Dynamics of Atmospheres and Oceans 78 (2017) 137–151

Contents lists available at ScienceDirect

Dynamics of Atmospheres and Oceans

journal homepage: www.elsevier.com/locate/dynatmoce

Vortex erosion in a shallow water model of the polar vortex

Robin Beaumont, Frank Kwasniok*, John Thuburn

College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF, UK

ARTICLE INFO

Article history:
Received 21 June 2016
Received in revised form 21 February 2017
Accepted 11 April 2017
Available online 13 April 2017

Keywords:
Polar vortex
Stratospheric sudden warming
Vortex erosion
Stripping of filaments of air
PV contour integral diagnostics

ABSTRACT

The erosion of a model stratospheric polar vortex in response to bottom boundary forcing is investigated numerically. Stripping of filaments of air from the polar vortex has been implicated in the occurrence of stratospheric sudden warmings (SSWs) but it is not understood in detail what factors determine the rate and amount of stripping. Here a shallow water vortex forced by topography is used to investigate the factors initiating stripping and whether this leads the vortex to undergo an SSW. It is found that the amplitude of topographic forcing must exceed some threshold (of order 200–450 m) in order for significant stripping to occur. For larger forcing amplitudes significant stripping occurs, but not as an instantaneous response to the forcing; rather, the forcing appears to initiate a process that ultimately results in stripping several tens of days later. There appears to be no simple quantitative relationship between the amount of mass stripped and the topography amplitude. However, at least over the early stages of the experiments, there is a good correlation between the amount of mass stripped and the global integral of wave activity, which may be interpreted as a measure of the accumulated topographic forcing. Finally there does not appear to be a simple correspondence between amount of mass stripped and the occurrence of an SSW.

1. Introduction

The stratospheric polar vortex is a dominant feature of the winter stratosphere. Variations in the polar vortex, such as stratospheric sudden warmings (SSWs) can have effects that influence surface climate (Baldwin and Dunkerton, 1999). At the start of winter a large cyclonic vortex is formed by radiative forcing. During the winter filaments of air are stripped from the edge of the vortex, reducing the area of the vortex and sharpening its edge (Mariotti et al., 1994), and are mixed into the surrounding region forming a relatively well mixed 'surf zone' (McIntyre and Palmer, 1984). The stripping process has implications for chemical transport (Waugh et al., 1994; Shepherd, 2007) and is also thought to 'precondition' the vortex, making it more susceptible to sudden warmings (Labitzke, 1981; Limpasuvan et al., 2004). However, it is not understood in detail what factors determine the rate and amount of stripping.

Shallow water models have long been used to investigate vortex dynamics, especially those of the polar vortex. The main advantage of using a shallow water model over more realistic, multi-layer models is that a larger area of the parameter space can be explored. As much of the motion in the stratosphere is along isentropic surfaces shallow water models are also well suited to investigating stratospheric dynamics.

* Corresponding author.
E-mail address: F.Kwasniok@exeter.ac.uk (F. Kwasniok).

http://dx.doi.org/10.1016/j.dynatmoce.2017.04.003

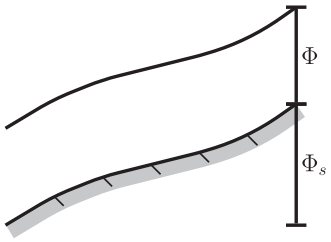

Fig. 1. The free surface geopotential is the sum of the geopotential of the topography Φ_s and the geopotential of the fluid layer Φ.

There have been two main approaches to investigation of vortex erosion in these experiments. Some studies have imposed a wave-like forcing on the lower boundary, while others have embedded a vortex in an externally imposed flow. Mixing in a model vortex under wave-like forcing has been investigated (e.g. Norton, 1994; Sobel and Plumb, 1999), and the effect of the forcing amplitude in single and multi-layer models (e.g. Polvani and Saravanan, 2000; Matthewman and Esler, 2011; Esler and Matthewman, 2011). The amount of mass mixed from the vortex into the surf zone has been found to be much less than that mixed from the tropics into the surf zone.

Matthewman and Esler (2011) looked at the effect of topographic forcing height and significance of background flow on the onset of vortex splitting. They found that for a given background flow, when the forcing increases past a particular value a bifurcation takes place causing the system to transition into a state in which a splitting can occur. This suggests that there may be a threshold in the amplitude of the forcing required to produce a vortex displacement or splitting event.

The other main approach is to embed a vortex in an externally imposed barotropic flow (e.g. Legras and Dritschel, 1993; Mariotti et al., 1994; Legras et al., 2001). This simulates the effects that other vortices have on the erosion of the polar vortex. In the presence of these externally imposed flows, if the externally imposed shear is very weak the vortex remains intact. If the externally imposed flow is increased but remains weak then the contours near the edge of the vortex become stripped until it can maintain a quasi-steady state. Under strong externally imposed flows the vortex can no longer maintain a steady state and breaks down. The first two regimes show that when the externally imposed strain growth is slow enough the vortex goes through a series of equilibrium states. This gives rise to the question in the third bullet point below, which will be investigated using a bottom boundary forcing rather than the externally imposed flow of Legras et al. (2001).

The main questions of interest in the following experiments are:

- Is there a threshold in the forcing amplitude to initiate stripping?
- Is there a quantitative relationship between the forcing amplitude and the amount of stripping?
- When the forcing increases at a slow enough rate does the vortex go through a series of quasi-steady states?
- What are the criteria for the vortex to break down; is there some threshold forcing amplitude above which the vortex breaks down?
- Is there a systematic difference between the response of the vortex to wave 1 and wave 2 forcings?

Answering these questions will aid understanding of the dynamics of the polar vortex and the conditions needed for SSWs to occur. In Section 2 the details of the model and experiments will be explained, and in Section 3 the integral diagnostics used are described. The results are shown in Section 4 and in Section 5 the conclusions are presented.

2. Experimental details

Stripping is investigated using the finite element shallow water model of Thuburn and Cotter (2015). The model uses a finite element based method to solve the rotating shallow water equations on a sphere and is capable of solving the dynamical equations (see below) on an unstructured grid. Here a cubed sphere grid is used.

The dynamical equations in the model take the form

$$\frac{\partial \Phi}{\partial t} + \nabla \cdot (\Phi \mathbf{u}) = 0 \tag{1}$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{k} \times (\Phi Q \mathbf{u}) + \nabla \left(\Phi + \Phi_s + \frac{1}{2} |\mathbf{u}|^2 \right) = 0 \tag{2}$$

where $\Phi = gh$ is the geopotential, g is the gravitational acceleration, h is the fluid depth, Φ_s is the geopotential of the topography, $\mathbf{u} = (u, v, 0)$ is the horizontal velocity and Q is the potential vorticity (PV). The free surface geopotential is $\Phi + \Phi_s$ (see Fig. 1). The prognostic variables of the model are Φ and $\mathbf{u}$. In the shallow water equations the PV can be formulated as

$$Q = \frac{\zeta + f}{\Phi} \tag{3}$$

where ζ is the relative vorticity and f is the Coriolis parameter.

The numerics of the model are formulated in such a way that the PV evolves as if the PV conservation law

$$\frac{\partial}{\partial t}(\Phi Q) + \nabla \cdot (\Phi Q \mathbf{u}) = 0 \tag{4}$$

itself were to be integrated rather than the dynamical Eqs. (1) and (2). This means that an accurate and conserving advection scheme may be used while maintaining desirable PV advection properties. This, in turn, helps to ensure that the diagnostics discussed in Section 3 are not contaminated by numerical artefacts.

In the model there are no explicit diabatic effects or frictional terms meaning that mass transport across contours of PV is only due to small-scale mixing within the advection scheme. When features of PV contours reach scales which are below those resolved by the model they become mixed into the background flow.

