The Stability of Mars’s Annular Polar Vortex

WILLIAM J. M. SEVIOUR AND DARRYN W. WAUGH
Department of Earth and Planetary Sciences, The Johns Hopkins University, Baltimore, Maryland

RICHARD K. SCOTT
School of Mathematics and Statistics, University of St Andrews, St Andrews, United Kingdom

(Manuscript received 17 October 2016, in final form 14 February 2017)

ABSTRACT

The Martian polar atmosphere is known to have a persistent local minimum in potential vorticity (PV) near the winter pole, with a region of high PV encircling it. This finding is surprising, since an isolated band of PV is barotropically unstable, a result going back to Rayleigh. Here the stability of a Mars-like annular vortex is investigated using numerical integrations of the rotating shallow-water equations. The mode of instability and its growth rate is shown to depend upon the latitude and width of the annulus. By introducing thermal relaxation toward an annular equilibrium profile with a time scale similar to that of the instability, a persistent annular vortex with similar characteristics as that observed in the Martian atmosphere can be simulated. This time scale, typically 0.5–2 sols, is similar to radiative relaxation time scales for Mars’s polar atmosphere. The persistence of an annular vortex is also shown to be robust to topographic forcing, as long as it is below a certain amplitude. It is therefore proposed that the persistence of this barotropically unstable annular vortex is permitted owing to the combination of short radiative relaxation time scales and relatively weak topographic forcing in the Martian polar atmosphere.

1. Introduction

In common with several other planetary bodies, Mars’s atmosphere exhibits regions of strong circum-polar zonal winds known as polar vortices. These form in both the Northern and Southern Hemispheres during their respective winters and are an important barrier for the mixing of polar and midlatitude air. As such, they act to intensify meridional temperature gradients, influence the rate of condensation of CO2 onto the polar ice cap, and limit the transport of dust and ice aerosols toward polar regions (Colaprete et al. 2008; Kuroda et al. 2013; Guzewich et al. 2016).

The most useful quantity for studying the dynamics of polar vortices is potential vorticity (PV) (e.g., Hoskins et al. 1985), which acts as a conserved tracer in the absence of diabatic effects. In Earth’s stratosphere, PV generally increases monotonically from the equator to the winter pole, with strong meridional gradients near 60° latitude at the edge of the polar vortex and relatively weak gradients in midlatitudes and within the vortex (e.g., Andrews et al. 1987). In contrast, the Martian polar vortices consist of an annulus of high PV with opposing meridional gradients on the equatorward and poleward sides and a local minimum at the geographic pole (e.g., Barnes 1984; Banfield et al. 2004).

Although this comparison of Earth’s stratospheric polar vortex with the Martian polar vortex is intuitive because of their similar spatial extent (Mitchell et al. 2015), it is important to note that they are dynamically quite distinct. Mars’s atmosphere does not have a stratosphere and its polar vortex is strongly associated with the global-scale Hadley circulation, forming on the poleward side of the descending branch. Changes in this Hadley circulation, for instance driven by an increase in atmospheric dust loading, can therefore significantly affect the strength and structure of the polar vortex (Wilson 1997; Guzewich et al. 2016). In contrast, Earth’s stratospheric polar vortex is not closely connected with the Hadley cell. A more direct comparison can perhaps be made between Mars’s polar vortex and Earth’s subtropical jet (Waugh et al. 2016), which also forms on the poleward edge of the Hadley cell. However, Earth’s subtropical jet is at much lower latitudes than Mars’s polar vortex, and it has a monotonic increase in PV from...
the jet to the pole, unlike the annulus with a local polar minimum found on Mars.

