Spontaneous inertia-gravity wave emission from a nonlinear critical layer in the stratosphere

Inna Polichtchouk[1] | **Richard K. Scott**[2]

[1]European Centre for Medium-Range Weather Forecasts, Reading, UK

[2]School of Mathematics, and Statistics, University of St Andrews, North Haugh, St Andrews, KY16 9SS, UK.

Correspondence
Inna Polichtchouk, European Centre for Medium-Range Weather Forecasts, Shinfield Rd, Reading RG2 9AX, UK.
Email: inna.polichtchouk@ecmwf.int

Funding information

Using a non-linear global primitive equation model, spontaneous inertia-gravity wave (IGW) emission is investigated in an idealized representation of the stratospheric polar night. It is shown that IGWs are spontaneously emitted in the interior of the fluid in a jet exit region that develops around a nonlinear Rossby wave critical layer. Two key ingredients for the generation are identified: the presence of a Rossby wave guide on the polar night jet; and a zero wind line on the jet flank that gives rise to nonlinear Rossby wave breaking and strong distortion of the flow. The emission of IGWs appears here as a quasi-steady process that begins at a well-defined time when the flow deformation becomes large enough. Part of the emitted IGWs undergoes wave capture by the cat's-eye flow in a Rossby wave critical layer. Another part — in a form of a well-defined IGW packet — escapes the wave capture limit, and propagates away into the far field. The propagating wave packet is numerically well-converged to increases in both vertical and horizontal resolution and thus provides an ideal test bed for understanding IGW emission and informing nonorographic gravity wave drag parametrization design.

KEYWORDS
Inertia-gravity waves, spontaneous emission, stratosphere, idealized numerical study, primitive equations

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/qj.3750

1 | INTRODUCTION

Spontaneous emission of inertia-gravity waves (IGWs) from jets and fronts provides an important source of gravity waves in the atmosphere (e.g. Fritts and Nastrom, 1992; Eckermann and Vincent, 1993; Hertzog et al., 2008). IGWs transport heat and momentum, contribute to mixing, and, have horizontal scales of $O(10-1000)$km and frequencies between the Brunt-Väisälä frequency and the Coriolis frequency. Such nonorographically generated IGWs are not fully resolved in most weather and climate models and so their effect on the larger scales is often parametrized using several simplifying assumptions (see discussion in e.g. Richter et al., 2010). Gaining a better understanding of nonorographic IGW generation can, therefore, help to inform parametrization design.

Several observational case studies have stressed spontaneous IGW emission from jet exit regions. While the emphasis has mostly been on emission from the tropospheric mid-latitude jet (e.g. Uccellini and Koch, 1987; Plougonven et al., 2003), a recent case study by Dörnbrack et al. (2018) also emphasizes emission from the jet exit region of the stratospheric polar night jet during the 2016 minor sudden stratospheric warming in the Northern Hemisphere. Radiosonde observations in Antarctica also reveal that the stratospheric polar night jet is an important source of IGWs (Yoshiki and Sato, 2000; Yoshiki et al., 2004; Sato and Yoshiki, 2008).

Motivated by the above observational studies, it is natural to consider whether spontaneous IGW emission from the stratospheric polar night jet can be represented in an idealized three-dimensional numerical model, where the emitted IGWs are well resolved and can be identified with more precision. In this regard, the present work complements previous numerical studies that have mostly focused on the IGW emission from tropospheric sources, namely surface fronts and baroclinic life cycles (e.g. Snyder et al., 1993; O'Sullivan and Dunkerton, 1995; Zhang, 2004; Plougonven and Snyder, 2007). The baroclinic life cycle studies have particularly emphasized IGW emission in the jet exit region near the tropopause, although a tropopause is not a necessary ingredient for this kind of IGW emission (Viúdez and Dritschel, 2006). At least part of the emitted waves in baroclinic life cycles have been explained as resulting from "wave capture" by the deforming background flow (Bühler and McIntyre, 2005; Plougonven and Snyder, 2005) in which IGWs are trained by the large scale horizontal deformation field and modified by the vertical shear. In the present study, the focus is again on IGW emission from a jet exit region, but in the absence of baroclinic instability and away from the boundaries such as the surface or tropopause.

A simple model of spontaneous IGW emission from jet exit regions that avoids the complicated flow structures of baroclinic instability is that of a steadily propagating cyclone-anticyclone dipole in a continuously stratified flow (McIntyre, 2009). Such a dipole represents an idealization of the flow in a jet exit region of the baroclinic flows without the complexity of time dependence and complicated spatial structure. Snyder et al. (2009); Viúdez (2007); Yasuda et al. (2015) have shown that the spontaneous IGW emission in such dipole flows can be understood as a linear response to forcing, which can be considered as arising from the small deviations from an exactly balanced flow and which increases with increasing Rossby number. The crucial ingredients for IGW emission here are strong velocities in the jet providing strong advection and significant along-jet variations producing the forcing (Plougonven and Zhang, 2014). The strong advection acts to project the otherwise slow forcing onto the fast IGW timescales, enabling spontaneous emission. Again, the structure of the emitted wave field is best understood in terms of wave capture by the large-scale dipole field.

Spontaneous IGW emission from vortex dipoles and baroclinic life cycles differs from the classic Lighthill theory on spontaneous emission in both the spatial scales and temporal nature of the emitted waves (for comprehensive discussion on this and the different forms of spontaneous emission see McIntyre (2009)). Under Lighthill type scenarios unsteady vortical motion spontaneously emits IGWs with horizontal scales larger than or comparable to the horizontal scales of the source region. In vortex dipole and baroclinic life cycle studies, on the other hand, the emission is a quasi-steady process (i.e. steady on a timescale of the IGWs), and, the emitted IGWs have small scales in comparison to their source,

due to the effects of wave capture.

The small scales of emitted IGWs in baroclinic life cycle and dipole experiments pose challenges for numerical studies, in which emitted waves are often sensitive to model configuration and where numerical convergence may be difficult or not practically possible (e.g. Zhang, 2004; O'Sullivan and Dunkerton, 1995; Plougonven and Snyder, 2007). This has hindered attempts to provide a more quantitative description of the emitted waves, for example relating frequency and wavenumbers to aspects of the balanced flow. One feature of the IGW emission in the stratospheric context presented here is that numerical convergence of most aspects of the flow is achievable at relatively modest resolutions.

A separate line of investigation of spontaneous emission has used simple analytic models to analyze the coupling between balanced and unbalanced motions asymptotically (Vanneste, 2004; Vanneste and Yavneh, 2004; Olafsdóttir et al., 2008). These range from low-order models, such as Lorenz-Krishnamurthy model discussed in Vanneste (2004) to more complex models that allow for a description of three-dimensional features in slowly evolving background flow with perturbations in a pure horizontal shear (Vanneste and Yavneh, 2004). In these studies, IGW emission is exponentially small in Rossby number and analysis typically requires severe idealizations of the background flow. The most sophisticated of these analytic models describe IGWs emitted from three-dimensional potential vorticity anomalies embedded in a background shear flow (Olafsdóttir et al., 2008; Lott et al., 2012). In this "sheared modes" description the flow is initially balanced, but evolves slowly in time as a result of the externally imposed shear. The spontaneous IGW emission turns on at a well-defined point in time when the balanced flow projects onto wave modes in the far field. This description can thus be contrasted with the dipole studies discussed above, in which the unbalance — and therefore IGW emission — is present in the beginning due to the particular flow configuration and in which the spatial structure of the dipole results in the wave capture.

In realistic atmospheric flows with more complicated time evolution and spatial structure, it is likely that elements of both dipole-like wave capture and the sheared mode-like spontaneous emission resulting from a structural change in the balanced flow would be present. So far, and to the best of the authors' knowledge, clean numerical simulations of spontaneous emission in three-dimensional, time-evolving flows, in which well-defined IGW packets are generated spontaneously at a well-defined point in time and propagate unambiguously away from an interior source region are still missing. One feature of the experiments presented below is that, while some evidence of wave capture can be seen, much of the emission is in the form of such a well-defined IGW packet that appears when the flow becomes sufficiently deformed and subsequently propagates away from the source region.

This paper is structured as follows. Section 2 describes the model, the experimental setup and the diagnostics used. Section 3 describes the results of spontaneous IGW emission from a deforming stratospheric polar night jet. This section also shows the convergence of the solutions to horizontal and vertical resolution as well as the robustness of the results to the chosen polar vortex profile. Finally, conclusions are given in Section 4.