The topography Φ_s consists of a zonally symmetric part, $\overline{\Phi}$, corresponding to the initial condition plus a time-varying mountain, Φ_{pert}, which imposes a wave-like forcing on the flow. The topography then takes the form

$$\Phi_s = \overline{\Phi} + \Phi_{\text{pert}}. \tag{5}$$

The mountain is described by the equation

$$\Phi_{\text{pert}} = gH(t)M(\lambda, \phi) \tag{6}$$

with λ denoting geographical longitude and ϕ denoting geographical latitude. The height of the mountain at time t is given as

$$H(t) = \begin{cases} \min\left\{\frac{1}{2}\left[1 - \cos\left(\frac{\pi t}{\tau}\right)\right] H^*, H_{\max}\right\}; & t \le \tau \\ H_{\max}; & t > \tau. \end{cases} \tag{7}$$

We choose $H^* = 2000$ m. τ is a parameter controlling the rate of growth of the mountain and $H_{\max}$ is the maximum height of the mountain ($0 \le H_{\max} \le H^*$). The height of the mountain is increased monotonically in time and then frozen at the maximum height $H_{\max}$ as soon as that height is attained. The full mountain height H^* would be reached at time $t = \tau$. $M(\lambda, \phi)$ is the shape of the mountain and takes the form

$$M(\lambda, \phi) = \exp\left[-\left(\frac{\phi - \phi_0}{\Delta\phi}\right)^2\right] \cos(m\lambda) \tag{8}$$

where m is the zonal wavenumber of the forcing, $\phi_0 = \frac{\pi}{3}$ and $\Delta\phi = \frac{\pi}{12}$. This mountain is similar in size and shape to the wave-1/wave-2 component of the geopotential in the wintertime stratosphere. A similar mountain has been used in previous work (e.g., Norton, 1994).

The topographic forcing here mimics the effect of vertically propagating planetary waves which disrupt and erode the polar vortex. The gradual increase of the forcing corresponds to a move from autumn into winter by ramping up longwave cooling. It is in line with observations which show a significant wave 1 amplitude of the geopotential in the polar stratosphere all year round, growing into winter, but with quite a lot of variability. Moreover, ramping up the forcing gradually ensures that the dynamics remain close to balance, and provides the cleanest experiment with the best hope of being able to interpret the results. Initialising the model with a large-amplitude mountain at the bottom boundary would trigger unrealistic large-amplitude gravity waves.

The fluid on top of the topography Φ_s initially takes a uniform depth, having geopotential $\Phi = 4\text{x}10^4$ m^2 s^{-2}. The initial velocity field is zonally symmetric with a zonal wind chosen to be similar to that observed in the stratosphere in a typical northern hemisphere polar winter. The zonal wind profile used in the model experiments is shown in Fig. 2 along with several observed zonal-mean zonal wind profiles. An analytical expression for the initial zonal wind profile is given in the appendix. The zonal mean part of the topography $\overline{\Phi}$ needed to balance the zonal velocity is then found by integrating Eq. (A.1).

3. Integral diagnostics

To examine stripping of mass from the vortex several diagnostic quantities will be used. These are Lagrangian integral diagnostics which are changed only by non-conservative processes. These quantities are suitable for helping to answer the questions posed above since they capture information about the vortex edge and irreversible PV mixing which is not well captured by Eulerian zonal-mean quantities (Thuburn and Lagneau, 1999).

3.1. Mass integral

The mass within the PV = Q contour is given by

$$\mathcal{M}(Q) = \int_{\text{PV} \ge Q} \Phi \, \mathrm{d}S \tag{9}$$

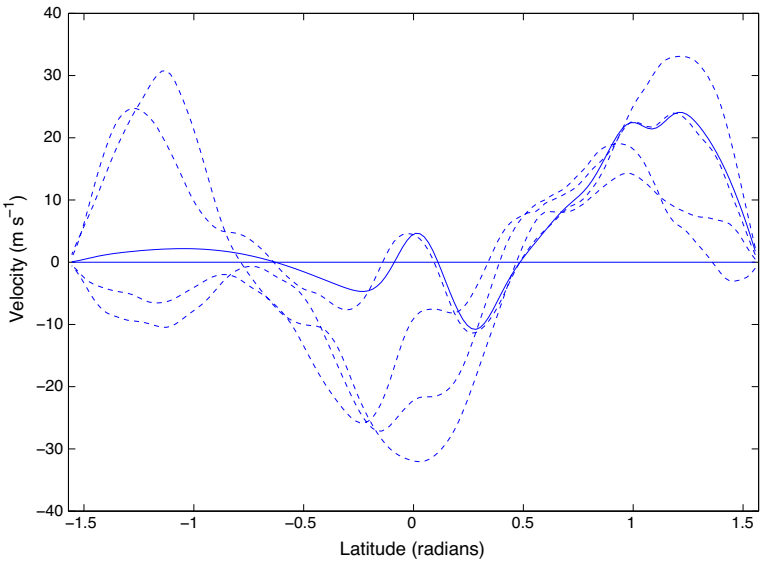

Fig. 2. Four examples of observed profiles of zonal-mean zonal wind at 60 °N and $p = 10$ hPa for different times during several winters (dashed), and the velocity profile used in the model experiments (solid).

as in Thuburn and Lagneau (1999). This is in fact a volume (multiplied by the gravity term g which is constant) and not a mass, but if density is taken to be constant unity then this quantity will have all the features of a true mass integral.

Since there is no diabatic heating (i.e. no mass source term in Eq. (1)), the amount of mass crossing a given PV contour is equal to the change of mass within that contour. Thus, transport across PV contours can be inferred from time series of $\mathcal{M}(Q)$. This will be used here as a Lagrangian diagnostic of vortex erosion by focussing on particular PV contours close to the vortex edge.

3.2. Circulation

The circulation around a PV=Q contour is defined as

$$\mathcal{C}(Q) = \oint_{\mathrm{PV}=Q} \mathbf{u}_{\mathrm{abs}} \cdot \mathrm{d}\mathbf{r} \tag{10}$$

where $\mathbf{u}_{\mathrm{abs}}$ is the fluid velocity in absolute reference frame. Using Stokes' theorem and Eq. (3) this can be expressed as

$$\mathcal{C}(Q) = \int_{\tilde{Q}\geq Q} \Phi\tilde{Q}\,\mathrm{d}S \tag{11}$$

which is the form which will be calculated in the model.

3.3. SSWs in the shallow water model

The WMO definition of an SSW states that an SSW has occurred when the zonal-mean temperature gradient between 60° and 85° north at 10 hPa is reversed, accompanied by a reversal of the zonal-mean zonal winds in the same region.

In order to identify if a simulated sudden warming occurs in the shallow water model, and if so at what point it can be said to occur, the sign of the zonal-mean zonal winds at 60° north is calculated. Using Stokes' theorem this can be expressed as an area integral of the relative vorticity:

$$\oint_{\phi=60^\circ} \mathbf{u} \cdot \mathrm{d}\mathbf{r} = \int_{\phi\geq 60^\circ} \zeta\,\mathrm{d}S. \tag{12}$$

An SSW can then be defined to have occurred if/when the value of this integral changes sign.

3.4. Wave activity

Wave activity diagnostics for the model can be calculated using the integral diagnostics outlined above. The conservation equation for the wave activity in a shallow water model takes the form (Brunet and Haynes, 1996; Thuburn and Lagneau, 1999)

$$\frac{\partial \mathcal{A}}{\partial t} + \nabla \cdot \mathcal{F} = \mathcal{B} + \mathcal{D} \tag{13}$$

where $\mathcal{F}$ is the wave activity flux, $\mathcal{B}$ represents lower boundary forcing and $\mathcal{D}$ represents sources and sinks due to irreversible mixing and diabatic and frictional processes. The wave activity is

$$\mathcal{A} = -\frac{\Phi_e u_e \cos\phi}{g} - \frac{\Phi}{2\pi a g}\int_{Q_{ref}}^{Q} (Q-\tilde{Q})\frac{d}{d\tilde{Q}} m_{ref}(\tilde{Q})\,d\tilde{Q} \tag{14}$$

where $m_{ref}(Q)$ is the mass north of the latitude of the contour PV=Q in a zonally symmetric reference state and subscript e indicates departures from the reference state. Assuming $Q_{ref}(\phi)$ is monotonic then $m_{ref}(Q)$ is single-valued and $m_{ref}(Q) = \mathcal{M}_{ref}(Q)$.