The annular structure of the Martian polar vortex was first inferred from Mariner 9 and Viking spacecraft observations (Barnes 1984), though its persistence was unclear given the sparse spatial and temporal sampling of measurements. Observations of Martian atmospheric temperature, pressure, and dust have now been collected over several Martian years (Smith 2008; Kleinböhl et al. 2009). Although there have been no direct measurements of winds in the free Martian atmosphere, these observations have been combined with Martian general circulation models (MGCMs) (Lewis 2003) to produce reanalysis datasets (Greybush et al. 2012; Montabone et al. 2014). These best estimates of the Martian atmospheric state also display annular PV structures in both hemispheres from the surface up to about 60-km altitude, which persist from autumn until spring (Mitchell et al. 2015; Waugh et al. 2016). Figure 1 shows the structure of the Martian polar vortex near northern winter solstice in the Mars Analysis Correction Data Assimilation (MACDA) reanalysis (Montabone et al. 2014). Over a 30-solar-day (sol) average (Fig. 1a), a coherent elliptical annulus of PV is seen, with the zonal wind having a maximum (shown by the dashed white lines) equatorward of the maximum gradient in PV, unlike Earth’s stratospheric polar vortex. However, when viewed as an instantaneous snapshot, the largest values of PV do not appear as a smooth ring but, rather, as a series of smaller patches of high PV surrounding the pole (Fig. 1b). Similar annular features are seen for Mars’s Southern Hemisphere polar vortex during southern winter, although their magnitude is generally weaker than the northern vortex (Mitchell et al. 2015).

In addition to these reanalysis results, annular vortices have also been found in MGCM simulations (Barnes and Haberle 1996; Waugh et al. 2016), as well as from direct inference of PV from temperature observations without the use of a MGCM (Banfield et al. 2004; McConnochie 2007).

The persistence of this annular polar vortex is surprising since it is known that a planar strip of uniform vorticity is barotropically unstable (Rayleigh 1880). This result was extended to two-dimensional flow on a sphere through a linear stability analysis by Dritschel and Polvani (1992), who showed that sufficiently thin and equatorward annuli are linearly unstable, though no linear instability was found for wide annuli or annuli with a maximum near the pole. Certain annular vortex structures were also found to be barotropically unstable in shallow-water (Bowman and Chen 1994) and primitive equation models (Ishioka and Yoden 1994) of
Earth’s stratosphere. Barotropic instability of an idealized Mars-like zonal jet (motivated by Mariner 9 observations) was studied by Michelangeli et al. (1987) using a barotropic model. They found the most unstable mode to have zonal wavenumber 2, with a growth rate of approximately 2 sols. However, no study has yet looked at the stability of the annular PV distribution specific to Mars.

As a result of this apparent instability, it has been proposed that an external forcing mechanism must be present in order to maintain a persistent annular vortex. Mitchell et al. (2015) speculated that dynamical heating due to the descending branch of the Hadley circulation over the winter pole, which is of global scale and much stronger on Mars than Earth, may be responsible. However, Toigo et al. (2017) have more recently shown that latent heat release from CO₂ condensation over the polar cap acts as a sink of PV and is necessary to simulate an annular vortex in a MGCM. When this mechanism was turned off in their model, they found PV to rapidly become monotonic. An additional factor, not part of the investigation by Toigo et al. (2017), is that the ability for CO₂ condensation to act as a sink of PV depends crucially on the rate at which the atmosphere radiatively cools following an input of latent heat; a shorter time scale will lead to more rapid condensation, greater heat input, and a greater sink of PV than a longer time scale.

The aim of this study is to develop an improved understanding of the stability of Mars’s annular polar vortex by placing it into wider context: what are the necessary conditions for a persistent annular vortex to exist? Using a simple numerical model that is able to capture the dynamics of barotropic instability, we first investigate how the stability of annular vortices depends upon their geometry. We go on to impose a forcing with a sink of PV near the pole and determine the necessary strength of forcing, analogous to the radiative relaxation time scale, required to maintain an annular vortex. We also investigate the dynamical mechanism that may lead to the “patchiness” of PV (Fig. 1b) and study the influence of topographic forcing on our results. Finally, we show where observed values for the Martian atmosphere fit into this broader picture.

2. Model and experimental design

a. Shallow-water model

We investigate integrations of the rotating single-layer shallow-water equations on a sphere. This system, which has been widely used in studies of Earth’s stratospheric polar vortex (e.g., Juckes 1989; Norton 1994; Rong and Waugh 2004; Liu and Scott 2015; Scott 2016), allows us to isolate the dynamics associated with barotropic instability from those due to vertically propagating waves. In vorticity–divergence form, these equations are

\[
\frac{\partial \zeta}{\partial t} = -\nabla \cdot (\mathbf{u} \zeta), \tag{1a}
\]

\[
\frac{\partial \delta}{\partial t} = -\frac{1}{2} \nabla^2 |\mathbf{u}|^2 + k \cdot \nabla \times (\mathbf{u} \zeta) - g \nabla^2 (h + h_b), \tag{1b}
\]

\[
\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}) = -(h - h_e)/t_r. \tag{1c}
\]

