

Zonal-mean dynamics of extended recoveries from stratospheric sudden warmings

Article

Published Version

Hitchcock, P. and Shepherd, T. G. ORCID:
<https://orcid.org/0000-0002-6631-9968> (2013) Zonal-mean dynamics of extended recoveries from stratospheric sudden warmings. *Journal of the Atmospheric Sciences*, 70 (2). pp. 688-707. ISSN 1520-0469 doi: <https://doi.org/10.1175/JAS-D-12-0111.1> Available at <https://centaur.reading.ac.uk/32107/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.1175/JAS-D-12-0111.1>

To link to this article DOI: <http://dx.doi.org/10.1175/JAS-D-12-0111.1>

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur



Central Archive at the University of Reading
Reading's research outputs online

Zonal-Mean Dynamics of Extended Recoveries from Stratospheric Sudden Warmings

PETER HITCHCOCK AND THEODORE G. SHEPHERD

Department of Physics, University of Toronto, Toronto, Ontario, Canada

(Manuscript received 7 April 2012, in final form 22 August 2012)

ABSTRACT

The recovery of the Arctic polar vortex following stratospheric sudden warmings is found to take upward of 3 months in a particular subset of cases, termed here polar-night jet oscillation (PJO) events. The anomalous zonal-mean circulation above the pole during this recovery is characterized by a persistently warm lower stratosphere, and above this a cold midstratosphere and anomalously high stratopause, which descends as the event unfolds. Composites of these events in the Canadian Middle Atmosphere Model show the persistence of the lower-stratospheric anomaly is a result of strongly suppressed wave driving and weak radiative cooling at these heights. The upper-stratospheric and lower-mesospheric anomalies are driven immediately following the warming by anomalous planetary-scale eddies, following which, anomalous parameterized nonorographic and orographic gravity waves play an important role. These details are found to be robust for PJO events (as opposed to sudden warmings in general) in that many details of individual PJO events match the composite mean.

A zonal-mean quasigeostrophic model on the sphere is shown to reproduce the response to the thermal and mechanical forcings produced during a PJO event. The former is well approximated by Newtonian cooling. The response can thus be considered as a transient approach to the steady-state, downward control limit. In this context, the time scale of the lower-stratospheric anomaly is determined by the transient, radiative response to the extended absence of wave driving. The extent to which the dynamics of the wave-driven descent of the stratopause can be considered analogous to the descending phases of the quasi-biennial oscillation (QBO) is also discussed.

1. Introduction

The Arctic polar vortex is one of the most variable components of the zonal-mean circulation of the atmosphere on intraseasonal to interannual time scales. The radiatively generated vortex is disrupted intermittently and irregularly by planetary-scale Rossby waves produced by the troposphere below. In the most spectacular cases, these bursts result in major stratospheric sudden warmings, during which the zonal-mean westerly winds reverse. For up to 3 months following roughly half of these warmings, however, Hitchcock et al. (2013, hereafter HSM) found that the recovery of the vortex follows a remarkably regular evolution, characterized by a persistently warm lower stratosphere and an elevated stratopause that descends steadily over the course of the recovery, separated by an extremely cold midstratosphere.

The behavior of these extended-time-scale recoveries from sudden warmings has been studied implicitly in other contexts. They contribute significantly, for instance, to the composites of sudden warmings presented by Charlton and Polvani (2007). However, HSM demonstrated that the phenomenology of those sudden warmings that are not followed by extended-time-scale recoveries differ significantly from those that are; most notably, planetary waves propagating into the vortex during the extended-time-scale recoveries are strongly suppressed.