Using the relationship between mass and circulation in PV-space (Thuburn and Lagneau, 1999),

$$\frac{\partial \mathcal{C}}{\partial Q} = Q\frac{\partial \mathcal{M}}{\partial Q}, \tag{15}$$

applied to the reference state, Eq. (14) can be written as

$$\mathcal{A} = -\frac{\Phi_e u_e \cos\phi}{g} - \frac{\Phi}{2\pi a g}\left\{Q\left[\mathcal{M}_{ref}(Q) - \mathcal{M}_{ref}(Q_{ref})\right] - \left[\mathcal{C}_{ref}(Q) - \mathcal{C}_{ref}(Q_{ref})\right]\right\}. \tag{16}$$

This expression consists of two terms, one of which, $-\frac{\Phi_e u_e \cos\phi}{g}$, approximately represents gravity waves while the other, $-\frac{\Phi}{2\pi a g}\left\{Q\left[\mathcal{M}_{ref}(Q) - \mathcal{M}_{ref}(Q_{ref})\right] - \left[\mathcal{C}_{ref}(Q) - \mathcal{C}_{ref}(Q_{ref})\right]\right\}$, is the approximate contribution from Rossby waves. The 'Rossby wave' term is several orders of magnitude ($\sim 10^4$) larger than the 'gravity wave' term, so in calculations of $\mathcal{A}$ it is sufficient to calculate only the contribution from the Rossby wave term. The Rossby wave term is easy to calculate once $\mathcal{M}$ and $\mathcal{C}$ are calculated, which are quantities of interest in themselves.

Since the initial state is a balanced flow the reference state is taken to be the initial state meaning that $Q_{ref} = Q_{init}$.

The integral of the wave activity $\mathcal{A}$ over the domain is related to the time integral of the lower boundary forcing (Thuburn and Lagneau, 1999) and can be interpreted as a measure of rearrangement of PV. This makes it a good candidate for being a useful quantity in studying vortex erosion.

4. Model results

The goal of our model experiments is to investigate vortex erosion by stripping of filaments of air as a function of three main parameters: the forcing amplitude given as the maximum mountain height H_{max}, the timescale of increase in topography τ and the zonal wavenumber m of the forcing.

Two sets of wave 1 experiments were performed. In the first H_{max} is fixed at 2000 m and the rate of growth of the forcing τ is varied. For a PV contour close to the vortex edge ($Q = 3.328 \times 10^{-9}$ s m^{-2}), the mass stripped is given by $\Delta\mathcal{M}(Q) = \mathcal{M}_{init}(Q) - \mathcal{M}(Q)$. For each run, both $\Delta\mathcal{M}(Q)$ and forcing height H are functions of time. Fig. 3 shows plots of $\Delta\mathcal{M}(Q)$ versus H for different values of τ.

To rule out numerical effects a range of resolutions were examined. Fig. 3 shows that for low resolutions numerical effects play a large role in the evolution of the vortex. This is seen for example in the bottom left panel where for $H < 1000$ m there is stripping of mass for all resolutions coarser than 52 km. When the resolution reaches 52 km the mass removed is close to zero; the numerical effects become small in comparison to other effects. The increase of mass stripping, seen at $H = 1000$ m in the bottom left panel coincides with the formation of tongues of air at the edge of the vortex.

In the second set of wave 1 experiments the value of τ is fixed at $\tau = 92.6$ days and the value of H_{max} is varied. The results of these experiments are shown in Fig. 4 for 2 PV contours on the vortex edge ($Q = 3.328 \times 10^{-9}$ s m^{-2} and $Q = 3.456 \times 10^{-9}$ s m^{-2}). The amount of mass removed is shown; negative values correspond to increases in mass. The vertical lines at the top of the plots show the times when each of the mountains reach their maximum heights. The two PV contours are highlighted in red in Fig. 5, and as they lie around the vortex edge the transports across them can be interpreted as changes in mass of the vortex.

4.1. Is there a threshold in the forcing amplitude to initiate stripping?

When the mountain is grown to only low heights there is a relatively small amount of mass removed from the vortex. For example when the mountain is grown to a height of 200 m the mass flux occurs at a very low rate (Fig. 4). The air surrounding

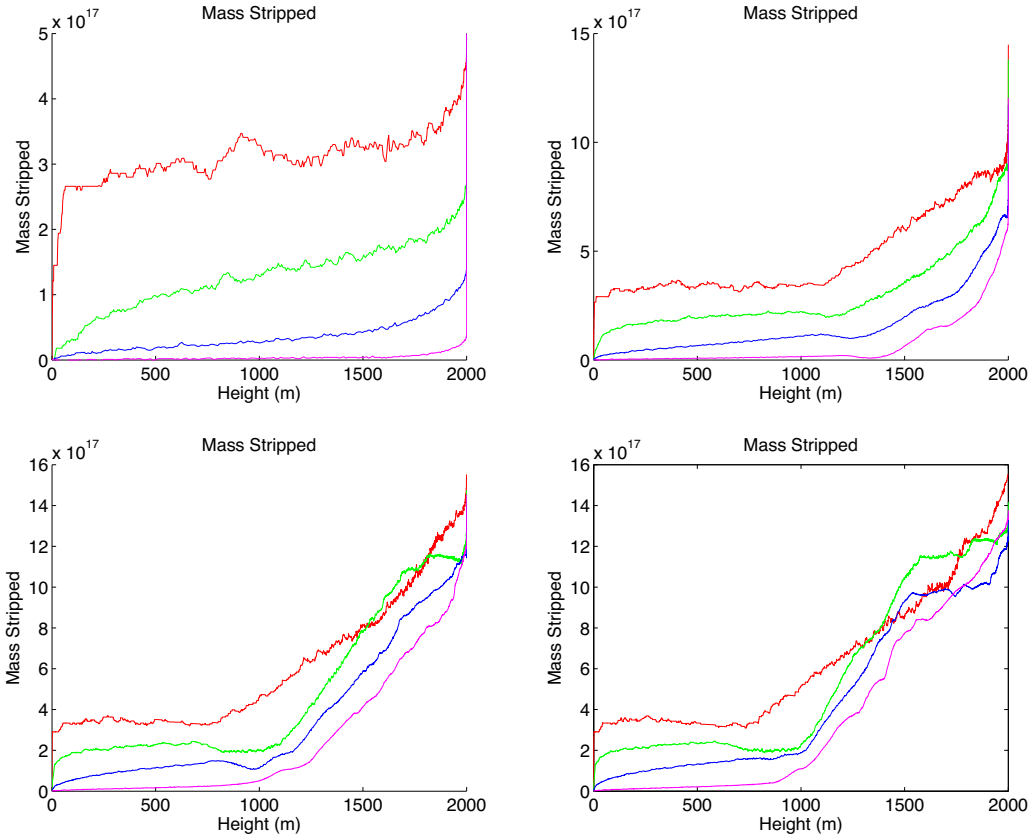

Fig. 3. Total mass removed $\Delta\mathcal{M}(Q)$ (in $m^4\,s^{-2}$) versus mountain height H for a PV contour on the vortex edge for $H_{max} = 2000$ m and $\tau = 11.6$ days (top left), $\tau = 46.3$ days (top right), $\tau = 92.6$ days (bottom left) and $\tau = 115.7$ days (bottom right). The red lines correspond to a resolution of 416 km, the green to 208 km, the blue to 104 km and the magenta to 52 km. The wavenumber of the topography is $m = 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