Here \( \zeta \) is the relative vorticity, \( \delta \) the divergence, and \( h \) the layer thickness. The absolute vorticity is \( \zeta_a = 2\Omega \sin \phi + \zeta \), where \( \Omega = 7.09 \times 10^{-5} \text{s}^{-1} \) is the planetary rotation rate, \( \phi \) is latitude, \( g = 3.71 \text{m s}^{-2} \) is the gravitational acceleration, and \( \mathbf{u} = (u, v) \) is the horizontal velocity (values of planetary constants appropriate for Mars have been chosen). The system is forced by topography, \( h_b \), and relaxation toward an equilibrium height profile, \( h_e \), with time scale \( t_r \). In the shallow-water system, PV is given by \( q = \zeta_a h \).

We use a mean layer depth of \( H = 17 \text{km} \), which is chosen to give a Rossby deformation length, \( L_D = \sqrt{gH/2\Omega} \approx 1800 \text{km} \), equal to the Rossby deformation length of the Lamb mode, \( ND/\sqrt{\kappa(1 - \kappa)} \), which is the relevant vertical structure of the barotropic mode in a compressible atmosphere (Liu and Scott 2015). Here \( \kappa = \frac{1}{4} \) (the appropriate value for a triatomic gas such as CO₂, which is the major constituent of the Martian atmosphere), \( f = 2\Omega \sin \phi \), \( N = 0.01 \text{s}^{-1} \) is the Brunt–Väisälä frequency, and \( D = 11 \text{km} \) is the density scale height (approximate values taken from http://nssdc.gsfc.nasa.gov/planetary/factsheet/marsfact.html).

The equations are solved in a spherical coordinate system using the “Built on Beowulf” (BOB) pseudospectral model in which the model fields are represented by a basis of spherical harmonic functions and the nonlinear terms are computed on a physical space grid (Rivier et al. 2002; Scott et al. 2004). Time stepping is by a semi-implicit leapfrog scheme. For numerical stability, a weak, scale-selective dissipation is included on the right-hand side of Eqs. (1a)–(1c) to prevent the buildup of enstrophy at small scales; here we use a dissipation of the form \( \Psi^8 \) with a damping rate of approximately \( 10 \text{sol}^{-1} \) at the smallest resolved wavenumber. In the experiments reported below, we use a triangular spectral truncation with maximum total wavenumber 170 (T170). The sensitivity of these experiments to this choice of resolution is discussed in the appendix.

b. Initial conditions

We initialize our model with a zonal-mean annular PV profile, as below:
This profile (shown in Fig. 2a) has a local minimum of PV, \( q_0 \), near the northern (winter) pole. Surrounding this is an annulus of higher PV, \( q_y \), followed by a drop to \( q_s \) from the vortex edge to the equator. While the Martian atmosphere has near-zero PV equatorward of the polar vortex region because of an approximately angular momentum-conserving Hadley circulation (Barnes and Haberle 1996), we find that setting \( q_s = 0 \) leads to unrealistically large easterlies at low latitudes in the shallow-water model. The precise choice of values of \( q_p \) and \( q_s \) is quite difficult to motivate physically, but the results presented below are not highly sensitive to this choice, and those selected here give, for appropriate values of \( \phi_1 \) and \( \phi_2 \), a similar climatological vortex as that observed on Mars. To satisfy the requirement that the integral of the relative vorticity over the sphere is zero, a constant is added to the initial PV distribution, which does not affect the dynamics of the subsequent evolution. To allow for easier comparison between different PV distributions, this constant offset is not included in the figures below. From the equator to the southern (summer) pole, PV is set equal to the planetary vorticity. To avoid unwanted Gibbs fringes, we apply a smoothing so that PV varies linearly over \( 3^\circ \) latitude across both the poleward and equatorward boundaries of the annulus.