2 | METHOD

2.1 | Numerical model

We use a global, pseudospectral, pressure coordinate model, which solves the dry hydrostatic primitive equations in a compressible (non-Boussinesq) atmosphere. The details of the model, Built on Beowolf (BOB) dynamical core, are documented in Scott et al. (2004). As the aim is to study gravity wave emission from the polar night jet, BOB is used in a stratosphere-only mode with a lower boundary placed at 200 hPa and the upper boundary at 0.02 hPa. The only external wave source is a zonal wavenumber-2 forcing at the lower boundary, representing wave generation in the

troposphere.

The initial condition is a height-independent polar night jet u, centered at 60°N:

$$u(\phi) = \begin{cases} u_0 \sin^3\phi \cos\phi, & \phi > 0 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where ϕ is latitude and $u_0 = 150$ m s^{-1}. This results in a ~50 m s^{-1} jet, which is shown in Figure 1 (black solid line). The initial $T_0 = 240$ K isothermal temperature distribution is specified, which results in a realistic stratospheric stratification with Brunt-Väisälä frequency $N = 2 \times 10^{-2}\ s^{-1}$ (see e.g. Figure 4 in Fueglistaler et al., 2011).

To assess the robustness of our results to the choice of the initial jet profile, we also consider a second initial condition, in which the polar vortex is represented by a simple patch of uniform Ertel potential vorticity, q, with

$$q(\phi) = \begin{cases} 2\Omega \sin\phi, & \phi < 0 \\ \frac{2\Omega q_0 \sin\phi}{\sin\frac{\pi}{3}}, & 0 \le \phi < \frac{\pi}{3} \\ 2\Omega q_1, & \phi \ge \frac{\pi}{3}, \end{cases} \quad (2)$$

where $q_1 = 1.15$ and $q_0 = \frac{1-2q_1(1-\sin\frac{\pi}{3})}{\sin\frac{\pi}{3}}$. The zonal wind profile corresponding to $q(\phi)$ is also shown in Fig. 1 (red dashed line). Spontaneous IGW emission for this initial condition is discussed in section 3.2.

Stationary planetary wave-2 is forced at the lower boundary by prescribing a geopotential height perturbation Φ' (similar to Scott and Polvani, 2006):

$$\Phi'(\phi, \lambda, t) = \begin{cases} g h_0(t) \sin^2(\pi \frac{\phi - \pi/6}{\pi/3}) \cos 2\lambda, & \phi > \frac{\pi}{6} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where λ is longitude and where a perturbation amplitude $h_0(t)$ grows linearly at a rate of 10 m day^{-1} to minimize spurious generation of gravity waves during initialization. While it is impossible to completely eliminate a small amount of gravity wave generation associated with the growth of the forcing, this was observed to be tiny relative to subsequent emission. In particular, slower amplitude growth rates were also investigated and were found to result in a similar IGW emission as for the 10 m day^{-1} rate. The model is integrated for 40 days. To prevent wave reflection from the model top, a sponge layer above 0.5 hPa is applied in a form of a linear damping at a rate $\tau^{-1} = \alpha \max\{0, [(0.5 - p)/0.5]^2\}$, where p is pressure and α =4(day)$^{-1}$. The damping is not applied on the mean flow.

Results at T170L80 resolution are mostly discussed. This resolution is equivalent to horizontal grid spacing of 0.7° and the vertical grid spacing of approximately 800 m. The convergence of results to both the horizontal and vertical resolutions will be investigated in section 3.1, by performing simulations also at T85 and T341 horizontal resolution (or 1.4° and 0.35°, respectively) and L40 and L160 vertical resolution (or 1600 m and 400 m, respectively). Enstrophy at small scales is removed by applying a scale-selective hyperdiffusion ∇^8, with a hyperdiffusion coefficient chosen such that the highest resolved wavenumber is dissipated at a timescale of 1.5 h at all horizontal resolutions.

2.2 | Gravity wave diagnostics

The gravity waves found in our simulations are low frequency IGWs. The dispersion relation for hydrostatic IGWs under the approximation $m >> k_h$ is (e.g. Fritts and Alexander, 2003; Holton, 1992):

$$\hat{\omega}^2 = N^2 \frac{k_h^2}{m^2} + f^2 \tag{4}$$

where $\hat{\omega} = \omega - \mathbf{k} \cdot \mathbf{u}$ is the intrinsic frequency, $\mathbf{u}$ is the horizontal background wind, $\mathbf{k}$ is the horizontal wavenumber vector (k, l), and m is the vertical wavenumber; $k_h = \sqrt{k^2 + l^2}$; $f = 2\Omega \sin\phi$ is the Coriolis parameter; and, N is the Brunt-Väisälä frequency. The ground-based group velocity $\mathbf{c_g}$ of an IGW packet is:

$$\mathbf{c_g} = \mathbf{u} + \frac{N^2}{\hat{\omega} m^2} \left(k, l, -\frac{k_h^2}{m} \right). \tag{5}$$

The convention is that for westward ($k < 0$) and upward ($m < 0$) group velocity, the intrinsic phase propagation is westward and downward.

In what follows, gravity wave packets are diagnosed from the divergence field, filtered by removing planetary and synoptic scales, necessary since the balanced flow itself contains a substantial divergent component. To achieve this it was found to be sufficient to simply remove all zonal wavenumbers 1–8 (i.e. horizontal scales larger than approximately 3500 km at mid-latitude). It should be noted that more reliable balance-imbalance decomposition methods exist (e.g. Viúdez and Dritschel, 2006; Kafiabad and Bartello, 2016; Eden et al., 2019), but are unnecessary for the relatively simple flows considered in this study.

The vertical flux of horizontal momentum of a gravity wave is calculated as follows:

$$(F_{px}, F_{py}) = \overline{\rho}(\overline{u'w'}, \overline{v'w'}), \tag{6}$$

where $\overline{\rho}$ is the background density, and u, v, w is the zonal, meridional and vertical wind, respectively. The overbar denotes a time and zonal mean and primes are obtained by removing all zonal wavenumbers 1-8. The total vertical flux of horizontal momentum due to gravity waves is then:

$$F_{ph} = \sqrt{F_{px}^2 + F_{py}^2}. \tag{7}$$

3 | RESULTS

Spontaneous IGW emission from the deforming idealized polar vortex is now discussed beginning with the initial condition (1). At $t = 0$, the topographic wave forcing begins to increase, generating zonal wavenumber-2 Rossby waves with broad meridional scale. These propagate upward and equatorward on the curvature of the background flow [1], and encounter an absorbing critical layer region near 30°N (i.e. in the region of diminishing easterlies), deforming the flow streamlines into the typical "cat's-eye" pattern. The location of the critical layer region here is consistent with that in the real stratosphere (e.g. Hitchman and Huesmann, 2009).

Figure 2 shows latitude-longitude cross sections of streamlines with filtered divergence field (shading) at different

[1] Refractive index (Matsuno, 1970) can be used to assess the meridional and vertical location of Rossby wave propagation, although the assumptions of scale separation and small amplitude disturbances are often not appropriate in the winter stratosphere (e.g. Scott, 2019).

altitudes and times in the evolution. As the flow deformations increase, two monochromatic gravity wave packets are emitted at 30°N, at $t = 26$ days (at which time the amplitude of the topographic forcing at 60°N has reached 260 m). The emission begins most strongly at $z = 20$ km, approximately one scale height up from the bottom boundary. Since the only explicit wave source in the system is the zonal wavenumber-2 topography, and since there is no obvious shear instability of the balanced flow, the gravity waves appear to be spontaneously emitted from the balanced flow in the anticyclone, in the jet exit region. The stronger emission in the anticyclones compared to the cyclones is consistent with previous analytical studies (Vanneste and Yavneh, 2004; Olafsdóttir et al., 2008).

The IGW packets in Fig. 2 in the jet exit region resemble those present in the jet exit region in tropospheric baroclinic life cycles (O'Sullivan and Dunkerton, 1995; Plougonven and Snyder, 2007), with the important distinction that here the gravity waves are generated in the interior of the stratosphere, away from boundaries or any strong variations in static stability. From Fig. 2, the estimate for the horizontal wavelength in the source region at 30°N and 120°E is $\lambda_x \approx 2500$ km (with $k >> l$). An estimate for the vertical wavelength in the source region is $\lambda_z \approx 8$ km, which can be obtained by examining vertical cross sections of filtered divergence at 30°N (not shown, but see ahead to Fig. 4). This results in a horizontal to vertical aspect ratio of 300, consistent with N/f (recall that $N = 2 \times 10^{-2}$ s^{-1}). From equation (4) this gives an intrinsic frequency of $\hat{\omega} \approx 1.4f$. Inspection of a vertical cross section at 30°N (not shown) reveals that the phase lines are tilted upward and westward ($k < 0, m < 0$). Thus, from equation (5) $\mathbf{c_g} \approx (-1.4 \times 10^3, 0, 4.5)$ km day^{-1} + $\mathbf{u}$. In the regions where the zonal wind speed (i.e. the component of velocity parallel to the local wave vector) exceeds intrinsic zonal group velocity, part of the emitted waves appear to the east (i.e. downstream) of the source region in the first column in Fig. 2. Another part is strained by the large scale flow and appears in the anticyclone south of the source region at 30°N, and, at later times undergoes wave capture by the anticyclone. This can be seen in Figure 3, which shows a close-up view of filtered divergence in the cat's eye at $z = 20$ km and at $t = 34$ days.