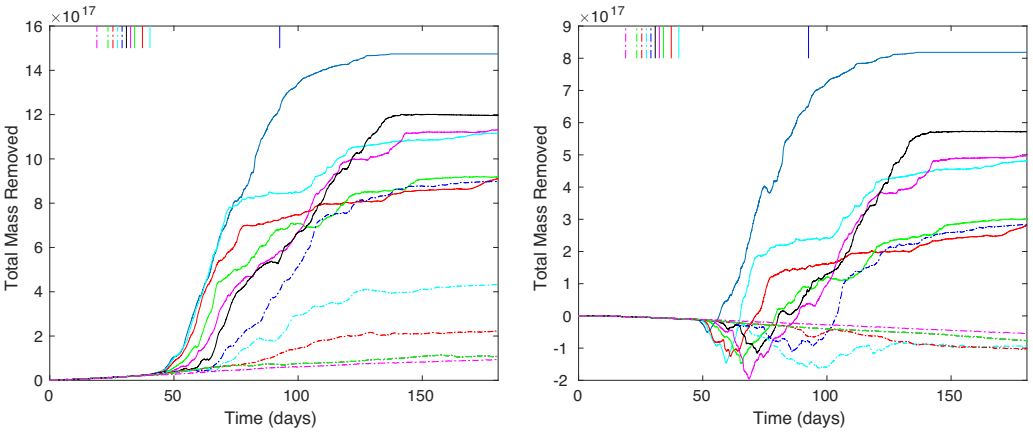

Fig. 4. Total mass removed $\Delta\mathcal{M}(Q)$ (in $m^4\,s^{-2}$) versus time for two PV contours on the vortex edge. The left panel refers to the outer red contour in Fig. 5 and the right panel to the inner red contour. The mountain is frozen at various maximum heights H_{max}: 2000 m (solid blue), 800 m (solid cyan), 700 m (solid red), 600 m (solid green), 550 m (solid magenta), 500 m (solid black), 450 m (dashed blue), 400 m (dashed cyan), 350 m (dashed red), 300 m (dashed green) and 200 m (dashed magenta). The mountain is grown with $\tau = 92.6$ days. The wavenumber of the topography is $m = 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

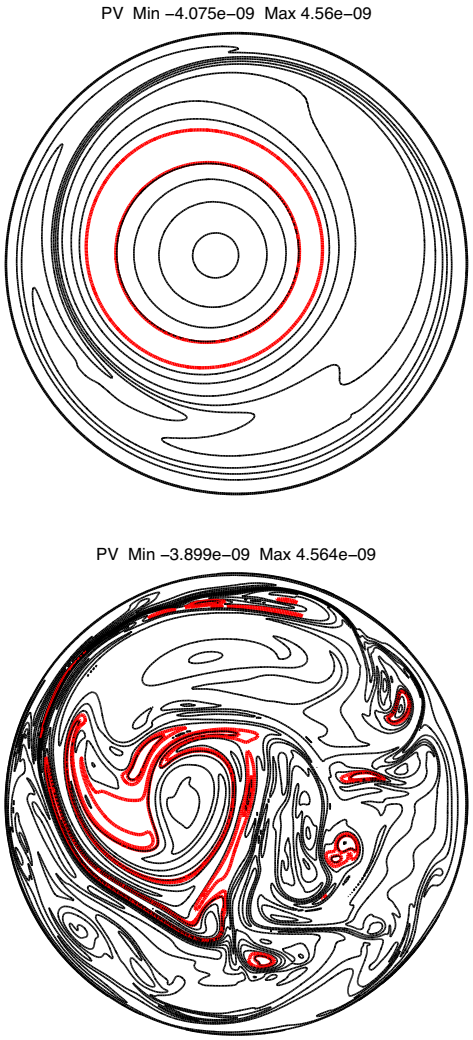

Fig. 5. PV maps (in $s\,m^{-2}$) before (top) and after (bottom) the onset of the model SSW for $H_{max} = 2000$ m, $\tau = 92.6$ days and $m = 1$. The black contours are evenly spaced, and the red contours indicate those referred to in Fig. 4. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the vortex becomes very turbulent after enough time has passed, and tongues of air can be seen to form at the edge of the vortex but no significant stripping occurs.

At larger heights the amount of mass removed from the vortex is very similar to the 200 m case for the first 70 model days. Following this the outward flux of mass increases. This increasing of the rate of mass loss occurs long after the mountain has reached its full height, particularly for cases with lower maximum mountain height. For mountains over 500 m stripping is initiated roughly 50 days after the start of the experiments.

Fig. 3 shows that stripping is not an instantaneous response to forcing; for the faster growing mountains there is no stripping until after the maximum height has been reached. This can also be seen in Fig. 4 where stripping does not occur until after the smaller mountains have reached their maximum height. The forcing does not initiate stripping directly, but initiates a process which then results in removal of mass from the vortex at some time later.

4.2. Is there a quantitative relationship between the forcing amplitude and the amount of stripping?

Once the rate of mass loss from the vortex increases, between 52 and 58 days, the average rate of loss is similar for all runs with a mountain of 500 m or more and continues at this rate for a period of 10 model days for all but the 2000 m mountain which proceeds at this rate for nearly twice the amount of time.

After a period of 140 model days many of the runs have reached a stage where little or no further mass flux across the contours occurs. In this case the flow is not steady, but $\mathcal{M}(Q)$ becomes almost unchanging at the edge of the vortex.

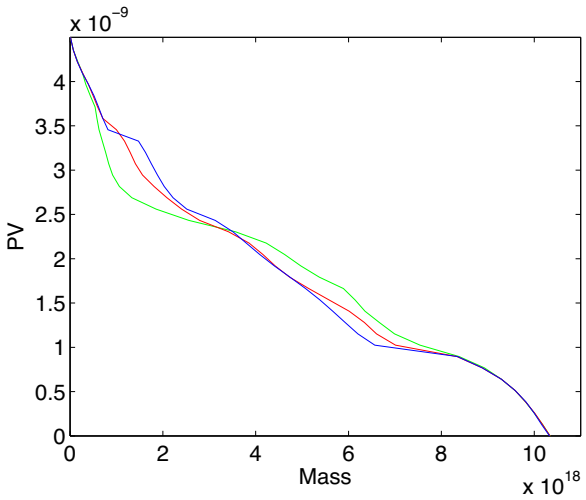

Fig. 6. PV (in $\mathrm{s\,m^{-2}}$) against mass $\mathcal{M}(Q)$ (in $\mathrm{m^4 s^{-2}}$) for $H_{\max} = 550$ m, $\tau = 92.6$ days and $m = 1$ at various times during the experiment. The blue line refers to the initial condition, the red line to 68 days and the green line to 105 days. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

There does not seem to be any threshold in the amount of mass removed from or added to either of the PV contours shown here before the rate of stripping starts to increase above the level of the 200 m mountain, rather it is the size of the forcing and not the amount of mass removed from or gained by the PV contour that is responsible for the contour starting to lose mass more rapidly. The final amount of mass removed from the vortex depends on $H_{\max}$ but there is not a linear relationship between the two quantities.

The PV contour shown in Fig. 4 (right) initially experiences an increase in mass, with the 200 m mountain case experiencing a gradual steady increase in mass for the entire period of the model run. When the mountain is grown to a height greater than 200 m the contour experiences a larger increase in mass, which persists for longer when the final mountain height is smaller. For those runs with a final mountain height over 400 m, following the initial gain in mass the contour then begins to decrease in mass to a level below that of the starting level.

As well as the period of the mass gain being longer as the final height of the mountain becomes smaller, the amount of mass gained by the PV contour also increases, attaining a maximum when the final height of the mountain is 550 m. For the runs where the mountain height is less than 450 m the amount of mass within this contour is larger at the end of the experiment than at the start as shown in Fig. 4 (right).