Given this definition of \( q_0 \), equilibrium values of \( h \) and \( u \) can be derived from Eq. (1b) by assuming nondivergence (\( \delta = 0 \)), and no topography (\( h_b = 0 \)), using the iterative method described in Liu and Scott (2015). These equilibrium and initial conditions are shown in Figs. 2b and 2c. To allow for the growth of instabilities, a small, unbalanced perturbation is applied to the absolute vorticity profile. This is a random number between \( \pm 1.5 \times 10^{-3} \bar{\Omega} \) applied at each grid point. The perturbation is therefore uncorrelated in space so as not to favor the growth of any particular mode.

c. Thermal relaxation

The Martian atmosphere responds to a warm or cold temperature anomaly by increased emission or absorption of longwave radiation, respectively. The right-hand side of Eq. (1c) provides a simple means of incorporating the effects of this radiative relaxation in our simulations through a relaxation toward an equilibrium height profile \( h_e \) with a time scale \( t_r \). Here we take \( h_e \) to be equal to the height profile associated with the initial PV distribution, shown in Fig. 2b. Similar relaxation schemes have been applied in studies of Earth’s stratospheric polar vortex (Rong and Waugh 2004; Scott 2016), although toward a monotonic rather than annular equilibrium profile.

d. Topographic forcing

The system is forced by topography \( h_b \) with zonal wavenumber 2 and maximum amplitude at the equator, given by

\[
h_b(\lambda, \phi) = A_0 \cos^2(\phi) \cos(2\lambda),
\]

where \( \lambda \) is longitude and the amplitude is \( A_0 \). This is a rough approximation to realistic Martian zonal topography, which is also dominated by a zonal wavenumber 2, with a maximum near the equator (Hollingsworth and Barnes 1996). Note, however, that the topography imposed in our model should be thought as representing the effect of topographically forced waves, so its amplitude cannot be straightforwardly compared with that of real Martian topography.

Form drag from the westerly flow of the vortex over this topography provides a source of negative angular
momentum that has no systematic sink in Eqs. (1a)–(1c). To allow for a stationary state to be reached, we impose a drag on the zonal-mean zonal velocity \( \pi \). Following Scott (2016), we add a term of the form

\[
-\mathbf{k} \cdot \nabla \times [r(\phi)\pi, 0]
\]

(4)

to the right-hand side of Eq. (1a), where \( r(\phi) \) is a latitude-dependent damping rate, which takes the form

\[
r(\phi) = r_0 \left( 1 - \tanh \frac{\phi - 10^\circ}{10^\circ} \right).
\]

(5)

This is chosen so that the drag is confined to the tropics and Southern Hemisphere, away from the Northern (winter) Hemisphere of interest. A value of \( r_0 = 1/20 \) sols was found sufficient to reach a stationary state over a wide range of values for \( A_0 \).

3. Results

a. Barotropic instability

We begin by considering the stability of an annular vortex in the adiabatic case \( t_e \rightarrow \infty \) (i.e., no thermal relaxation) and without topographic forcing (\( A_0 = 0 \)). Figure 3 shows the evolution of the vortex from three different initial conditions: \([\phi_1, \phi_2] = [65^\circ N, 70^\circ N], [70^\circ N, 75^\circ N], \) and \([60^\circ N, 70^\circ N] \). In all three cases initial disturbances are seen to grow faster on the inner (poleward) edge of the annulus. At 16 sols following the initialization, waves on the inner and outer edges of the annulus have phase locked, leading to a polygonal structure in all three cases; a hexagon, pentagon, and square, respectively. The wave on the inside edge is the first to break, drawing filaments of the relatively low PV near the pole into the annulus. These filaments stretch and fold and dissipate when reaching sufficiently short length scales, so that the annular structure gradually degrades (see sol 28). Finally, at 60 sols all three cases have formed a vortex patch with a monotonic increase in PV from the equator to pole. We also find this behavior to be robust for a range of horizontal resolutions (see the appendix).

This evolution can be further illustrated by considering a Fourier decomposition of PV into zonal wavenumbers (Fig. 4). Over the initial growth of the instability (approximately the first 10 sols) the amplitude \( Z_m \) of the dominant zonal wavenumber \( m \) grows nearly exponentially as

\[
Z_m = Z_{m_0} e^{\nu t_{\text{inst}}},
\]

(6)

where \( Z_{m_0} \) is the initial amplitude and \( t_{\text{inst}} \) is the time scale for the growth of the instability. This exponential growth is shown by the fit of the dashed lines in Fig. 4 and is characteristic of a linear instability. The fit is from sol 0 to the point at which the amplitude becomes greater than 9, though the results presented below are not highly sensitive to this choice. Beyond about 15 sols from initialization, other wavenumbers grow and more complex nonlinear behavior is seen as the vortex wraps up from an annulus into a patch.