The downward migration of temperature anomalies on monthly time scales (which are also seen in the zonal wind field) projects strongly onto the EOFs used by Kuroda and Kodera (2004) to define the polar-night jet oscillation (PJO). The PJO is conceived of as a slow mode of variability, active throughout the winter season but with varying amplitudes (Kuroda and Kodera 2001). The most coherent, largest-amplitude manifestations of this mode, however, occur during these extended-time-scale recoveries. To emphasize these events as a well-defined subset of sudden warmings, and to emphasize

Corresponding author address: Peter Hitchcock, Dept. of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB1 0WA, United Kingdom.
E-mail: aph42@cam.ac.uk

the connection with the monthly time scales and downward migration of the PJO, we follow HSM and refer to these extended-time-scale recoveries as PJO events. Note that in this perspective they are not truly oscillatory.

PJO events have also been of interest recently because of the elevated polar stratopause that accompanies them (Manney et al. 2008; Siskind et al. 2010; Limpasuvan et al. 2012; Tomikawa et al. 2012). The behavior of CMAM agrees well with these studies, confirming the robustness of the dynamics in the upper stratosphere across a number of comprehensive middle-atmosphere models.

The regularity of the zonal-mean evolution during PJO events suggests that the processes responsible are common to all the events, and are thus presumably predictable; indeed, enhanced predictability in the stratosphere has been found following sudden warmings in forecasting case studies (Mukougawa et al. 2009). Since these events have been shown to influence the tropospheric circulation (Baldwin and Dunkerton 2001), this has practical implications for seasonal predictability at the surface. It also implies that a detailed understanding of the zonal-mean dynamics of a small number of events could yield more insight on the general case than might otherwise be expected.

Moreover, the statistical characterization of these events in HSM suggests that the dynamics underlying the persistence of the lower-stratospheric anomaly are essentially zonal in character. For instance, counter to the implications of some studies (Yoden et al. 1999; Charlton and Polvani 2007), the longer recovery time scales apparent following vortex splits are simply a result of there having occurred more PJO events in the observations following vortex splits than have occurred following displacements. If one considers only those warmings that are followed by PJO events, there is no difference in the time scales of splits and displacements. Provided the initial disturbance reaches the lower stratosphere, the duration of the recovery is insensitive to this zonally asymmetric classification.

The purpose of this paper is to analyze in detail the zonal-mean evolution during PJO events, emphasizing in particular the factors that give rise to persistence of the lower-stratospheric anomalies, which are likely to be responsible for setting the time scale of the events as a whole. The dynamics of this evolution can be divided into two components: the behavior of the eddies given a certain zonal flow, and the response of the zonal-mean flow to those eddies (and any other zonal-mean forcing). We focus here primarily on the latter. One would ideally like to attribute any changes in the zonal-mean state to a particular zonal-mean forcing. This, however, is complicated by the process of Eliassen adjustment, by which

the meridional circulation induced by any imposed torque or heating acts to maintain thermal wind balance (Eliassen 1951). Under quasigeostrophic scaling assumptions appropriate to the extratropics, the meridional circulation is determined instantaneously by the structure of any such zonal-mean forcings (Plumb 1982); the full response to a given forcing is then the net effect of that forcing and the circulation it induces.

Since the diabatic heating in the middle atmosphere is strongly determined by local temperature anomalies [Rodgers and Walshaw 1966; see Hitchcock et al. (2010) for a more recent discussion], one can consider the latter to be caused, ultimately, by the torques alone. In this framework, the zonal response to the torque becomes time dependent, and relaxes toward the steady-state, “downward control” limit. The nature of this transient approach was thoroughly investigated by Haynes et al. (1991) in the case of constant radiative damping. Note that while the torques are thus causally prior to the diabatic heating, the details of the latter are, as will be seen below, essential for determining the structure and persistence of the lower-stratospheric anomalies.

We demonstrate here that, provided care is taken in specifying appropriate boundary conditions, radiative damping rates, and static stability, this framework can reproduce in some detail the zonal-mean dynamics of PJO events as produced by the Canadian Middle Atmosphere Model (CMAM), a comprehensive chemistry-climate model. Moreover, since this diagnostic framework is nearly linear, it permits a full decomposition of the response to eddy flux convergences of different types of waves.

