberger, M., Rapp, M., Gerding, M., Söder, J., Žagar, N., and Jelić, D.: Gravity waves excited during a minor sudden stratospheric warming, Atmospheric Chemistry and Physics, 18, 12 915–12 931, 2018.

Duck, T. J., Whiteway, J. A., and Carswell, A. I.: Lidar observations of gravity wave activity and Arctic stratospheric vortex core warming, Geophysical Research Letters, 25, 2813–2816, 1998.

ECMWF: ECMWF Confluence Wiki, [online] Available from: https://confluence.ecmwf.int/display/S2S/Models (Accessed 09 May 2022), 2022.

Ern, M., Preusse, P., Alexander, M. J., and Warner, C. D.: Absolute values of gravity wave momentum flux derived from satellite data, Journal of Geophysical Research: Atmospheres, 109, 2004.

Fritts, D. C. and Alexander, M. J.: Gravity wave dynamics and effects in the middle atmosphere, Reviews of geophysics, 41, 2003.

Gupta, A., Birner, T., Dörnbrack, A., and Polichtchouk, I.: Importance of gravity wave forcing for springtime southern polar vortex breakdown as revealed by ERA5, Geophysical Research Letters, 48, e2021GL092 762, 2021.

Haynes, P., McIntyre, M., Shepherd, T., Marks, C., and Shine, K. P.: On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces, Journal of the Atmospheric Sciences, 48, 651–678, 1991.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, 2020.

Hindley, N., Wright, C., Hoffmann, L., Moffat-Griffin, T., and Mitchell, N.: An 18-year climatology of directional stratospheric gravity wave momentum flux from 3-D satellite observations, Geophysical research letters, 47, e2020GL089 557, 2020.

Hitchcock, P. and Shepherd, T. G.: Zonal-mean dynamics of extended recoveries from stratospheric sudden warmings, Journal of the atmospheric sciences, 70, 688–707, 2013.

355 Hitchcock, P., Shepherd, T. G., Taguchi, M., Yoden, S., and Noguchi, S.: Lower-stratospheric radiative damping and polar-night jet oscillation events, Journal of the Atmospheric Sciences, 70, 1391–1408, 2013.

Hitchcock, P., Butler, A., Charlton-Perez, A., Garfinkel, C., Stockdale, T., Anstey, J., Mitchell, D., Domeisen, D. I., Wu, T., Lu, Y., et al.: Stratospheric Nudging And Predictable Surface Impacts (SNAPSI): A Protocol for Investigating the Role of the Stratospheric Polar Vortex in Subseasonal to Seasonal Forecasts, Geoscientific Model Development Discussions, pp. 1–29, 2022.

360 Holton, J. R.: The influence of gravity wave breaking on the general circulation of the middle atmosphere, Journal of Atmospheric Sciences, 40, 2497–2507, 1983.

Karpechko, A. Y., Charlton-Perez, A., Balmaseda, M., Tyrrell, N., and Vitart, F.: Predicting sudden stratospheric warming 2018 and its climate impacts with a multimodel ensemble, Geophysical Research Letters, 45, 13–538, 2018.

Kautz*, L.-A., Polichtchouk*, I., Birner, T., Garny, H., and Pinto, J. G.: Enhanced extended-range predictability of the 2018 late-winter
365 Eurasian cold spell due to the stratosphere, Quarterly Journal of the Royal Meteorological Society, 146, 1040–1055, 2020.

Kodera, K., Mukougawa, H., Maury, P., Ueda, M., and Claud, C.: Absorbing and reflecting sudden stratospheric warming events and their relationship with tropospheric circulation, Journal of Geophysical Research: Atmospheres, 121, 80–94, 2016.

Kuchar, A., Sacha, P., Eichinger, R., Jacobi, C., Pisoft, P., and Rieder, H.: On the impact of Himalaya-induced gravity waves on the polar vortex, Rossby wave activity and ozone, EGUsphere, 2022, 1–22, https://doi.org/10.5194/egusphere-2022-474, 2022.

370 Lawrence, Z. D., Abalos, M., Ayarzagüena, B., Barriopedro, D., Butler, A. H., Calvo, N., de la Cámara, A., Charlton-Perez, A., Domeisen, D. I., Dunn-Sigouin, E., et al.: Quantifying stratospheric biases and identifying their potential sources in subseasonal forecast systems, Weather and Climate Dynamics Discussions, pp. 1–37, 2022.

Limpasuvan, V., Thompson, D. W., and Hartmann, D. L.: The life cycle of the Northern Hemisphere sudden stratospheric warmings, Journal of Climate, 17, 2584–2596, 2004.