Following the three examples above, the linear instability over a wider range of annular vortex widths and latitudes is considered (Fig. 5a). These have been calculated from simulations taking 5° steps in \( \phi_1 \) and \( \phi_2 \). The two main results are that 1) the instability for thinner annuli has higher wavenumbers than thicker annuli and 2) annuli with a maximum nearer the pole have lower-wavenumber instabilities than those with a more equatorward maximum. Note that although Fig. 5a appears to show jumps between, say, wavenumber-6 instabilities and wavenumber 4 as the annulus width is increased, there is in fact a smooth transition with wavenumber 5 for intermediate widths. This is not seen because of the 5° steps we have taken between simulations. Furthermore, instabilities with wavenumbers greater than 6 are possible, but simulations with sufficiently thin annuli are not shown in Fig. 5.

Result 1 can be seen in analogy with Rayleigh (1880), who showed that the wavenumber of the most unstable mode of a planar strip of vorticity is inversely proportional to its width. Similar dependence of the most unstable wavenumber on annulus width has also been found in studies of two-dimensional planar flow (Dritschel 1986; Waugh and Dritschel 1991), as well as \( f \)-plane studies of the stability of annular hurricane eyewall structures (Rozoff et al. 2009; Hendricks et al. 2014). Result 2 is consistent with Dritschel and Polvani (1992), who investigated the stability of annular vortices in two-dimensional flow on the sphere. They found that annuli with a maximum near the pole tended to have lower wavenumbers of the most unstable mode—a result of the curvature of the sphere, which becomes more significant near the pole, limiting the ability of vorticity strips to “roll up.” The time scale for the growth of the most unstable mode \( t_{\text{inst}} \) is approximately between 1 and 3 sols for each simulation (Fig. 5b), consistent with the time scale of Michelangeli et al. (1987). There is a tendency for thinner annuli and annuli with a maximum nearer the equator to have shorter time scales than those that are thicker or have a maximum nearer the pole, which is also consistent with Dritschel and Polvani (1992).

The gray regions in Fig. 5a indicate simulations for which no linear instability was found. Specifically, this is defined as where no wavenumber grows beyond an
amplitude of 9 within the first 40 sols of the simulation. These simulations show qualitatively different evolution to those with a linear instability, and an example is shown in Fig. 6. Although a wavenumber-3 structure develops on the inner edge of the annulus, this does not significantly affect the outer edge, and by 55 sols the PV distribution is not much changed from the initial state. Eventually, a wavenumber-1 mode grows and by
100 sols the PV distribution becomes monotonic. This is typical of these simulations without linear instability, none of which maintain an annular vortex beyond 100 sols. A decomposition of the evolution into zonal wavenumbers (as in Fig. 4) does not show the exponential growth of any mode (not shown), again indicating the lack of linear instability. These findings are again consistent with Dritschel and Polvani (1992), who also found that sufficiently wide annuli or annuli with a maximum near the pole do not show linear instabilities in cases where the vorticity inside and outside the annulus is unequal.

Overall we have demonstrated, through these unforced simulations of the barotropic instability of annular vortices, that the shallow-water system investigated here produces qualitatively similar results to analogous studies of two-dimensional or planar flow. It is instructive to consider where the real Martian atmosphere may fit into these broader results. The annular vortex observed in Mars’s atmosphere is relatively wide with a PV maximum near 75°N (Fig. 1; Waugh et al. 2016); hence, we propose that the high-wavenumber, highly unstable modes found in this analysis are unlikely to apply on Mars. It is more likely that either a low-wavenumber mode or a distribution which is not linearly unstable is more relevant. These results provide strong grounds to go on to investigate the role of thermal relaxation in maintaining an annular vortex, which is the topic of the next section.

b. Effect of thermal relaxation

PV evolution under three different relaxation time scales \( t_r = 10, 2, \) and 0.5 sols is shown in Fig. 7 for the equilibrium profile of \( \phi_1, \phi_2 = [60°N, 70°N] \) (chosen for comparison with Fig. 3). The unforced evolution of this profile was shown in the right-hand column of Fig. 3 and found to have a wavenumber-4 instability with time scale \( t_{\text{inst}} = 2.5 \) sols (Fig. 5). These three relaxation time scales shown here can therefore be thought to represent the cases \( t_r > t_{\text{inst}}, t_r \approx t_{\text{inst}}, \) and \( t_r < t_{\text{inst}}. \) In the case of the weakest relaxation, \( t_r = 10 \) sols, the initial growth of the linear instability (at sol 16) is very similar to the unforced case (Fig. 3). By 60 sols, the initial annulus has collapsed to form a monotonic distribution which then persists, so that the long-term average is also monotonic (bottom row of Fig. 7). The overall strength of the vortex is, however, weaker than the unforced case as a result of the relaxation causing a net PV sink poleward of 70°N.