Some understanding of why the inner contour in Fig. 4 gains mass temporarily can be obtained by using the relationship between mass and circulation of Eq. (15) to show that the area under the curves in $\mathcal{M}$-Q plots should be conserved. On PV contours with negligible mixing the circulation will be conserved. Consider two such contours Q_1 and Q_2; for example the contours initially at the equator and north pole. Then, during the evolution, $\Delta\mathcal{C}(Q_1) = 0$ and $\Delta\mathcal{C}(Q_2) = 0$. Hence,

$$\begin{aligned} 0 = \Delta\mathcal{C}(Q_2) - \Delta\mathcal{C}(Q_1) &= \Delta \int_{Q_1}^{Q_2} \frac{\partial \mathcal{C}}{\partial Q}\, dQ \\ &= \Delta \int_{Q_1}^{Q_2} Q \frac{\partial \mathcal{M}}{\partial Q}\, dQ \\ &= \Delta \int_{\mathcal{M}(Q_1)}^{\mathcal{M}(Q_2)} Q\, d\mathcal{M}. \end{aligned} \tag{17}$$

This means that the area under the curves in $\mathcal{M}$-Q space between the PV values Q_1 and Q_2 must remain constant for all instants during the model run. It then follows that if the mass within some PV contours decreases, there must be other region(s) in PV space where the mass increases to compensate.

Fig. 6 shows PV versus mass for various instants during the run with 550 m wave 1 forcing. (The mass within PV contours in the southern hemisphere, not shown, is almost constant over the entirety of the model run.) The red line is for day 68, which is just before the inner contour in Fig. 4 reaches its maximum mass (day 70). Comparison with the initial $\mathcal{M}$-Q profile (blue) shows that there is one main area where contours decrease in mass, between 2.5 and $3.5 \times 10^{-9}\ \mathrm{s\,m^{-2}}$, and two areas that gain mass compared with the initial state; the small area near PV value $3.5 \times 10^{-9}\ \mathrm{s\,m^{-2}}$ corresponds to the inner contour in Fig. 4, while the larger area between 1 and $1.5 \times 10^{-9}\ \mathrm{s\,m^{-2}}$ corresponds to the beginning of the formation of the surf zone.

By day 105 the mass of the contours within the vortex (those above $2.4 \times 10^{-9}\ \mathrm{s\,m^{-2}}$) has decreased significantly. Here there is no region of compensating mass increase in the higher valued PV contours, instead there is a large increase in mass

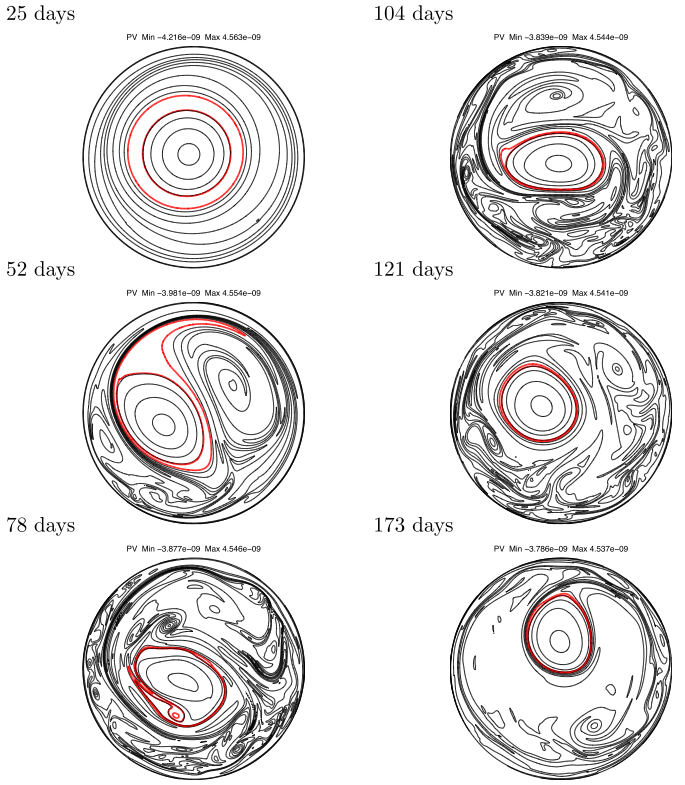

Fig. 7. PV maps (in $s\,m^{-2}$) at several times during the experiment for $H_{max} = 700$ m, $\tau = 92.6$ days and $m = 1$. The red contours have the same PV values as in Fig. 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

within the 1 to 2.4×10^{-9} $s\,m^{-2}$ region. This region clearly corresponds to the surf zone region which is by this point well established.

Following the period of rapid mass loss there is a dramatic reduction in the rate of stripping. The period of rapid mass loss continues for differing periods of time for different values of H_{max}.

Around the time of the sharp decrease in the rate of mass flux from the vortex in the case of the 700 m forcing the tongues of air attached to the vortex start to disappear; as time passes the existing tongues get pulled from the vortex and no new tongues form. This can be seen from 104 days in Fig. 7, where the remaining tongues are pulled from the vortex and the formation of new tongues is inhibited. The 700 m experiment is unusual in the fact that this occurs very soon after the rapid stripping commences. The edge of the vortex appears to become sharper and more oval shaped than is typical up until this point (see Fig. 7, day 104). The sharper vortex edge created by the removal of mass then appears to inhibit further tongues of air from forming leading to the dramatic reduction in the rate of mass removal. This is similar to what has been found in other shallow water experiments and termed 'Rossby elasticity' (Juckes and McIntyre, 1987; Norton, 1994). A latitudinal gradient of PV provides a restoring mechanism that tends to return displaced fluid parcels to their original latitude and gives the mechanism for Rossby wave propagation. As the PV gradient on the vortex edge becomes sharper the restoring mechanism (on scales comparable to the edge thickness) becomes stronger, thus inhibiting the formation of further tongues of air. The disappearance of tongues from the vortex edge does not occur until over 100 days for the other experiments.

At the time of the initial decrease in the rate of mass removal in Fig. 4, around 70 days, the tongues of air that are attached to the vortex have mostly disappeared and those which remain are very small compared to those previously present. Over the next 35 days there is a much reduced rate of mass removal where the parts of the vortex which form into small tongues are pulled from the vortex but do not pull any further mass with them, therefore disappearing very quickly after they form. This then makes the vortex into a more regular oval shape, gradually losing the non-oval shaped elements as they become eroded. After 100 days the vortex has lost all tongues attached to it, and the mass removal rate becomes negligible, around the same rate as that of the 200 m mountain. Following this there are short periods where tongues form and become eroded from the vortex. Following the erosion the vortex again becomes oval shaped. One of these can be seen in Fig. 7 at day 173.

In the cases of the other final mountain heights the same effect can be seen in the period where the rate of mass removal significantly decreases, around 115 model days in most cases. In the case of the 2000 m mountain the vortex has become almost completely eroded.

As noted above, the amount of mass stripped from the vortex does not appear to be proportional to the amplitude of the forcing. The wave activity $\mathcal{A}$ (Eq. (16)), however, can be related to a measure of the accumulated forcing.

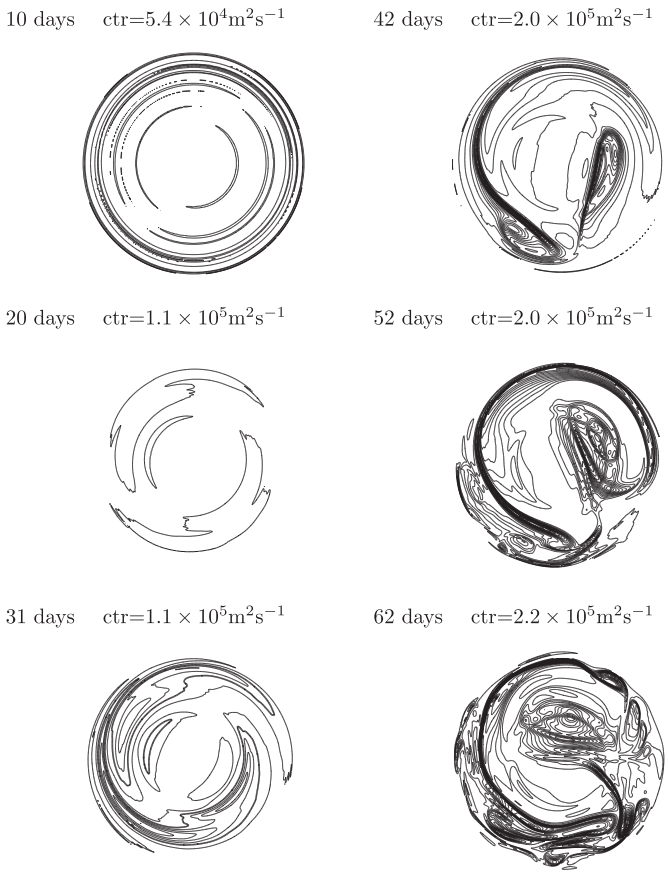

Fig. 8. Wave activity $\mathcal{A}$ (in $m^2 s^{-1}$) at various times during the experiment for $H_{max} = 700$ m, $\tau = 92.6$ days and $m = 1$. The contour interval is indicated by ctr.