To the north of the source region, a well-defined gravity wave packet escapes the critical layer region and "leaks" into the upper stratosphere into the far field (see middle panel in Fig. 2). To illustrate more clearly the extent of the vertical and poleward propagation away from the source region, Figure 4 shows vertical cross sections through the line segment shown in Fig. 2h. The horizontal wavelength of this IGW packet is $\lambda_x \approx 1200$ km, and the vertical wavelength varies from $\lambda_z \approx 3$ km at low latitudes to $\lambda_z \approx 8$ km at 60°N, at the latitude of the jet core. The increase in the vertical wavelength with increasing latitude appears consistent with the increase in the vertical wavelength expected from increase in the horizontal wind speed for waves obeying mid-frequency dispersion relation — i.e. on neglect of f in equation (4). Note that in Fig. 4 another IGW is present just above the jet core at $t = 29$ days (i.e. at 25°N and 110°E) that is part of the waves seen in the first column of Fig. 2. This wave does not propagate further upward and appears to be captured by the local jet structure.

The background flow strongly affects where the "leaked" IGW packet propagates: the propagation is not directly upward. Instead the packet propagates upward and north-eastward, focusing into the core of the polar night jet at 60°N as it propagates up. Such focusing has been previously discussed in Dunkerton (1984); Sato et al. (2009, 2012) and can be explained by the alignment of the wave vector and mean flow as follows (see also an enlightening schematic in Fig. 6 of Sato et al. (2012)). From the middle panels of Fig. 2 it can be seen that the wavenumber vector $\mathbf{k}$ points from southeast to northwest. The intrinsic horizontal group velocity $\hat{\mathbf{c_g}}$ is parallel to $\mathbf{k}$, whereas the ground-based horizontal group velocity $\mathbf{c_g} = \hat{\mathbf{c_g}} + \mathbf{u}$ (recall that $\mathbf{u}$ is the horizontal background wind vector). Since the background wind is predominantly westerly, the ground-based group velocity $\mathbf{c_g}$ is rotated clockwise from the direction of the wavenumber vector $\mathbf{k}$. The result is that the wave packet moves with this $\mathbf{c_g}$ in a more northeast direction and into the jet core.

The gravity wave emission from the jet exit region appears to be associated with a strong along-jet flow variation combined with moderate 15-20m s^{-1} jet velocities. As discussed in (McIntyre, 2009), the spontaneous emission relies on spatial and/or temporal scale matching in the source region, which occurs most efficiently for $O(1)$ Rossby number.

To measure unbalance we calculate the advective Rossby number (e.g. Uccellini et al., 1984; O'Sullivan and Dunkerton, 1995)

$$Ro \equiv \frac{|\mathbf{u} \cdot \nabla \mathbf{u}|}{|\mathbf{u}| f}. \tag{8}$$

Initially the flow is balanced (i.e. $Ro = 0$ at $t = 0$) but subsequently slowly time evolving as a result of the growing planetary wave. At $t \approx 26$ days — when the flow deformation becomes large enough — $Ro \gtrsim 0.4$ in the source region at 30°N and $z = 20$ km, indicating a less balanced flow and the possibility for significant IGW emission. From this time on, the IGW emission from the jet exit region is a quasi-steady process. It should be noted that IGW emission can occur for smaller Ro, but the amplitude of the emitted waves decreases rapidly with Ro.

While the IGW emission is a quasi-steady process, the amplitude of the "leaked" IGW packet increases with time (see Figures 2 and 4). To elucidate the time evolution of the "leaked" IGW packet and to relate the emission to the slowly evolving background flow, Figure 5a shows the time series of the IGW amplitude and the planetary wave amplitude together with horizontal strain in the source region (i.e. the cat's eye). The IGW amplitude is defined as the maximum absolute value of filtered divergence at $z = 35$ km between 35°N and 60°N and the planetary wave amplitude is defined as the maximum absolute value of zonal wavenumber-2 relative vorticity. A measure of mean horizontal strain is given by (e.g. Plougonven and Zhang, 2014):

$$\mathcal{S} = \frac{1}{2}\sqrt{\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)^2}, \tag{9}$$

where u and v are zonal and meridional background wind respectively, and the derivatives are taken in spherical coordinates. In the figure, the IGW amplitude begins to grow exponentially at $t \approx 26$ days. After $t \approx 28$ days a subsequent increase in the IGW amplitude appears to be independent from the planetary wave amplitude, which saturates at $t \approx 28$ days. However, the horizontal strain in the source region continues to increase and inspection of IGW amplitude versus horizontal strain in Fig. 5b reveals that the IGW amplitude grows exponentially with increase in horizontal strain.

As the amplitude of planetary wave-2 forcing continues to increase linearly with time, it provides a continuous wave forcing in the source region and hence a continuous emission of IGWs. To investigate whether the amplitude of the IGW decreases once the planetary wave-2 forcing is switched off at an earlier time, we perform a simulation in which the wavenumber-2 forcing is first grown and then, from $t = 25$ days, ramped back down to zero at the same rate of 10 m/day. The relationship between IGW amplitude, planetary wave-2 amplitude and horizontal strain for this simulation is also shown in Fig. 5 (dotted lines in a) and circles in b)). When the forcing is ramped down, the planetary wave amplitude in the source region decreases concomitantly, but the IGWs continue to be emitted due to an increase in horizontal strain. It is only when the strain begins to decrease at $t = 32$ days that the IGW amplitude also begins to decrease.

To establish the ability of the spontaneously emitted IGWs to affect the background flow, it is of interest to estimate how much momentum they carry. The total as well as meridional and zonal vertical flux of horizontal momentum due to the waves shown in Figs. 2-4 is shown in Figure 6 (note the nonlinear contour interval). The figure shows that most of the momentum is concentrated in the 15 km $< z <$ 30 km region at 30°N, with relatively little momentum flux being carried by the "leaked" waves to higher altitudes and latitudes. This means that most of momentum is deposited near the source region in the lower stratosphere, whereas the saturation and/or breaking of the upward and north-eastward propagating "leaked" IGW will deposit little momentum in the upper stratosphere. On the other hand, mean flow changes at these altitudes, where the density is very low, may yet be non-negligible. Note that the signs $\rho\overline{u'w'} < 0$ and

$\rho\overline{v'w'} > 0$ at latitudes poleward of 35°N and altitudes $z > 25$ km are consistent with the direction of the horizontal wave-vector **k** in the region of propagation.

Compared to the typical observed distribution of the absolute gravity wave momentum flux in the winter lower stratosphere (see e.g. Figure 13 in Ern et al. (2018)), the momentum fluxes from these waves are three orders of magnitude smaller. Ern et al. (2004) showed that for mid-frequency approximation, equation (7) can be written as:

$$F_{ph} = \frac{1}{2}\overline{\rho}\frac{k_h}{m}\left(\frac{g}{N}\right)^2\left(\frac{T'}{\overline{T}}\right)^2, \tag{10}$$

where g is gravitational acceleration, T' is a temperature perturbation due to the gravity wave and $\overline{T}$ is a background temperature. While the mid-frequency approximation is not strictly valid for low frequency IGWs discussed here, equation (10) is a reasonably good approximation to F_{ph} in Fig. 6a (not shown). Equation (10) implies that for a fixed vertical wavelength, the IGWs with their large horizontal wavelengths are not expected to carry a large amount of momentum and therefore are not likely to have a large effect on the time-mean zonal-mean circulation in comparison to small horizontal wavelength waves. Having said this, a combined momentum flux from a large number of such IGW packets might be significant. Because of the sensitivity of the evolution of the polar vortex to small changes in the background flow (e.g. Birner and Williams, 2008; Matthewman and Esler, 2011; Albers and Birner, 2014), it is not inconceivable that the cumulative effect of such waves may play a role in modulating larger changes in state of the winter stratosphere.

3.1 | Convergence to horizontal and vertical resolutions

The IGWs in this study are of a large enough vertical and horizontal scale to be well resolved by the T170L80 resolution. However, as previous studies have found the properties of the spontaneously emitted gravity waves to be sensitive to model resolution (e.g. Zhang, 2004; Plougonven and Snyder, 2007), it is pertinent to assess the robustness of our results to resolution.