The datasets analyzed here are summarized in section 2. The variability of the Arctic polar vortex in the CMAM integration is discussed in section 3, then in section 4 by a more thorough review of Eliassen adjustment in the framework of the transformed Eulerian mean (TEM). Composites of PJO events are provided in section 5, followed by a detailed decomposition of the zonal-mean evolution in two case studies in section 6 with a particular emphasis on the persistence of the lower-stratospheric anomaly. In section 7 the gravity wave–driven descent of the stratopause is considered explicitly. Discussion and conclusions are presented finally in section 8.

2. Data

The analysis in this study is primarily based on a 96-yr time-slice run of CMAM (Scinocca et al. 2008). The annular mode variability of this run has been analyzed by Simpson et al. (2011). The integration was performed at horizontal resolution of T63, with 72 vertical levels to 0.001 hPa. No interactive chemistry was used for this

run, and sea surface temperatures and sea ice were specified climatologically. Similar analyses have been performed on shorter time slice simulations that include fully interactive chemistry and an interactive ocean, the results of which resemble closely those presented here, suggesting that neither is necessary for obtaining PJO-like variability in the Arctic stratosphere. We note as well that the parameters of the orographic gravity wave drag parameterization have been adjusted from the CMAM simulations submitted to the second Chemistry–Climate Model Validation Activity (CCMVal) Intercomparison Project (Butchart et al. 2011) in order to recover the observed frequency and seasonality of sudden warmings. A similar adjustment was performed by Richter et al. (2010). As a result, these features match observations fairly well, as was found in the version analyzed by HSM.

Zonal-mean zonal winds and temperatures from the Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011) are also presented below to demonstrate that the composited PJO variability in the model closely resembles that of the reanalyses, as was also demonstrated in detail in the version of CMAM considered by HSM. The model underlying MERRA has a horizontal resolution of 0.5° latitude by 0.75° longitude, and 72 levels from the surface to 0.01 hPa.

All fields X are decomposed into a climatology X_c and an anomaly X' . Climatologies of all relevant fields are computed from daily fields by averaging over all years, then retaining the first four harmonics of the annual cycle. The full quantity X will be referred to as the net or the absolute value to differentiate it from its anomaly.

3. Abacus plot

To show an overview of the variability of the Arctic vortex in the CMAM simulation, a visualization of the principal component time series ts_1 and ts_2 of the first two EOFs of polar cap (70° – 90° N)–averaged temperature anomalies we term an “abacus plot” is shown in Fig. 1. A full discussion of their construction and further examples of their utility can be found in HSM; however, the details are reviewed here. Both EOFs describe vertical dipoles, the first with maxima in the upper stratosphere and upper mesosphere, and the second with maxima in the midstratosphere and mesosphere. Together they capture nearly 90% of the total polar cap–averaged temperature variance from the surface to the upper mesosphere (not weighted by mass). The time series are first transformed into polar coordinates r and θ , defined by $r^2 = ts_1^2 + ts_2^2$ and $\tan\theta = ts_2/ts_1$. Each year (from September to September) is

represented by a single, vertical ribbon, the width of which is given by r and the color of which is given by θ . Red corresponds to the positive phase of EOF 1, or a warm upper stratosphere. Green corresponds to the positive phase of EOF 2, or a warm midstratosphere. Blue corresponds to a cold upper stratosphere and yellow to a cold midstratosphere, while intermediate phases are interpolated.