Under strong relaxation, \( t_r = 0.5 \) sols, the evolution is very different. At 16 sols following initialization the wavenumber-4 instability is present, but much suppressed. The relaxation prevents the barotropic instability from developing, and an annular vortex is maintained with relatively small zonal asymmetries. The long-term average vortex is similar to the equilibrium profile, although its maximum is slightly shifted poleward, and with weaker meridional gradients. Like the strong relaxation case, intermediate relaxation, \( t_r = 2 \) sols \( \approx t_{\text{inst}}, \) prevents the instability from growing and maintains an annular vortex. There are, however, greater zonal asymmetries seen in the instantaneous
snapshots owing to the closer competition between the instability and relaxation. These asymmetries somewhat resemble the patchy PV structure seen in the reanalysis data (Fig. 1b), although the simulated annulus is slightly more coherent than that of the reanalysis. Much of the differences are due to the lower resolution of the reanalysis, which is unable to represent the development of features at the smaller scales of the simulation. In the long-term average, the intermediate relaxation case shows an annular vortex but its maximum is shifted poleward and is weaker than the strong relaxation case.

In simulations where the relaxation is sufficiently strong to maintain an annular vortex, we find the strongest meridional PV gradients to be on the outside

FIG. 5. (a) Dominant wavenumber of the linear instability for a range of annular vortex positions and widths. Regions labeled “n” represent no linear instability at any wavenumber. (b) Time scale $t_{\text{inst}}$ for the growth of the most unstable mode. In these panels, moving from the lower left toward the upper right corresponds to an annulus of constant width but with a maximum moving toward the pole, while moving from the upper left toward the lower right corresponds to a wider annulus but with a fixed position of its maximum.

FIG. 6. Polar stereographic projection of PV evolution, with no thermal relaxation and no topography, for the initial condition $[\phi_1, \phi_2] = [65^\circ N, 80^\circ N]$. Color scale and latitudinal range as in Fig. 3.
of the annulus, with weaker gradients in the interior (illustrated in Fig. 8 for the case $t_r = 0.5$ sols). This is the reverse of the result seen in reanalysis data, where the strongest gradients are found near the pole (Mitchell et al. 2015; Waugh et al. 2016). Also seen in reanalysis data is the result that the maximum meridional PV gradient is some way poleward of the maximum zonal winds (Fig. 1; Waugh et al. 2016), unlike on Earth where the two are approximately collocated. In our simulations, like the Earth, but unlike Martian reanalysis, we also find the maximum meridional PV gradient and zonal wind to be collocated (Fig. 8).

The above caveats aside, we find it possible to simulate persistent annular vortices for a range of different...
equilibrium profiles (Fig. 9). The equilibrium profiles $[\phi_1, \phi_2] = [65^\circ \text{N}, 70^\circ \text{N}]$ and $[70^\circ \text{N}, 75^\circ \text{N}]$ (Figs. 9a,b) have $t_{\text{inst}} = 1.37$ and 1.42 sols, respectively. In these cases, $t_r = 0.5$ sol produces an annular mean vortex, but weaker relaxation ($t_r \geq 1$ sol) leads to an approximately monotonic PV profile. For the case $[\phi_1, \phi_2] = [65^\circ \text{N}, 70^\circ \text{N}]$, as shown in Fig. 7, for which the instability time scale is longer $t_{\text{inst}} = 2.5$ sols, annular mean vortices are seen for $t_r \leq 2$ sols (Fig. 9c). In general, therefore, we find that an annular vortex can be maintained provided $t_r \leq t_{\text{inst}}$. Interestingly, these results do not change much when considering an equilibrium profile that is not linearly unstable (Fig. 9d shows the profile as in Fig. 6). Here an annular mean vortex is seen for $t_r \leq 2$ sols but monotonic profiles for weaker relaxation values. The same is true for other equilibrium profiles that are not linearly unstable, all of which require relaxation time scales less than about 3 sols to maintain an annular vortex beyond 100 sols from initialization.