We have performed a number of simulations in which the horizontal and vertical resolutions are changed both separately and together. Here we show results of the latter, in particular, experiments at resolutions of T85L40, T170L80 and T341L160. This is equivalent to horizontal grid spacing of 1.4°, 0.7° and 0.35°, and, vertical grid spacing of 1000 m, 800 m, and 400 m respectively. Recall that the hyperdiffusion time scale is fixed to 1.5h for the highest resolved wavenumber in each experiment. The time step size is 180 s, 90 s and 45 s for the T85L40, T170L80 and T341L160 resolutions respectively.

Figures 7 and 8 compare the emitted gravity wave packet at these resolutions at t=30 days. Fig. 7 shows latitude-longitude cross sections of filtered divergence and PV fields at two different altitudes and Fig. 8 shows vertical cross sections of filtered divergence and wind magnitude. The figures show a clear convergence at T170L80 resolution, as panels c) and d) are essentially indistinguishable from panels e) and f) in Fig. 7, and panel b) is essentially indistinguishable from panel c) in Fig. 8. The convergence of the horizontal wavelengths appears to be achieved already at T85L40 resolution (cf. panels a) and b) with panels c) and d) in Fig. 7). However, the IGW packet amplitude is too weak. This is because the hyperdiffusion is having a larger impact on the large scale flow at T85L40 resolution than at T170L80 or T341L160 resolutions. The convergence of vertical wavelengths is achieved at 800 m (or L80) vertical resolution (cf. panel b) to c) in Fig. 8). Thus to resolve vertical wavelengths of $\lambda_z \sim 3$ km (as observed at 30°N in the 25 km<z<40 km region in Fig. 4), at least four vertical grid points are needed across a single wave.

At later times, the part of the IGW that is captured by the anticyclone (see Fig. 3) is more sensitive to resolution. As

shown in Figure 9, the resolved wavelength at these times becomes smaller as resolution becomes finer, consistent with the wave capture theory of Bühler and McIntyre (2005) in which wavelengths are strained by the background balanced flow to the smallest available scale.

3.2 | Sensitivity to the polar night jet profile

We next consider briefly the sensitivity of the spontaneous IGW emission to the details of the polar night jet. To address this, we perform a simulation similar to the above but with the initial jet profile specified by a PV patch (see equation (2) and red line in Fig. 1). Figure 10 shows the horizontal cross sections of filtered divergence and PV and Figure 11 shows the vertical cross section of filtered divergence and wind magnitude at different times in the evolution for this PV patch profile. The cross sections again demonstrate clear IGW emission in the jet exit region, similar to that seen in Figs. 2-4 for the smoother initial condition, again with distinct regions of wave capture in the critical layer and wave packet propagation. In addition to the IGWs emitted in the jet exit region, there is also a region of IGW emission centered on the PV jump of the vortex edge (i.e. at the core of the jet). These waves originate at the lower boundary and quickly fill out the jet core region at all longitudes. They thus appear to be related to the jump in the PV and its intersection with the lower boundary, being absent in the smooth vortex case examined above.

By comparing the basic characteristics of the smooth and PV patch profiles, we may obtain an idea of the key flow features needed for spontaneous emission in the current stratospheric context. Both profiles have (i) a well-defined Rossby wave guide in the form of the polar night jet, enabling vertical propagation of Rossby waves from the surface, as well as (ii) a broad surf zone region of weak winds decreasing to zero in the tropics. Less obvious from the zonal wind profiles, but easily verifiable, is that both profiles also have weak but non zero potential vorticity gradients throughout this weak wind region. The Rossby waves generated by the wave-2 forcing at 60°N are therefore able to propagate equatorward from their source region — on the background of this non-vanishing PV gradient — and encounter their absorbing critical layer in low latitudes. We hypothesize that this is an important ingredient for IGW emission in the jet exit region in this study.

The above hypothesis is supported by two further experiments that were performed using variations of the PV patch profile (not shown). Both represented the polar vortex by a patch of uniform PV, but differed in the PV profile equatorward of the vortex edge. In one case , PV gradients equatorward of the jet were exactly zero. In the following time evolution, the vortex edge acted as a near perfect wave guide, with Rossby waves propagating straight up vertically from the surface to the upper atmosphere. The flow at low latitudes remained relatively undisturbed and no IGW emission was observed. In the second case, PV gradients equatorward of the vortex edge were set to the planetary value. In this case, Rossby waves were able to propagate freely away from the vortex edge. At low latitudes, however, the flow being everywhere westerly resulted in Rossby waves simply propagating through the tropics without dissipating and into the opposite hemisphere. Again, there was no clearly defined low-latitude critical layer region, no jet exit region and no IGW emission. We thus suggest that non-zero PV gradients in the surf zone and a low-latitude region of weak winds are both necessary flow features for IGW emission of the form presented here.

4 | CONCLUSIONS

Spontaneous inertia-gravity wave (IGW) emission from the stratospheric polar night jet is relatively unexplored. The examples presented above provide a first clean illustration that such emission can occur in relatively simple, yet physically realistic, flows, and provide a link with the recent observational studies of IGW in the winter polar stratosphere

(Dörnbrack et al., 2018). In particular, they illustrate two important points. First, that stratospheric sources of IGWs, though often neglected in nonorographic gravity wave drag schemes used in current general circulation models, may actually be relatively common. Second, they illustrate that IGW wave packets leaving their source region may travel substantial horizontal distances, in addition to their vertical propagation. The assumption of purely vertical propagation made in most current gravity wave drag schemes therefore may not be strictly appropriate.

It should be borne in mind that, while the momentum fluxes found here are many times smaller than those found in observational studies, much larger fluxes might be expected from more complicated flow fields where emission might be either stronger or more widespread. In other words, the care taken to construct as clean as possible an example of spontaneous emission here necessarily limits the magnitude of the effect. We anticipate that future work with more complex flow fields will provide a clearer idea (i) of how important spontaneous emission is likely to be in typical stratospheric conditions and (ii) how momentum fluxes might eventually be linked to details of the balanced flow such as the EP flux divergence associated with planetary scale waves.

To put our study in the context of previously studied idealized paths to spontaneous emission from balanced flows, we note that the emission here has elements of both the dipole and the "sheared modes" paradigms discussed in the introduction. As in the sheared modes case the flow is initially balanced but slowly time evolving due to the growing planetary wave forcing; the IGW emission then appears to turn on at a well-defined point in time when the flow deformation becomes large enough. Once it begins however, the IGW emission from the jet exit region is quasi-steady and closer in character to that of the dipole cases. An element of wave capture is present in our study, but in contrast to the dipole studies much of the wave propagates away from the source region into the far field. The IGW emission from baroclinic life cycles also has this relation to the dipoles and sheared modes: the flow is time evolving from a state of balance to one of unbalance, as in the sheared modes, but once it begins the emission appears quasi-steady and closer to the dipole case. What distinguishes our simulations from those of baroclinic life cycles is then (i) the relatively simple, stable vertical structure of the background flow, (ii) that the spatial distortion of the balanced flow develops simply through the linear growth of planetary wave forcing, and (iii) that the IGW emission occurs in the interior of the domain, away from the surface or tropopause. Moreover, we note that all aspects of IGW packet emission here appear to be well-resolved numerically, with convergence at relatively modest horizontal and vertical resolutions, so that the example may be a useful test case for further study. In general terms, the key feature of the balanced flow is the nonlinear development of the cat's-eye flow in a Rossby wave critical layer, with its associated jet exit region. As far as the authors are aware, the present work provides the first examples of IGW emission from a such nonlinear critical layer development. The basic balanced flow configuration, however, is relatively straightforward and may be constructed in even simpler flows and geometries; we anticipate therefore that it may be used advantageously to examine fundamental aspects of IGW emission.

ACKNOWLEDGEMENTS

The authors thank Riwal Plougonven for extremely helpful discussions and comments that shaped this manuscript. Sonja Gisinger is thanked for useful insights on observations. Chris Snyder and two anonymous reviewers are thanked for their comments that have greatly improved this manuscript. This work originated at the Kavli Institute for Theoretical Physics Program on "Wave-Flow Interaction in Geophysics, Climate, Astrophysics, and Plasmas".

REFERENCES

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

Birner, T. and Williams, P. D. (2008) Sudden stratospheric warmings as noise-induced transitions. *Journal of the Atmospheric Sciences*, **65**, 3337–3343.

Bühler, O. and McIntyre, M. E. (2005) Wave capture and wave–vortex duality. *Journal of Fluid Mechanics*, **534**, 67–95.

Dörnbrack, A., Gisinger, S., Kaifler, N., Portele, T. C., Bramberger, M., Rapp, M., Gerding, M., Söder, J., Žagar, N. and Jelić, D. (2018) Gravity waves excited during a minor sudden stratospheric warming. *Atmospheric Chemistry and Physics*, **18**, 12915–12931.