An example of a PJO event can be seen beginning in January of model year 18. The ribbon widens rapidly in the red phase, corresponding to a rapid warming in the mid- to upper stratosphere. The ribbon remains wide through to nearly the end of April, while the color changes rapidly from red to green, then more gradually to blue and finally to yellow, corresponding to the rapid descent of the initial warming to the lower stratosphere, while the upper stratosphere cools and the stratopause reforms at what are typically mesospheric heights. The descent of the cold anomaly and elevated stratopause are then represented by the steady change of color in the abacus plot. PJO events are identified algorithmically by first smoothing the principal component (PC) time series with a 5-day low-pass filter. Central dates are then defined by when the phase θ rotates counter clockwise through a given phase $\theta_c = 2\pi/3$, provided that r is greater than a threshold $r_c = 2\sigma$. The duration of the event then corresponds to the period around this central date during which r exceeds a lower threshold $r_m = 1.5\sigma$. These episodes are identified on Fig. 1 by the vertical black lines to the left of the abacus ribbons.

Also marked on Fig. 1 are sudden warmings, identified and classified as vortex displacements (horizontal lines) or vortex splits (horizontal lines with upticks) following Charlton and Polvani (2007). The ubiquity and regularity of PJO events can be seen in Fig. 1, as can their tendency to follow a subset of sudden warmings. Two PJO events are highlighted—one in model year 42 and one in model year 94. These events will be analyzed in detail in section 6.

4. Transient approach to downward control

On the intraseasonal time scales relevant for PJO events, the zonal-mean circulation in the stratosphere is assumed to be stable and in thermal wind balance. We use the TEM formulation of zonal-mean quasi geostrophy on the sphere, following Plumb (1982) and Haynes et al. (1991):

$$\frac{\partial u}{\partial t} - f\bar{v}^* = \mathcal{F}, \quad (1a)$$

$$\frac{\partial T}{\partial t} + S\bar{w}^* = Q, \quad (1b)$$

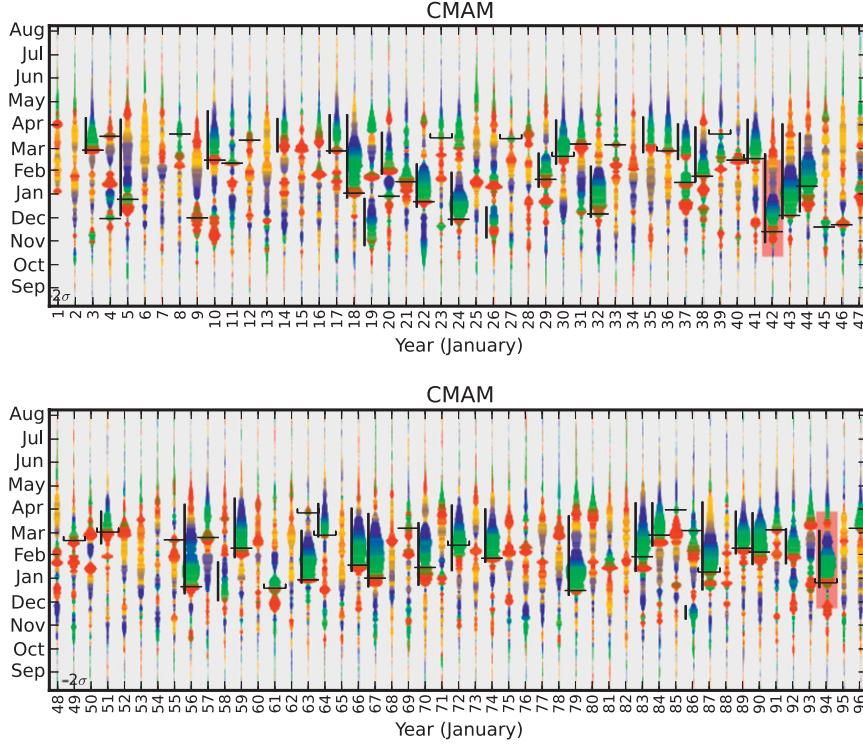


FIG. 1. “Abacus” plot of the CMAM time-slice run. See text for details of construction. Each vertical “ribbon” corresponds to a single year. The width of the ribbon is a measure of the root-mean-square departure of polar cap-averaged temperature from climatology, while the color indicates the vertical structure of the anomalies. PJO events are indicated by the vertical lines to the left of the ribbons. Sudden warmings are shown as horizontal lines, with additional upturns on vortex splits to differentiate them from vortex displacements.