c. Effect of topography

It is known that topographically driven waves are present in the Martian atmosphere, so it is important to consider if the above results hold in the presence of this forcing. Figure 10 shows the initial evolution and climatology for the equilibrium profile $[\phi_1, \phi_2] = [65^\circ \text{N}, 70^\circ \text{N}]$ with $t_r = 0.5$ sol, under three different amplitudes of topographic forcing. In the weakest case, $A_0 = 0.0125H$, an initial zonal wavenumber-6 instability is seen, as in the unforced case (Fig. 3, left column). The relaxation prevents collapse to a monotonic PV distribution, and the long-term average is an approximately circular annulus. For the intermediate case, $A_0 = 0.075H$, the initial instability is less clear, although relaxation maintains an annular vortex. Unlike the weaker topography case, this climatological vortex is an elliptical annulus, remarkably similar to the climatological vortex structure seen in reanalysis data (Fig. 1; Mitchell et al. 2015; Waugh et al. 2016). Under the strongest topographic forcing, $A_0 = 0.125H$, the initial annulus breaks up with a dominant zonal wavenumber-2 pattern. Unlike the previous two cases, the topographic forcing here overwhelms the relaxation, and the climatological vortex is elliptical and monotonic.

This result, that topographic forcing beyond a certain amplitude leads to a monotonic mean PV distribution, also holds for a range of other equilibrium profiles. While topography amplitudes in our shallow-water model cannot be compared directly to real Martian topography, these results suggest that Martian topographic forcing is most analogous to the case $A_0 = 0.075H$ in showing an elliptical yet annular climatological vortex.

4. Discussion and conclusions

We have shown that a range of annular vortices are linearly unstable with instabilities growing with time scales, $t_{\text{inst}} \approx 1–3$ sols. These annular vortices can be maintained with relaxation, time scale $t_r$, toward an annular equilibrium profile, so long as $t_r \leq t_{\text{inst}}$ (and the topographic forcing is below a certain amplitude). How does this relate to radiative relaxation time scales in the real Martian atmosphere? For a simple gray atmosphere, the radiative relaxation time scale $t_{\text{rad}}$ can be estimated (e.g., James 1994) as

$$t_{\text{rad}} \approx \frac{c_p p}{4(2 - \varepsilon) g \alpha T^3}, \quad (7)$$

where $c_p$ is the constant-pressure heat capacity, $p$ is the pressure, $\varepsilon$ is the infrared opacity, $T$ is the temperature, and $\alpha$ is the Stefan–Boltzmann constant. Typical values for polar winter at approximately 1 hPa are $c_p = 860 \text{ J kg}^{-1} \text{K}^{-1}$, $T = 150 \text{ K}$, and $\varepsilon = 0.2$, giving a time scale $t_{\text{rad}} \approx 0.2$ sols. In reality, the radiative relaxation time scale is dependent upon vertical scale, and for typical scales for polar vortex variability (10–40 km), values of $t_{\text{rad}} \approx 0.5–2$ sols are more appropriate (Eckermann et al. 2011). These values are similar to or less than typical barotropic instability time scales $t_{\text{inst}}$ and therefore consistent with our understanding that $t_r \leq t_{\text{inst}}$ is a necessary condition in order for a persistent annular vortex to exist.

In the case that $t_r \approx t_{\text{inst}}$, we showed that although there was a persistent annular vortex, there were also significant zonal asymmetries owing to the close competition between the linear instability and the relaxation (Fig. 7). Given the relatively similar values of $t_{\text{rad}}$
and $t_{\text{inst}}$ described above, it is possible that a similar mechanism is responsible for the zonal asymmetries seen in reanalysis data (Waugh et al. 2016; Fig. 1b). Interestingly, this mechanism is entirely independent of vertically propagating waves forced from the surface.

The relaxation toward a specified height profile in our simulations can be viewed as the sum of a mass source, driving a positive flux of PV for $\phi > \phi_1$, and a mass sink, driving a negative flux of PV for $\phi > \phi_2$. A limitation of our study is that these source and sink regions represent the combined effect of several distinct physical processes which can affect PV, including diabatic descent in the Hadley circulation as well as latent heat release from CO$_2$ condensation. Since these processes are combined into a single term, it is not possible to vary them independently. It would, however, be possible to construct a system in which this is can be achieved using the shallow-water Hadley cell model of Held and Phillips (1990) or through extension to a three-dimensional system. This would allow investigation into whether an annular vortex can be generated from a Hadley cell alone [as proposed by Mitchell et al. (2015)] or if latent heating is necessary.