Dunkerton, T. J. (1984) Inertia–gravity waves in the stratosphere. *Journal of the Atmospheric Sciences*, **41**, 3396–3404.

Eckermann, S. D. and Vincent, R. A. (1993) VHF radar observations of gravity-wave production by cold fronts over southern Australia. *Journal of the Atmospheric Sciences*, **50**, 785–806.

, C., Chouksey, M. and Olbers, D. (2019) Gravity wave emission by shear instability. *Journal of Physical Oceanography*, **49**, 2393–2406.

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

Ern, M., Trinh, Q. T., Preusse, P., Gille, J. C., Mlynczak, M. G., Russell III, J. M. and Riese, M. (2018) GRACILE: a comprehensive climatology of atmospheric gravity wave parameters based on satellite limb soundings. *Earth System Science Data*, **10**, 857–892.

Fritts, D. C. and Alexander, M. J. (2003) Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics*, **41**. 1003.

Fritts, D. C. and Nastrom, G. D. (1992) Sources of mesoscale variability of gravity waves. Part II: Frontal, convective, and jet stream excitation. *Journal of the Atmospheric Sciences*, **49**, 111–127.

Fueglistaler, S., Haynes, P. H. and Forster, P. M. (2011) The annual cycle in lower stratospheric temperatures revisited. *Atmospheric Chemistry and Physics*, **11**, 3701–3711.

Hertzog, A., Boccara, G., Vincent, R., Vial, F. and Coquerez, P. (2008) Estimation of gravity-wave momentum fluxes and phase speeds from long-duration stratospheric balloon flights. 2. Results from the Vorcore campaign in Antarctica. *Journal of the Atmospheric Sciences*, **65**, 3056–3070.

Hitchman, M. H. and Huesmann, A. S. (2009) Seasonal influence of the quasi-biennial oscillation on stratospheric jets and Rossby wave breaking. *Journal of the Atmospheric Sciences*, **66**, 935–946.

Holton, J. (1992) *An introduction to dynamic meteorology Third Edition.* Academic Press, London, 507pp.

Kafiabad, H. A. and Bartello, P. (2016) Balance dynamics in rotating stratified turbulence. *Journal of Fluid Mechanics*, **795**, 914–949.

Lott, F., Plougonven, R. and Vanneste, J. (2012) Gravity waves generated by sheared three-dimensional potential vorticity anomalies. *Journal of the Atmospheric Sciences*, **69**, 2134–2151.

Matsuno, T. (1970) Vertical Propagation of Stationary Planetary Waves in the Winter Northern Hemisphere. *Journal of the Atmospheric Sciences*, **27**, 871–883.

Matthewman, N. J. and Esler, J. (2011) Stratospheric sudden warmings as self-tuning resonances. Part I: Vortex splitting events. *Journal of the Atmospheric Sciences*, **68**, 2481–2504.

McIntyre, M. E. (2009) Spontaneous imbalance and hybrid vortex–gravity structures. *Journal of the Atmospheric Sciences*, **66**, 1315–1326.

Olafsdóttir, E., Daalhuis, A. O. and Vanneste, J. (2008) Inertia–gravity-wave radiation by a sheared vortex. *Journal of Fluid Mechanics*, **596**, 169–189.

O'Sullivan, D. and Dunkerton, T. J. (1995) Generation of inertia–gravity waves in a simulated life cycle of baroclinic instability. *Journal of the Atmospheric Sciences*, **52**, 3695–3716.

Plougonven, R. and Snyder, C. (2005) Gravity waves excited by jets: Propagation versus generation. *Geophysical Research Letters*, **32**.

(2007) Inertia–gravity waves spontaneously generated by jets and fronts. Part I: Different baroclinic life cycles. *Journal of the Atmospheric Sciences*, **64**, 2502–2520.

Plougonven, R., Teitelbaum, H. and Zeitlin, V. (2003) Inertia gravity wave generation by the tropospheric midlatitude jet as given by the Fronts and Atlantic Storm-Track Experiment radio soundings. *Journal of Geophysical Research: Atmospheres*, **108**.

Plougonven, R. and Zhang, F. (2014) Internal gravity waves from atmospheric jets and fronts. *Reviews of Geophysics*, **52**, 33–76.

Richter, J. H., Sassi, F. and Garcia, R. R. (2010) Toward a physically based gravity wave source parameterization in a general circulation model. *Journal of the Atmospheric Sciences*, **67**, 136–156.

Sato, K., Tateno, S., Watanabe, S. and Kawatani, Y. (2012) Gravity Wave Characteristics in the Southern Hemisphere Revealed by a High-Resolution Middle-Atmosphere General Circulation Model. *Journal of the Atmospheric Sciences*, **69**, 1378–1396.

Sato, K., Watanabe, S., Kawatani, Y., Tomikawa, Y., Miyazaki, K. and Takahashi, M. (2009) On the origins of mesospheric gravity waves. *Geophysical Research Letters*, **36**.

Sato, K. and Yoshiki, M. (2008) Gravity wave generation around the polar vortex in the stratosphere revealed by 3-hourly radiosonde observations at Syowa Station. *Journal of the Atmospheric Sciences*, **65**, 3719–3735.

Scott, R. and Polvani, L. M. (2006) Internal variability of the winter stratosphere. Part I: Time-independent forcing. *Journal of the Atmospheric Sciences*, **63**, 2758–2776.

Scott, R., Rivier, L., Loft, R. and Polvani, L. (2004) BOB, model description and user's guide. *NCAR Technical Note NCAR/TN-404+IA*.

Scott, R. K. (2019) Nonlinear latitudinal transfer of wave activity in the winter stratosphere. *Quarterly Journal of the Royal Meteorological Society*, **145**, 1933–1946.

Snyder, C., Plougonven, R. and Muraki, D. (2009) Forced linear inertia-gravity waves on a basic-state dipole vortex. *Journal of the Atmospheric Sciences*, **66**, 3464–3478.

Snyder, C., Skamarock, W. C. and Rotunno, R. (1993) Frontal dynamics near and following frontal collapse. *Journal of the Atmospheric Sciences*, **50**, 3194–3212.

Uccellini, L. W. and Koch, S. E. (1987) The synoptic setting and possible energy sources for mesoscale wave disturbances. *Monthly Weather Review*, **115**, 721–729.

Uccellini, L. W., Kocin, P. J., Petersen, R. A., Wash, C. H. and Brill, K. F. (1984) The Presidents' Day cyclone of 18–19 February 1979: Synoptic overview and analysis of the subtropical jet streak influencing the pre-cyclogenetic period. *Monthly weather review*, **112**, 31–55.

Vanneste, J. (2004) Inertia–gravity wave generation by balanced motion: Revisiting the Lorenz–Krishnamurthy model. *Journal of the Atmospheric Sciences*, **61**, 224–234.

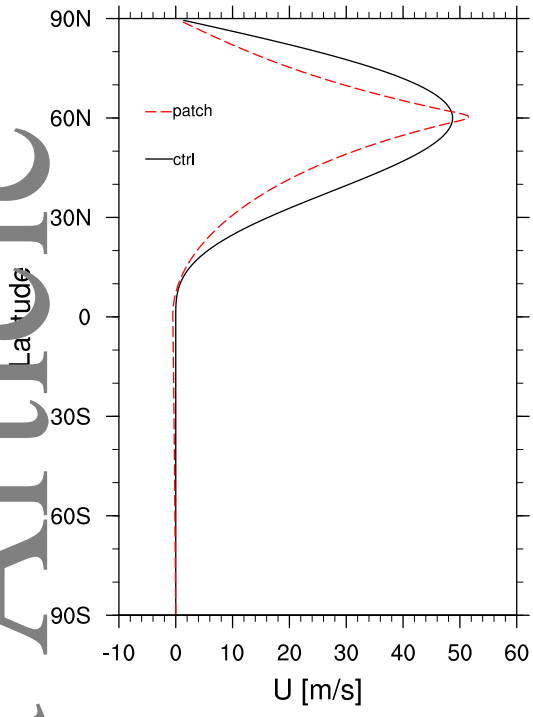

FIGURE 1 Initial height-independent zonal wind profiles (in [m/s]) for the control case (solid black line) and the potential vorticity patch case (dashed red line).

Vanneste, J. and Yavneh, I. (2004) Exponentially small inertia–gravity waves and the breakdown of quasigeostrophic balance. *Journal of the Atmospheric Sciences*, **61**, 211–223.

Viúdez, Á. (2007) The origin of the stationary frontal wave packet spontaneously generated in rotating stratified vortex dipoles. *Journal of Fluid Mechanics*, **593**, 359–383.