Using the initial zonally symmetric state as the reference state in Eq. (16) the wave activity was calculated for various model runs. Fig. 8 shows the wave activity $\mathcal{A}$ for the wave 1 case with a mountain of height 700 m.

In the period up until the mountains reach their maximum heights the profiles of the wave activity are identical to that of the 2000 m mountain. As the smaller mountains reach their maximum heights the profile then diverges from that of the larger forcing.

The wave activity begins very small, but after a period of 20 model days a wave 2 pattern starts to emerge. The terms in the wave activity Eq. (16) are second order in disturbance quantities, so maps of $\mathcal{A}$ for a wavenumber 1 disturbance will show wavenumber 2 patterns, which is the case here for small amplitude forcing. As it starts to become more asymmetric $\mathcal{A}$ starts to pick out some of the anticyclones surrounding the vortex. For example at day 52 the area of the largest anticyclone is clearly picked out as well as several of the smaller ones.

These synoptic maps of wave activity $\mathcal{A}$ pick out many of the features seen in corresponding PV maps and do not seem to provide significantly more information than the PV does. The spatial integral of $\mathcal{A}$ is related to the space-time integral of mountain forcing as well as the amplitude of the disturbance and its phase relative to the mountain forcing. This means that the integrated wave activity $\mathcal{A}$ is a diagnostic quantity which contains information about several important features of the flow.

Fig. 9 shows the mass removed from the vortex plotted against the wave activity integrated over the area of the northern hemisphere, $\int_{\phi \geq 0^o} \mathcal{A}\, dS$, for a selection of the final mountain heights from the wave 1 experiments. The figure shows the part of the graph for the initial part (about 65 days) of the experiment where the integrated wave activity is increasing, in order to provide a monotonic x-coordinate. The points where the integrated wave activity reaches its maximum value for the 2000 m and the 700 m mountains are both very similar, with the wave activity for the 2000 m mountain reaching its maximum a couple of days earlier than that for the 700 m mountain. Following this the wave activity starts to decrease due to the mixing occurring around the vortex.

On the inner PV contour (right panel in Fig. 4 and Fig. 9) the period where mass is being entrained into the vortex corresponds almost exactly to the period when the integrated wave activity is increasing. Other than the case of the 2000 m mountain, where the mass flux changes from inward flux to outward flux slightly before the wave activity starts to decrease, the moment when the wave activity starts to decrease is the same moment as the change from inward to outward mass flux.

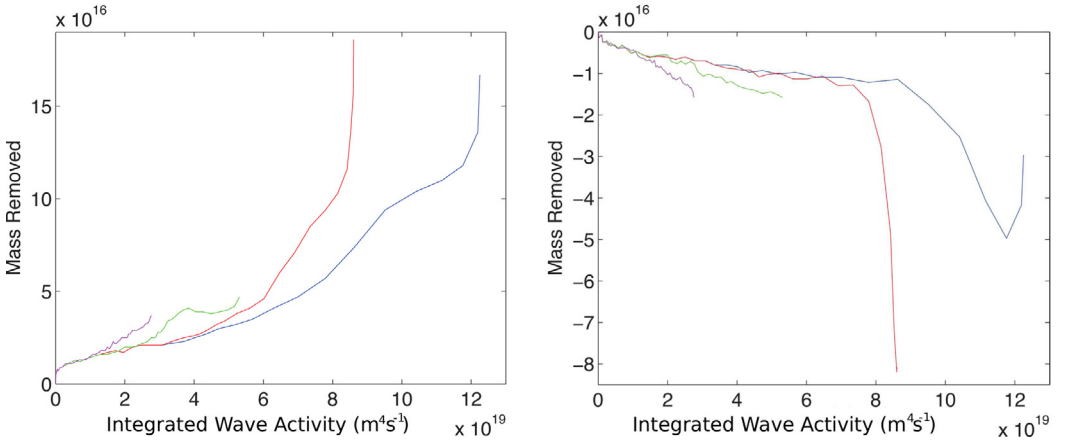

Fig. 9. Total mass removed $\Delta\mathcal{M}(Q)$ (in $m^4 s^{-2}$) against integrated wave activity $\mathcal{A}$ for the outer red PV contour (left) and the inner red PV contour (right) for various mountain heights H_{max}: 2000 m (blue), 700 m (red), 450 m (green) and 200 m (magenta). The negative values in the right panel correspond to mass added to the contour. The mountain is grown with $\tau = 92.6$ days. The wavenumber of the topography is $m = 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

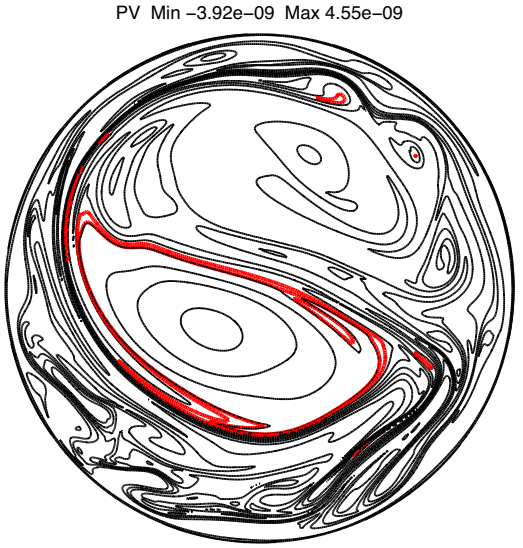

Fig. 10. PV map (in $s\,m^{-2}$) showing the displaced vortex following a model SSW for $H_{max} = 600$ m, $\tau = 92.6$ days and $m = 1$. The vortex can be seen in the lower left part of the plot with the characteristic 'comma shape' that is observed in SSWs in the real atmosphere.

The mass increase is approximately linearly correlated with integrated wave activity up until the wave activity becomes large, from around $8x10^{19}$ m^4 s^{-1}, when the rate of mass increase becomes significantly larger.

4.3. When the forcing increases at a slow enough rate does the vortex go through a series of quasi steady states?

The hypothesis that the vortex evolves through a series of quasi-steady states for a slow enough forcing growth rate would suggest that the amount of mass removed from the vortex is the same function of the mountain height for differing values of τ. From the model results in Fig. 3 this is not the case. As the rate of growth of the forcing decreases the erosion of the vortex occurs at lower forcing amplitudes. If the vortex were to go through a series of quasi-steady states we would expect to see that the total amount of mass removed is related to the forcing amplitude which is not the case in these experiments.

4.4. What are the criteria for the vortex to break down; is there some threshold forcing amplitude above which the vortex breaks down?

Using the criterion of Eq. (12), when the mountain is grown to heights of 450 m or higher the sign of the vorticity integral given by Eq. (12) reverses at some time during the experiment, indicating that an SSW-like event has occurred, and the vortex can be seen to have displaced from the pole (see Fig. 10 for the 600 m mountain for example). When the mountain is

Table 1
Times of occurrence of model SSWs for $\tau = 92.6$ days, $m = 1$ and various maximum mountain heights H_{max}.