$$\frac{1}{a \cos\phi} \frac{\partial \cos\phi \bar{v}^*}{\partial \phi} + \frac{1}{\rho_0} \frac{\partial \rho_0 \bar{w}^*}{\partial z} = 0, \quad \text{and} \quad (1c)$$

$$f \frac{\partial u}{\partial z} = -\frac{R}{aH} \frac{\partial T}{\partial \phi}. \quad (1d)$$

The zonal wind u and temperature T are forced by torques \mathcal{F} and heating Q , respectively, and by the effects of the residual velocities \bar{v}^* and \bar{w}^* . Here $\rho_0 \propto \exp(-z/H)$ is the reference density profile that also defines the density scale height H , $f = 2\Omega \sin\phi$ is the Coriolis parameter, and Ω and a are the angular velocity and radius, respectively, of the earth. The stratification parameter is related to the square of the buoyancy frequency N^2 by $S = HN^2/R$, where R is the dry gas constant. Our approach will be to specify \mathcal{F} and Q , either by some idealized form or using output directly from CMAM, and compute their effects on the zonal-mean circulation.

The mechanical forcing from CMAM that will be considered can be written as the sum of the resolved eddy-driving \mathcal{F}_r given by the divergence of the Eliassen–Palm (EP) flux \mathbf{F}_r , and the acceleration due to parameterized

gravity waves \mathcal{F}_p , of which there are both orographic and nonorographic varieties:

$$\mathcal{F} = \frac{1}{a \cos\phi \rho_0} \nabla \cdot \mathbf{F}_r + \frac{1}{\rho_0} \frac{\partial F_o^z}{\partial z} + \frac{1}{\rho_0} \frac{\partial F_n^z}{\partial z}. \quad (2)$$

The resolved EP flux can be further subdivided by the zonal wavenumber of the eddies; we will consider the planetary-scale EP flux \mathbf{F}_p to be due to wavenumbers 1–3.

Full expressions for these quantities and details of their calculation are given in the appendix. The diabatic heating Q in the middle atmosphere is dominated by radiative heating, though the output from all parameterized processes is used here.

The meridional circulation is determined by requiring that thermal wind balance in (1d) be maintained under the imposed \mathcal{F} and Q . The instantaneous response to localized forcing of either type is to drive a residual circulation that opposes the effect of the forcing locally, and forms two cells that close in both directions. The influence of the enhanced planetary wave driving during the initial warming, for instance, can lead to anomalous

upwelling in the mesosphere above. The set (1) can be solved for the vertical residual velocity, yielding a linear equation of the form

$$\mathcal{L}\bar{w}_T^* = \mathcal{L}_F\mathcal{F} + \mathcal{L}_Q Q, \quad (3)$$

where \mathcal{L} , \mathcal{L}_F , and \mathcal{L}_Q are linear differential operators defined in the appendix. One particular difficulty in inverting this equation to find \bar{w}_T^* is specifying a lower boundary condition in log-pressure coordinates. Critically, \bar{w}^* does not vanish at the pressure surface used as a lower boundary for this inversion. This occurs for two reasons. First, the overturning circulation can produce a net meridional transport of mass, leading to transient changes in the surface pressure (Haynes and Shepherd 1989). Secondly, the meridional heat fluxes at the lower boundary need not vanish. It has been argued for small-amplitude topography that this term is balanced by Eulerian mean meridional flow within the topographic layer (McIntyre 1980). For the realistic topography in CMAM, however, this was found not to be the case, and the term leads to a nonzero \bar{w}^* as will be shown below. At any rate, these remain an approximation to the true lower boundary condition, though the results are found to be sufficiently accurate for the present purposes. More details of the model and the numerical methods used to solve it are given in the appendix for the purposes of reproducibility.