Viúdez, Á. and Dritschel, D. G. (2006) Spontaneous generation of inertia–gravity wave packets by balanced geophysical flows. *Journal of Fluid Mechanics*, **553**, 107–117.

Yasuda, Y., Sato, K. and Sugimoto, N. (2015) A theoretical study on the spontaneous radiation of inertia–gravity waves using the renormalization group method. Part II: Verification of the theoretical equations by numerical simulation. *Journal of the Atmospheric Sciences*, **72**, 984–1009.

Yoshiki, M., Kizu, N. and Sato, K. (2004) Energy enhancements of gravity waves in the Antarctic lower stratosphere associated with variations in the polar vortex and tropospheric disturbances. *Journal of Geophysical Research: Atmospheres*, **109**.

Yoshiki, M. and Sato, K. (2000) A statistical study of gravity waves in the polar regions based on operational radiosonde data. *Journal of Geophysical Research: Atmospheres*, **105**, 17995–18011.

Zhang, F. (2004) Generation of mesoscale gravity waves in upper-tropospheric jet–front systems. *Journal of the atmospheric sciences*, **61**, 440–457.

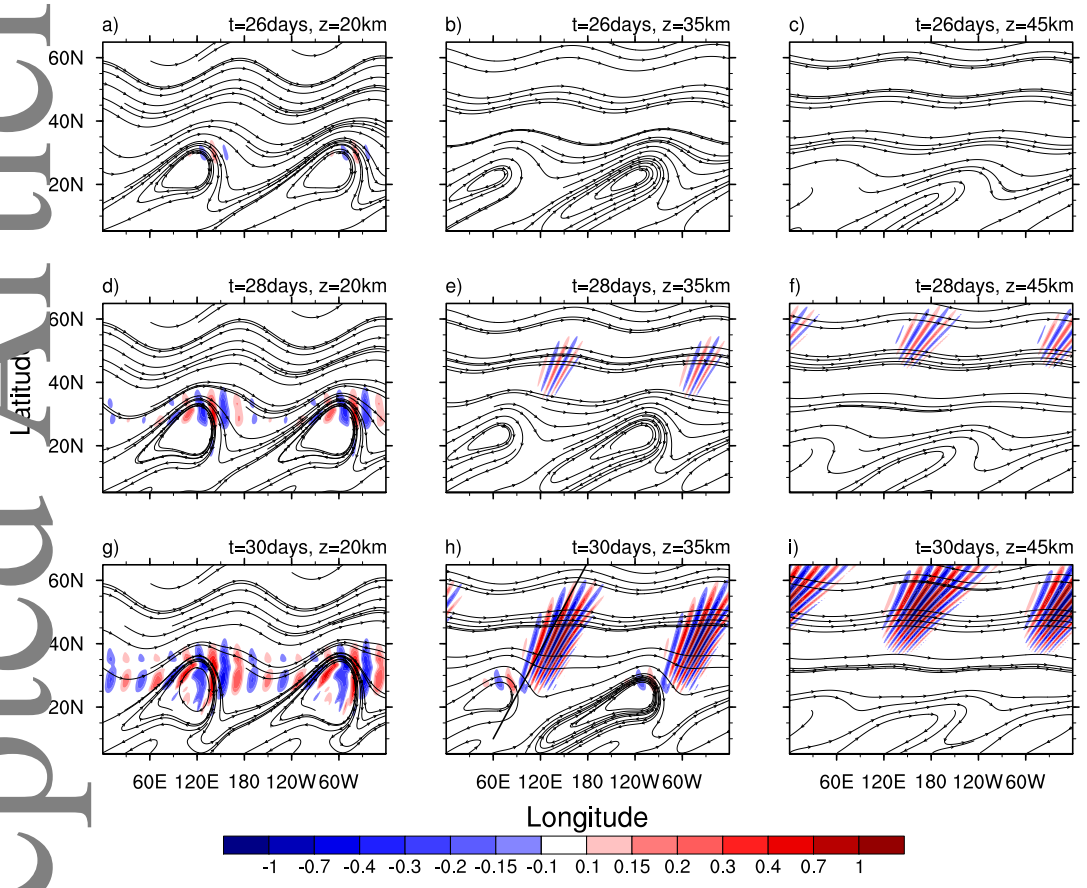

FIGURE 2 Longitude-latitude cross sections of filtered divergence 3×10^6 (shading, [1/s]) and streamlines [m/s]) at (a-c) t=26 days, (d-f) t=28 days, and (g-i) t=30 days into the evolution at altitude (a,d,g) z=20 km, (b,e,h) z=35 km, and (c,f,i) z=45km, respectively. Note inertia-gravity wave generation in the jet exit region at z=20km. Part of the gravity wave remains in the altitude and latitude region of the cat's eye, and part propagates upward and north-eastward. The height z is estimated as $z = -H \log[p/p_0]$, where p_0 is 1000 hPa, and $H = 7$km.

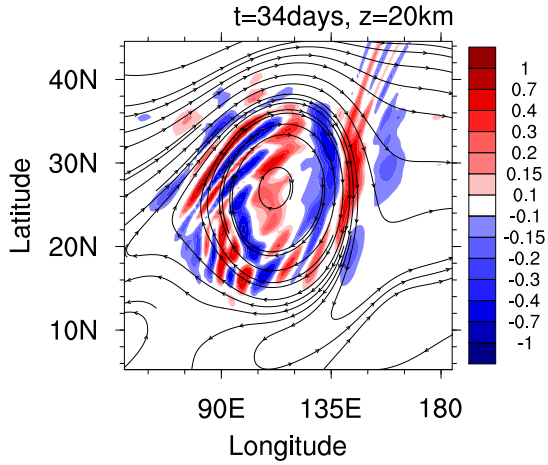

FIGURE 3 Longitude-latitude cross section of filtered divergence 2×10^6 (shading, [1/s]) and streamlines [m/s]) at z=20 km and t=34 days into the evolution, illustrating wave capture. Close-up view of the cat's eye is shown.

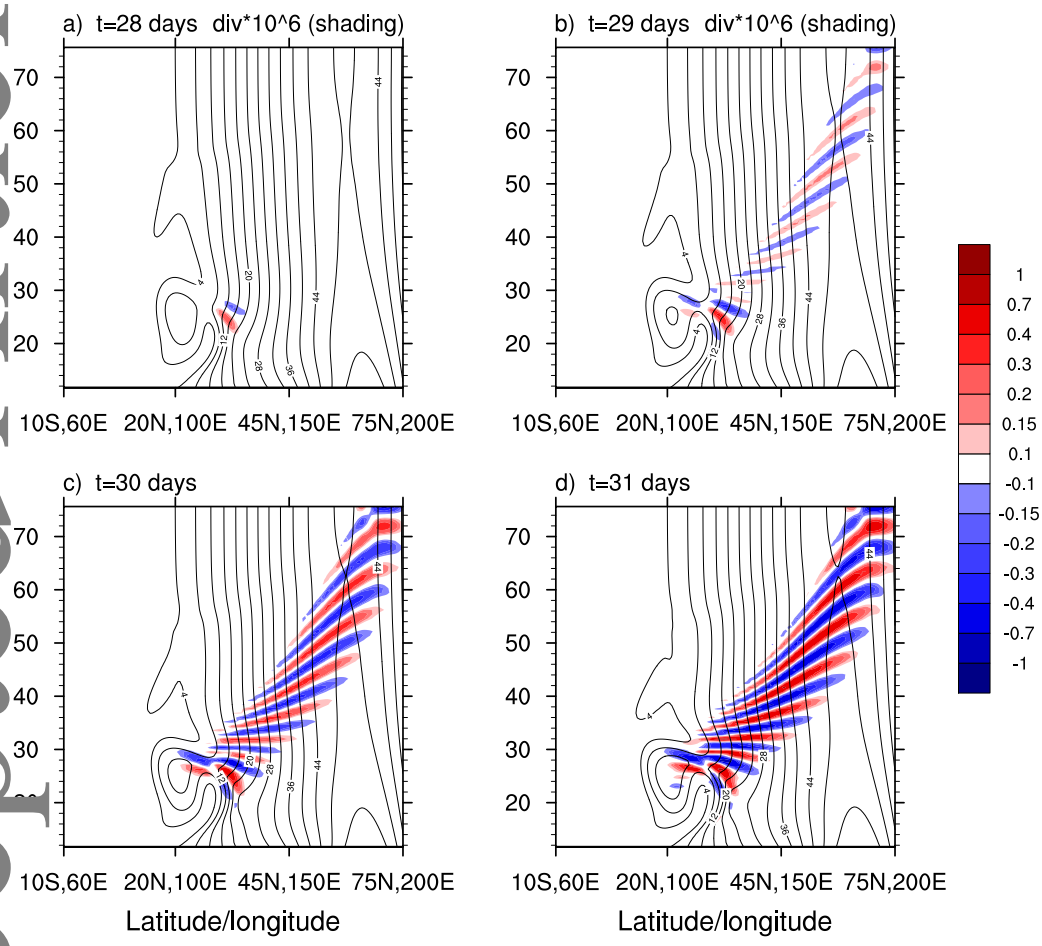

FIGURE 4 Vertical cross sections of filtered divergence $\times 10^6$ (shading, [1/s]) and horizontal wind magnitude $\sqrt{u^2 + v^2}$ (contours, [m/s]) through the line segment shown in Fig. 2h at (a) t=28 days, (b) t=29 days, (c) t=30 days, and (d) t=31 days, respectively. Contour interval for wind magnitude is 4 m s^{-1}.