375 Lindzen, R. S. and Fox-Rabinovitz, M.: Consistent vertical and horizontal resolution, Monthly Weather Review, 117, 2575–2583, 1989.

Matsuno, T.: A dynamical model of the stratospheric sudden warming, Journal of Atmospheric Sciences, 28, 1479–1494, 1971.

Polichtchouk, I. and Scott, R. K.: Spontaneous inertia-gravity wave emission from a nonlinear critical layer in the stratosphere, Quarterly Journal of the Royal Meteorological Society, 146, 1516–1528, https://doi.org/https://doi.org/10.1002/qj.3750, 2020.

Polichtchouk, I., Stockdale, T., Bechtold, P., Diamantakis, M., Malardel, S., Sandu, I., Vana, F., and Wedi, N.: Control on stratospheric
380 temperature in IFS: resolution and vertical advection, European Centre for Medium-Range Weather Forecasts, 2019.

Polichtchouk, I., van Niekerk, A., and Wedi, N.: Resolved gravity waves in the extra-tropical stratosphere: Effect of horizontal resolution increase from O (10 km) to O (1 km), Journal of the Atmospheric Sciences, 1, 2022a.

Polichtchouk, I., Wedi, N., and Kim, Y.-H.: Resolved gravity waves in the tropical stratosphere: Impact of horizontal resolution and deep convection parametrization, Quarterly Journal of the Royal Meteorological Society, 148, 233–251, 2022b.

385 Preusse, P., Dörnbrack, A., Eckermann, S. D., Riese, M., Schaeler, B., Bacmeister, J. T., Broutman, D., and Grossmann, K. U.: Space-based measurements of stratospheric mountain waves by CRISTA 1. Sensitivity, analysis method, and a case study, Journal of Geophysical Research: Atmospheres, 107, CRI–6, 2002.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, Journal of Climate, 19, 3771–3791, 2006.

390 Sato, K.: A statistical study of the structure, saturation and sources of inertio-gravity waves in the lower stratosphere observed with the MU
        radar, Journal of atmospheric and terrestrial physics, 56, 755–774, 1994.

    Skamarock, W. C., Snyder, C., Klemp, J. B., and Park, S.-H.: Vertical resolution requirements in atmospheric simulation, Monthly Weather
        Review, 147, 2641–2656, 2019.

    Song, B.-G., Chun, H.-Y., and Song, I.-S.: Role of gravity waves in a vortex-split sudden stratospheric warming in January 2009, Journal of
395   the Atmospheric Sciences, 77, 3321–3342, 2020.

    Torrence, C. and Compo, G. P.: A practical guide to wavelet analysis, Bulletin of the American Meteorological society, 79, 61–78, 1998.

    Tripathi, O. P., Baldwin, M., Charlton-Perez, A., Charron, M., Eckermann, S. D., Gerber, E., Harrison, R. G., Jackson, D. R., Kim, B.-M.,
        Kuroda, Y., et al.: The predictability of the extratropical stratosphere on monthly time-scales and its impact on the skill of tropospheric
        forecasts, Quarterly Journal of the Royal Meteorological Society, 141, 987–1003, 2015a.

400 Tripathi, O. P., Charlton-Perez, A., Sigmond, M., and Vitart, F.: Enhanced long-range forecast skill in boreal winter following stratospheric
        strong vortex conditions, Environmental Research Letters, 10, 104 007, 2015b.

    Venkat Ratnam, M., Tsuda, T., Jacobi, C., and Aoyama, Y.: Enhancement of gravity wave activity observed during a major Southern Hemi-
        sphere stratospheric warming by CHAMP/GPS measurements, Geophysical Research Letters, 31, 2004.

    Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., et al.: The
405   subseasonal to seasonal (S2S) prediction project database, Bulletin of the American Meteorological Society, 98, 163–173, 2017.

    Waite, M. L.: Dependence of model energy spectra on vertical resolution, Monthly Weather Review, 144, 1407–1421, 2016.

    Wang, L. and Alexander, M. J.: Gravity wave activity during stratospheric sudden warmings in the 2007–2008 Northern Hemisphere winter,
        Journal of Geophysical Research: Atmospheres, 114, 2009.

    Wu, R. W.-Y., Wu, Z., and Domeisen, D. I.: Differences in the sub-seasonal predictability of extreme stratospheric events, Weather and
410   Climate Dynamics, 3, 755–776, 2022.

    Yamashita, C., Liu, H.-L., and Chu, X.: Gravity wave variations during the 2009 stratospheric sudden warming as revealed by ECMWF-T799
        and observations, Geophysical Research Letters, 37, 2010.