Under a steadily imposed forcing, however, the temperatures (and thus the diabatic heating) will adjust until the net induced circulation closes entirely downward (Haynes et al. 1991). Following a switch-on forcing in an atmosphere with a constant rate of radiative damping, the time scale on which steady state is expected to be obtained is given by $\Delta z H / \alpha H_R^2$, where Δz is the vertical distance from the height of the forcing, and H_R is a Rossby height that depends on the horizontal scale of the forcing. In the real atmosphere (and in the simulations considered here), there is substantial vertical structure in the radiative damping rates. An important question to address is whether the weak circulation in the lower stratosphere has approached this steady-state limit.

Indeed for an accurate calculation of this adjustment, one requires a sufficiently accurate representation of the diabatic heating. In general, radiative heating is a non-local and nonlinear function of the vertical temperature profile (and that of radiatively active constituents). However, an accurate effective description can be obtained by assuming a linear relaxation of the form $Q = Q_c - \alpha T'$. We follow the regression methodology of Hitchcock et al. (2010), but use zonal-mean temperature anomalies from CMAM during PJO events to obtain a profile of α that is most representative of the radiative damping during these

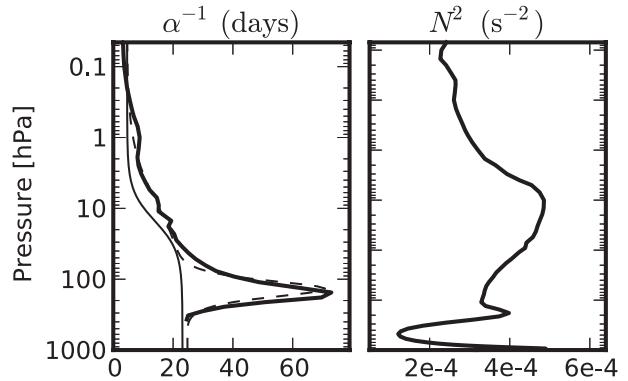


FIG. 2. (a) Vertical profile of radiative damping time scales: Holton-Mass profile (thin solid line), the regressed damping time scales from the events in CMAM (thick solid line), and the analytical approximation to the latter (thin dashed line). (b) Vertical profile of climatological polar cap-averaged (70° – 90° N) Brunt-Väisälä frequency squared in the winter season taken from CMAM.

events. The profile of effective damping time scales computed from the PJO events simulated by CMAM is shown in black in Fig. 2a. For comparison, the profile used by Holton and Mass (1976) (which has been used by a number of subsequent studies) and an analytical approximation to the CMAM rates are also shown. The analytical approximation (day⁻¹) is given by

$$\alpha_{\text{CMAM}} = \alpha_T + \frac{1}{2}(\alpha_M - \alpha_T) \left[1 + \tanh\left(\frac{z - 45 \text{ km}}{14 \text{ km}}\right) \right] - \alpha_{\text{LS}} \exp\left[-\frac{(z - 14 \text{ km})^2}{2 \times (3 \text{ km})^2}\right], \quad (4)$$

with $\alpha_T = 0.04 \text{ day}^{-1}$, $\alpha_M = 0.22 \text{ day}^{-1}$, and $\alpha_{\text{LS}} = 0.028 \text{ day}^{-1}$. Note in particular the extended time scales in the lowermost stratosphere. Also required for inverting (3) is a representative profile of N^2 . Except where otherwise noted, the climatological profile, area averaged from 60° to 90° N over the winter season, is used. This is shown for reference in Fig. 2b.

To test explicitly where the steady-state limit has been achieved, we compute the downward control vertical residual velocity following Haynes et al. (1991). Under the same quasigeostrophic scaling, this is given by

$$\bar{w}_{\text{DC}}^* = -\frac{1}{\rho_0 a \cos\phi} \int_z^\infty \frac{\partial}{\partial \phi} \left(\frac{\rho_0 \cos\phi \mathcal{F}}{f} \right) dz, \quad (5)$$

In this case the circulation is attributed solely to the torques, with the assumption that the diabatic heating field has fully adjusted to balance the implied adiabatic heating.