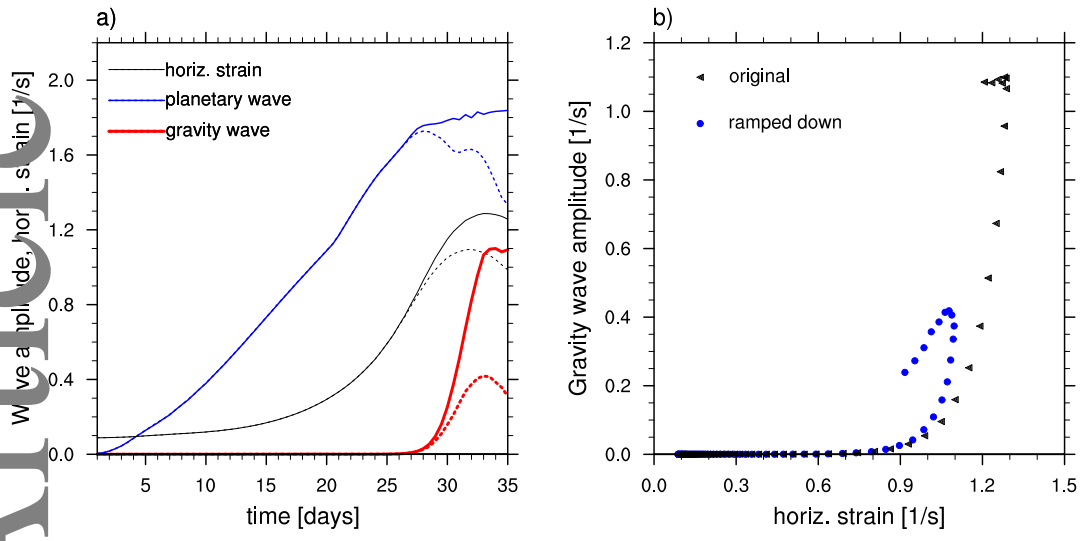

FIGURE 5 (a) Time evolution of "leaked" IGW amplitude 5×10^5 (red solid line, [1/s]) at z=35 km, and, planetary wave-2 amplitude $\times 10^5$ (blue solid line, [1/s]) and horizontal strain 2.5×10^4 (black solid line, [1/s]) in the source region at z=20 km (see text for definition). The multiplication factor for each line is chosen to show all lines in one panel. Dashed lines show the fields for simulation where the planetary wave-2 forcing is ramped down from t=25 days onwards. (b) IGW amplitude versus horizontal strain for the original simulation (triangles) and the simulation where the planetary wave-2 forcing is ramped down from t=25 days onwards (circles).

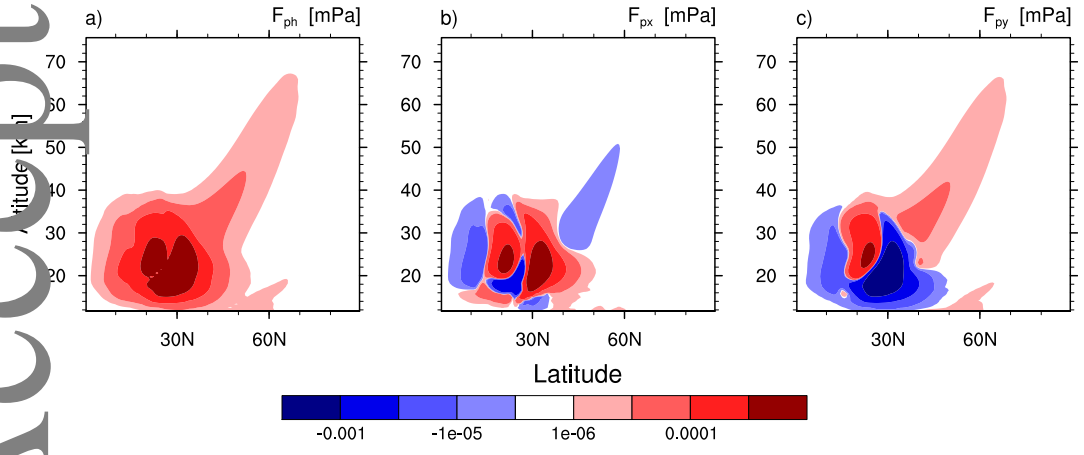

FIGURE 6 Latitude-altitude distribution of the vertical flux of horizontal momentum [mPa] averaged in time over days t=[26,32]. (a) Total vertical flux of horizontal momentum; (b) vertical flux of zonal momentum; and (c) vertical flux of meridional momentum. Note the non-linear contour interval to accentuate distribution at higher levels.

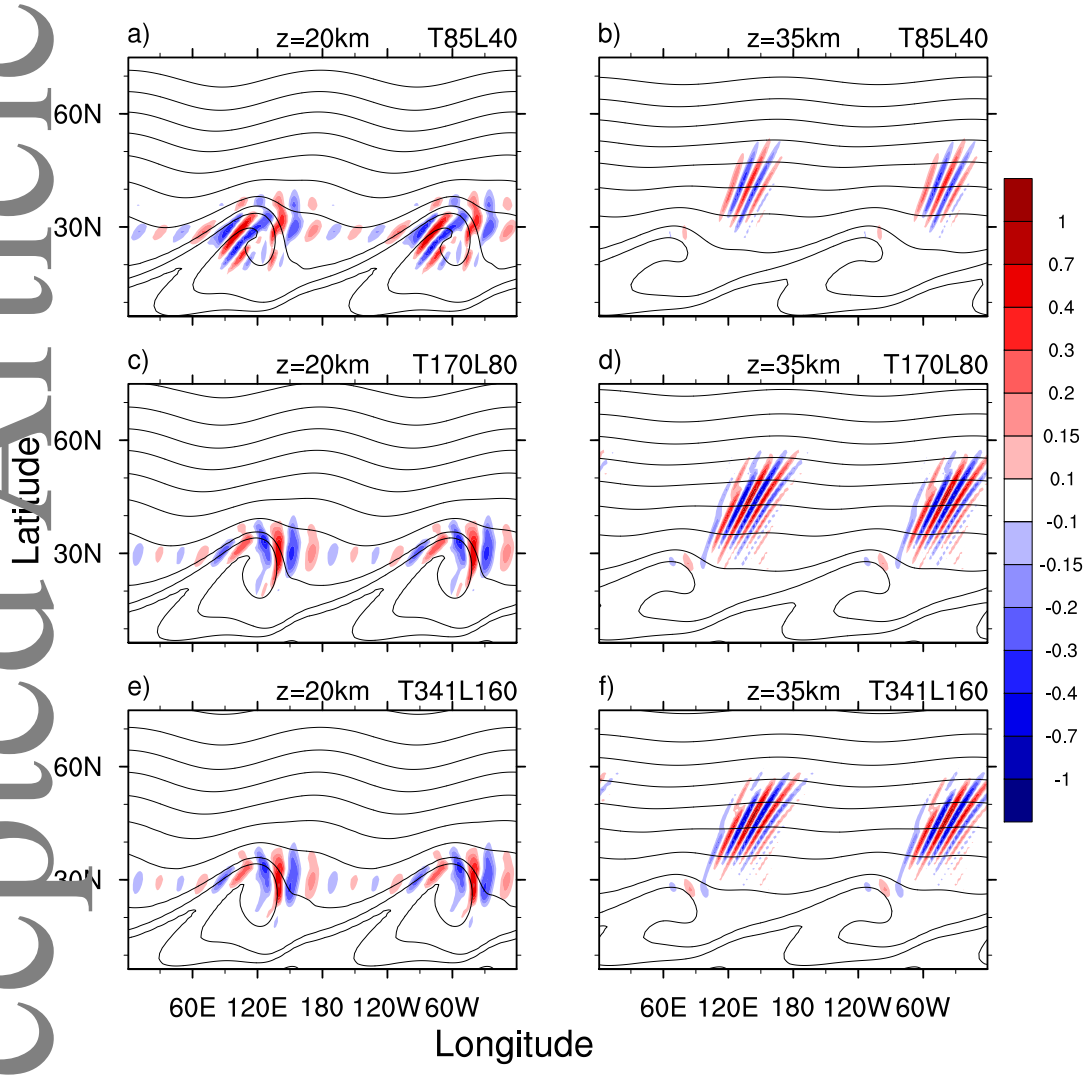

FIGURE 7 Longitude-latitude cross sections of filtered divergence 3×10^6 (shading, [1/s]) and Ertel potential vorticity [K $m^2 kg^{-1} s^{-1}$]) at t=30 days at altitude (right column) z=20 km and at (right column) z=35 km for different resolutions: (a-b) T85L40; (c-d) T170L80; and (e-f) T341L160. All the fields have been interpolated onto 128×256 Gaussian grid. Note convergence at T170L80 resolution.