A further significant limitation of this study is that several dynamical processes that cannot be captured by our simple shallow-water model are also likely to play an important role in the dynamics of Mars’s polar vortex. These include baroclinic eddies, forced stationary eddies, and thermal tides, all of which may produce strong meridional mixing of PV. The analysis of the role of these phenomena in relation to the stability of the annular polar vortex can be carried out using MGCM simulations and would be an important future step.

Although the focus of this study has been Mars’s polar vortex, the principles developed here are easily translatable to other planetary bodies. Of particular interest is Titan, for which an annular polar structure has also been inferred in the stratosphere (Achterberg et al. 2011). As well as being an interesting dynamical feature, annular polar vortices may have significant consequences for a planet’s atmospheric chemistry and transport. On Earth, the confinement of air inside the stratospheric polar vortex is crucial for the chemical processes that lead to the Antarctic ozone hole. The relative effectiveness of annular vortices at isolating polar air is
an open question. Given the results above, it seems likely that the zonal asymmetries formed in cases when \( t_r \approx t_{\text{rot}} \) would allow for greater mixing of midlatitude and polar air, but this hypothesis remains to be tested.

**Acknowledgments.** The authors thank Anthony Toigo, Scott Guzewich, Lorenzo Polvani, and three anonymous reviewers for their valuable comments. The code necessary to reproduce the simulations and analysis described here, as well as the data from those simulations, may be accessed from Seviour et al. (2017). The MACDA data, used in Fig. 1, are publicly available from Montabone et al. (2011). This research was partially supported by a NASA grant.
APPENDIX

Sensitivity to Spatial Resolution

The simulations presented in this study have used a triangular spectral truncation with maximum total wavenumber 170 (T170), which corresponds to 512 × 256 longitude–latitude Gaussian grid points, or approximately 42 km near the equator. To illustrate the sensitivity of our results to this choice of resolution we here compare simulations run at different resolutions. Fig. A1 shows unforced simulations for the equilibrium profile \([\phi_1, \phi_2] = [60^\circ N, 70^\circ N]\) (as shown in Fig. 3), at four different spectral resolutions; T42, T85, T170, and T341, each representing a quadrupling in the total number of grid points, and corresponding to a spacing of 42 km near the equator.
approximately 2.8°, 1.4°, 0.7°, and 0.35° near the equator, respectively.

In common with the T170 simulation, both T85 and T341 simulations show wavenumber-4 instability. The instability appears faster in the T85 case, though the time scale $t_{\text{inst}}$ is actually similar in these three cases; the faster appearance of the instability is a result of the relatively larger initial perturbation at T85 because of its coarser grid. Progression from the linear to nonlinear evolution is slower at higher resolutions, and this may be understood in that thinner filaments of PV are permitted at higher resolution, which take longer to reach length scales at which they dissipate. While the T85, T170, and T341 cases have broadly similar evolution, the T42 simulation is strikingly different. This shows an initial wavenumber-3 pattern, followed by a wavenumber-2 mode, which take longer to reach length scales at which they dissipate. While the T85, T170, and T341 cases have broadly similar evolution, the T42 simulation is strikingly different. This shows an initial wavenumber-3 pattern, followed by a wavenumber-2 mode, which take longer to reach length scales at which they dissipate. While the T85, T170, and T341 cases have broadly similar evolution, the T42 simulation is strikingly different. This shows an initial wavenumber-3 pattern, followed by a wavenumber-2 mode, which take longer to reach length scales at which they dissipate.

We conclude that our results are broadly insensitive to an increase or moderate decrease in resolution, but at sufficiently coarse resolution, they may become qualitatively different. We find similar resolution dependence when considering forced simulations (not shown). It is somewhat surprising that the T42 simulations have such different behavior, given that they can easily resolve a wavenumber-4 mode. This may have some important implications for MGCM or reanalysis studies, which typically use resolutions closer to T42 than T170, but further investigation is needed into whether these results apply to more complex models.

REFERENCES