Mountain height H_{max}	SSW time
2000 m	72 days
800 m	104 days
600 m	126 days
500 m	73 days
400 m	–

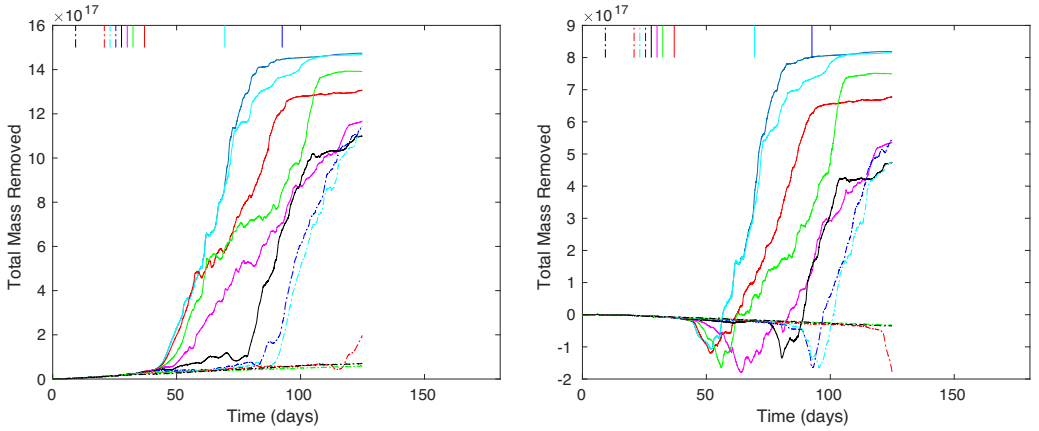

Fig. 11. Total mass removed $\Delta\mathcal{M}(Q)$ (in $m^4 s^{-2}$) versus time for the same two PV contours on the vortex edge as in Fig. 4. The left panel refers to the outer contour and the right panel to the inner contour. The mountain is frozen at various maximum heights H_{max}: 2000 m (solid blue), 1500 m (solid cyan), 800 m (solid red), 700 m (solid green), 650 m (solid magenta), 600 m (solid black), 550 m (dashed blue), 500 m (dashed cyan), 450 m (dashed red), 400 m (dashed green) and 200 m (dashed black). The mountain is grown with $\tau = 92.6$ days. The wavenumber of the topography is $m = 2$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

grown to lower heights (400 m and below) the sign of the vorticity integral does not change indicating that for these forcing amplitudes the model vortex does not undergo an SSW event.

The time of occurrence of an SSW as a function of the maximum mountain height H_{max} is listed in Table 1. As with the total amount of mass removed from the vortex the timing of the model equivalent of an SSW does not follow a linear relationship to the final mountain height. The 2000 m mountain is the first to experience a model SSW around 70 days from the start of the run. There does not seem to be a threshold for the amount of mass removed from the vortex (or equivalently the total mass within the vortex) for an SSW to occur. Shortly after the onset of the warming in the 2000 m case an SSW occurs for the 500 m mountain. The mass within the vortex at this point is much larger than that of the 2000 m case.

4.5. Is there a systematic difference between the response of the vortex to wave 1 and wave 2 forcings?

Fig. 11 shows the results for the same experiments as Fig. 4 performed for a wave 2 forcing to answer the question of whether there are systematic differences in the response to the wave 1 and wave 2 forcings. Again the mountain is grown with a value of $\tau = 92.6$ days to several intermediate heights, remaining at that height for the rest of the model run. The lines at the top of the plots show the times when the various mountains reach their maximum heights.

When the mountain is grown to the full height of 2000 m the removal of mass begins at around 50 model days, similar to that of the wave 1 forcing above. The amount of mass initially moved into the vortex in the right panel is much larger than for wave 1 forcing. The stripping rate is also larger on the inner contour than for the wave 1 forcing, though it is similar for the outer contour.

When the mountain is grown to heights greater than 650 m the initial mass loss is similar to that of the 2000 m mountain. Smaller forcing amplitudes start losing mass at later times and at slower rates. For forcings of 400 m and below there is negligible mass loss for the entirety of the experiments, in contrast to the wave 1 case where only the 200 m mountain produces negligible mass loss. Of those forcings which produce significant mass loss those of lower amplitudes generally have shorter periods of rapid mass loss.

In the case of the wave 1 forcing the SSWs produced showed many of the features of real SSWs. For wave 2 forcings there are some forcings which show features similar to real SSWs but there are also some which produce behaviour unlike SSWs seen in the real stratosphere. The vortex starts being stretched by the forcing but is slowly eroded instead of splitting and

remains centered on the pole. The first 100 days of the simulations for which the vortex splits are highly symmetrical as the vortex gets stretched and eventually breaks into two symmetrical pieces. The evolution of these two sub-vortices continues to be similar to each other until after 100 days when one of these smaller vortices breaks down while the other remains intact.

For both the wave 1 and wave 2 forcing there is not a linear relationship between the maximum height of the mountain and the total amount of mass removed from the vortex. For both wave 1 and wave 2 there are several cases where smaller forcing produces more mass loss than larger forcing. This is due to the edge of the vortex becoming strengthened for some of the forcings, preventing further significant stripping, while this does not happen for others.

5. Discussion and conclusions

A shallow water model has been used to investigate stripping of filaments of air from a model stratospheric polar vortex in relation to the amplitude and rate of growth of the topographic forcing. The model SSWs produced here exhibit many of the features which can be seen in real stratospheric sudden warmings.

The stripping of mass from the vortex for wave 1 forcing is initiated at a point some time after the maximum forcing amplitude has been reached, particularly for lower mountain heights. This indicates that the forcing does not directly initiate stripping of mass from the vortex. A plausible hypothesis is that it takes longer for enough wave activity to build up to initiate stripping.

Wave activity $\mathcal{A}$ is related to the amplitude of the disturbance in the flow as well as its phase relative to the forcing. The point where the vortex starts to lose mass corresponds almost exactly to the point where the integrated wave activity reaches its maximum value, corresponding to wave breaking and mixing of filaments.

The total amount of mass removed from the vortex over the course of the experiment is not proportional to the maximum amplitude of the forcing. There are certain cases, most prominently in the case of the wave 1 forcing but also present to a lesser degree for the wave 2 case, when the vortex edge appears to sharpen after a certain amount of mass has been removed from it. This then has the effect of inhibiting any further significant tongues of air from forming and being stripped. This does not happen for all cases and it is not quite clear what may cause this to happen in some cases but not in others. This would be an area for further study. It is a reasonable hypothesis that the non-monotonic dependence of mass stripped on the maximum mountain height $H_{\max}$ is due to the chaotic nature of wave breaking; we have no other plausible explanation.

Matthewman and Esler (2011) looked at the effect of topographic forcing height and significance of the background flow on the onset of vortex splitting. Comparing their results to those obtained here suggests a possible explanation for the behavior seen here. In the results of Matthewman and Esler (2011) once the topography reaches a certain height the system undergoes a bifurcation and enters either a splitting or a high wave activity regime, with an area of oscillating vortex activity on the boundary of this region (see their fig.2). This is consistent with the results seen here where once the mountain reaches 450 m stripping starts to occur. For 500 m and above the vortex undergoes SSW like behaviour. For some values of the forcing amplitude the vortex looks similar to the high wave activity regime. This suggests that the system may be on the border of the two regimes (splitting and high wave activity) from fig.2a of Matthewman and Esler (2011). This would also account for the fact that some larger forcings do not induce SSWs while smaller ones do.

In contrast, Esler and Matthewman (2011) found that in the case of displacement SSWs the dynamics are much more complex than they are for the case of vortex splitting events. The more varied results for the wave 1 experiments performed here appear to be in agreement with the conclusion of Esler and Matthewman (2011) in this respect. The theory of Esler and Matthewman (2011) does not explain the fact that some smaller forcings can produce SSWs while intermediate ones do not.