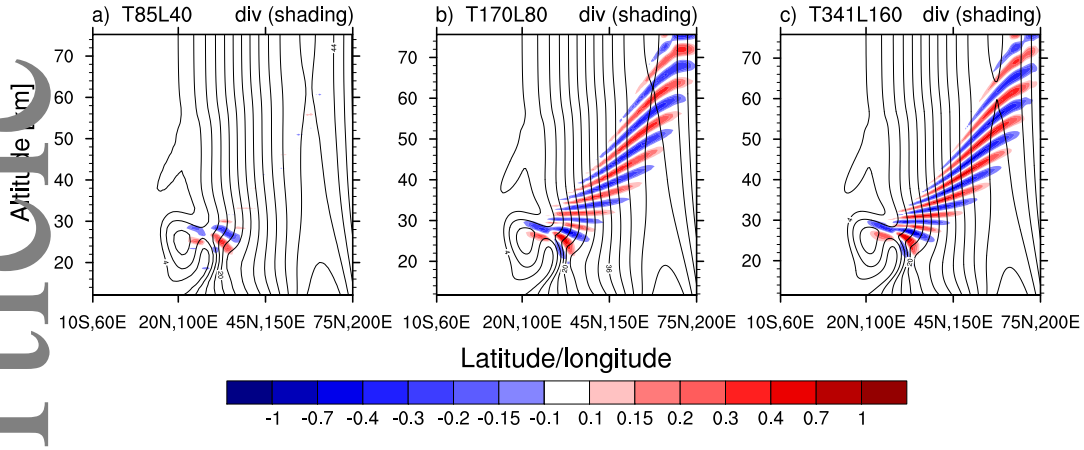

FIGURE 8 Vertical cross sections of filtered divergence $\times 10^6$ (shading, [1/s]) and horizontal wind magnitude (contours, [m/s]) through the line segment shown in Fig. 2 h at t=30 days for different resolutions: (a) T85L40; (b) T170L80, and (c) T341L160. Contour interval for wind magnitude is 4 m s^{-1}. Note convergence at T170L80 resolution. Panel b) is the same as Fig. 2c).

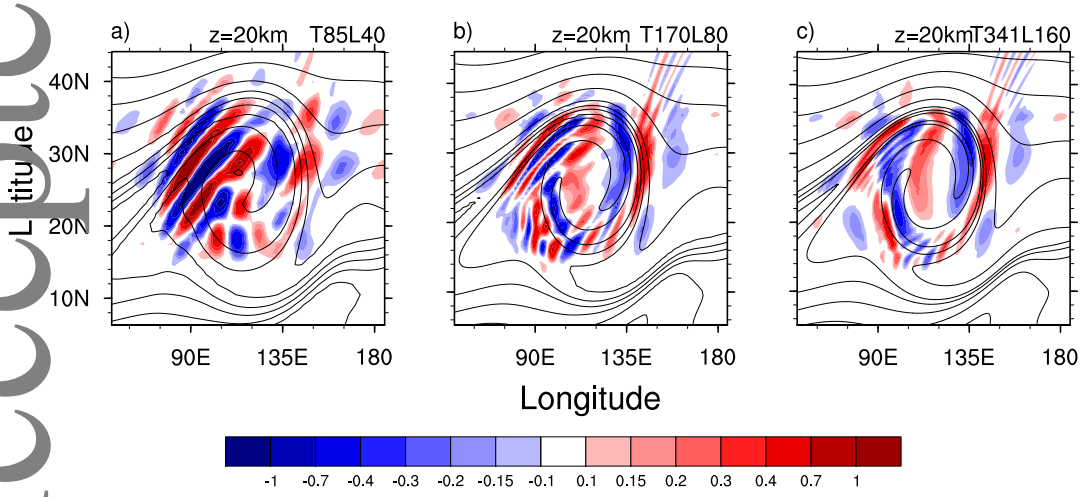

FIGURE 9 Longitude-latitude cross section of filtered divergence 2×10^6 (shading, [1/s]) and Ertel potential vorticity (contours, [K m^{2}kg^{-1}s^{-1}]) in the cat's eye at t=34 days and z=20 km for (a) T85L40; (b) T170L80; and (c) T341L160 resolution.

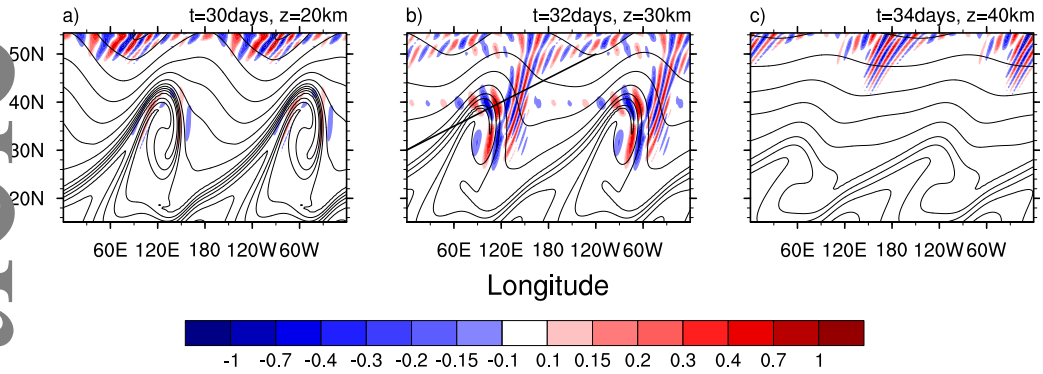

FIGURE 10 Longitude-latitude cross sections of filtered divergence 2×10^6 (shading, [1/s]) and Ertel potential vorticity (contours, [K m^{2}kg^{-1}s^{-1}]) for vortex patch initial condition (red line in Fig. 1). Evolution is shown at (a) t=30 days and z=20 km, (b) t=32 days and z=30 km, and (c) t=34 days and z=40 km, respectively. Note the different latitude range shown in this figure compared to Figs. 2 and 7.

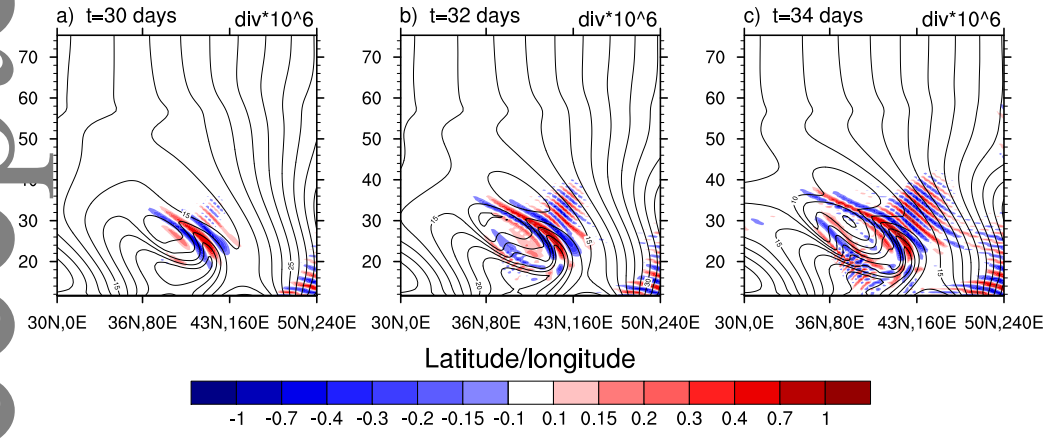

FIGURE 11 Vertical cross sections of filtered divergence $\times 10^6$ (shading, [1/s]) and horizontal wind magnitude $\sqrt{u^2 + v^2}$ (contours, [m/s]) through the line segment shown in Fig. 10b at (a) t=30 days, (b) t=32 days and (c) t=34 days, respectively. Contour interval for wind magnitude is 4 m s^{-1}.