In both the wave 1 and wave 2 cases there seems to be a certain amplitude of forcing below which there is no significant stripping of air initiated. This threshold forcing is larger for the wave 2 forcing (around 450 m) than for the wave 1 case (where all mountains larger than 200 m initiate significant amounts of stripping). The difference in mass removed between the negligible case and the smallest where stripping is initiated is also much larger in the wave 2 case than the wave 1 case where once stripping starts it appears to carry on for far longer.

Using the criterion of Eq. (12) to identify when SSW-like events have occurred shows that they can be produced for fairly small forcing amplitudes. In the case of the wave 1 forcing SSWs are produced for forcings over 500 m. In these the vortex looks similar to that of the real northern hemisphere polar vortex during a displacement event. In the case of the wave 2 forcing not all of the warmings produced resembled those seen in the stratospheric vortex. For smaller forcings the vortex remains centred exactly over the pole while Eq. (12) defines an SSW to have occurred. For larger forcings the vortex can be seen to split into two parts and resemble more closely the split vortex composite of Seviour et al. (2013).

To answer the questions posed in the introduction:

- **Is there a threshold in the forcing amplitude to initiate stripping?** In the experiments studied here 200 m is the threshold to initiate stripping of mass from the vortex for wave 1 forcing and 450 m for wave 2.

- **Is there a quantitative relationship between the forcing amplitude and the amount of stripping?** There does not appear to be any simple, direct relation between the amount of air stripped and the mountain height. However, the amount of stripping is a function of the integrated wave activity, which is itself related to the accumulated strength of the forcing.
- **When the forcing increases at a slow enough rate does the vortex go through a series of quasi steady states?** The vortex does not experience a series of quasi-steady states on its way to being eroded. Rather, the forcing appears to initiate a process that results in stripping several tens of days later.
- **What are the criteria for the vortex to break down; is there a threshold forcing amplitude above which the vortex breaks down?** The criteria for the vortex to break down are still unknown. It does not appear to be related in any simple way to the strength of the forcing or the amount of mass stripped from the vortex.
- **Is there a systematic difference between the response of the vortex to wave 1 and wave 2 forcings?** There is a systematic difference in the forcing amplitude needed to initiate the stripping (first bullet). The qualitative conclusions of the second, third and fourth bullets hold for both wave 1 and wave 2 forcing.

The jet shape and location in the initial velocity field have been kept fixed throughout the experiments, as well as the location and shape of the mountain. It is likely that these factors play a role in the rate and total amount of stripping experienced by the vortex. The values adopted in this study were chosen to be similar to what can be seen in the real winter stratosphere, but further experiments looking at these parameters would allow determination of the relative importance of these factors in relation to the amplitude or rate of growth of the forcing.

Acknowledgement

Robin Beaumont was supported during this research with a PhD studentship funded by an EPSRC Doctoral Training Grant.

Appendix A. Initial conditions

The initial conditions are set by starting with a zonally symmetric velocity field $\overline{u}(\phi)$ similar to that which would be found in the winter stratosphere. This is then integrated using the gradient wind balance

$$f\overline{u} + \frac{\overline{u}^2}{a}\tan\phi = -\frac{1}{a}\frac{\partial\overline{\Phi}}{\partial\phi} \tag{A.1}$$

to find the geopotential $\overline{\Phi}$.

The specified form of the velocity profile $\overline{u}(\phi)$ is

$$\begin{aligned}\overline{u}(\phi) \;=\; & 4.5\sin\left(\frac{20\phi-3\pi}{7}\right)\exp\left(\frac{11\phi+5}{10}\right) \;-4.5\sin\left(-\frac{13\pi}{7}\right)\exp\left(\frac{-11\pi+10}{20}\right)\exp\left[-\frac{(\phi+\frac{\pi}{2})^2}{0.015}\right] \\ & +12\exp\left[-\frac{(\phi+0.02)^2}{0.02}\right] \;-6\exp\left[-\frac{(\phi-0.29)^2}{0.015}\right] \;-2\exp\left[-\frac{(\phi-0.8)^2}{0.01}\right] \;+2\exp\left[-\frac{(\phi-1)^2}{0.01}\right] \\ & -4\exp\left[-\frac{(\phi-1.09)^2}{0.01}\right] \;+\exp\left[-\frac{(\phi-1.185)^2}{0.01}\right].\end{aligned} \tag{A.2}$$

This profile is constrained to be equal to 0 at both the North and South poles for continuity. Fig. 2 shows several examples of observed stratospheric zonal-mean zonal wind fields along with the initial velocity profile of Eq. (A.2).

References

Baldwin, M.P., Dunkerton, T.J., 1999. Propagation of the Arctic Oscillation from the stratosphere to the troposphere. J. Geophys. Res.: Atmos. 104, 30937–30946.
Brunet, G., Haynes, P.H., 1996. Low-latitude reflection of Rossby wave trains. J. Atmos. Sci. 53, 482–496.
Esler, J.G., Matthewman, N.J., 2011. Stratospheric sudden warmings as self-tuning resonances. Part II: Vortex displacement events. J. Atmos. Sci. 68, 2505–2523.
Juckes, M.N., McIntyre, M.E., 1987. A high-resolution one-layer model of breaking planetary waves in the stratosphere. Nature 328, 590–596.
Labitzke, K., 1981. The amplification of height wave 1 in January 1979: a characteristic precondition for the major warming in February. Mon. Weather Rev. 109, 983–989.
Legras, B., Dritschel, D., 1993. Vortex stripping and the generation of high vorticity gradients in two-dimensional flows. Adv. Turbul. IV, 445–455, Springer.
Legras, B., Dritschel, D.G., Caillol, P., 2001. The erosion of a distributed two-dimensional vortex in a background straining flow. J. Fluid Mech. 441, 369–398.
Limpasuvan, V., Thompson, D.W.J., Hartman, D.L., 2004. The life cycle of the Northern Hemisphere sudden stratospheric warmings. J. Clim. 17, 2584–2596.
Mariotti, A., Legras, B., Dritschel, D.G., 1994. Vortex stripping and the erosion of coherent structures in two-dimensional flows. Phys. Fluids 6, 3954–3962.
Matthewman, N.J., Esler, J.G., 2011. Stratospheric sudden warmings as self-tuning resonances. Part I: Vortex splitting events. J. Atmos. Sci. 68, 2481–2504.
McIntyre, M.E., Palmer, T.N., 1984. The 'surf zone' in the stratosphere. J. Atmos. Terres. Phys. 46, 825–849.
Norton, W.A., 1994. Breaking Rossby waves in a model stratosphere diagnosed by a vortex-following coordinate system and a technique for advecting material contours. J. Atmos. Sci. 51, 654–675.
Polvani, L.M., Saravanan, R., 2000. The three-dimensional structure of breaking Rossby waves in the polar wintertime stratosphere. J. Atmos. Sci. 57, 3663–3685.

Seviour, W.J.M., Mitchell, D.M., Gray, L.J., 2013. A practical method to identify displaced and split stratospheric polar vortex events. Geophys. Res. Lett. 40, 5268–5273.
Shepherd, T.G., 2007. Transport in the middle atmosphere. J. Meteorol. Soc. Jpn. 85, 165–191.
Sobel, A.H., Plumb, R.A., 1999. Quantitative diagnostics of mixing in a shallow water model of the stratosphere. J. Atmos. Sci. 56, 2811–2829.
Thuburn, J., Lagneau, V., 1999. Eulerian mean, contour integral, and finite-amplitude wave activity diagnostics applied to a single-layer model of the winter stratosphere. J. Atmos. Sci. 56, 689–710.
Thuburn, J., Cotter, C.J., 2015. A primal-dual mimetic finite-element scheme for the rotating shallow water equations on polygonal spherical meshes. J. Comput. Phys. 290, 274–297.
Waugh, D.W., et al., 1994. Transport out of the lower stratospheric Arctic vortex by Rossby wave breaking. J. Geophys. Res.: Atmos. 99, 1071–1088.