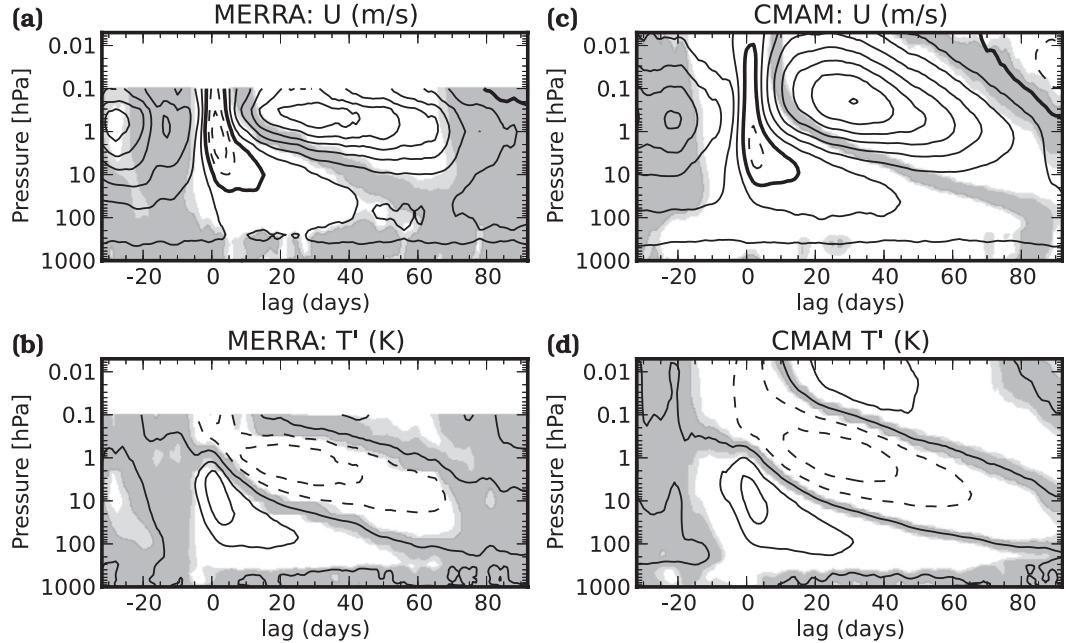


FIG. 3. Composites of PJO events from (a),(b) MERRA and (c),(d) CMAM. (a),(c) Absolute zonal wind, area averaged from 50° to 70° N, at an interval of 10 m s^{-1} . Zero contour is thick, and negative values are dashed. (b),(d) Temperature anomaly, area averaged from 70° to 90° N, at an interval of 10 K . In both cases, gray shading indicates where the composite does not differ significantly from climatology at the 95% (light) and 99% (dark) confidence levels.

5. Composites

To quantify the common aspects of PJO events we first present composite averages of the relevant dynamical fields. To properly resolve the wind reversal that initiates the warming, the definition of the central date used by the composite is modified slightly from the criteria used to identify PJO events in section 3. We take the first day following the initial date when the winds reverse at 10 hPa as the central date; if no such wind reversal occurs during the event, then it is discarded (four such events are found in CMAM and three in MERRA). In this way, we are essentially analyzing only those PJO events that are also classified as sudden warmings.

Figure 3 shows composites of zonal-mean temperatures and winds during these PJO events from MERRA (10 events) and the CMAM simulation (38 events). Figures 3a,c show the absolute zonal-mean zonal wind (area averaged from 50° to 70° N) with a thick zero wind line; the wind reversal at 10 hPa marking the initial date of the sudden warming is evident. The plots are shaded where the composite is indistinguishable from climatology. Winds between 100 and 10 hPa remain weak for 60 – 80 days following the warming in the CMAM composite. The lower-stratospheric winds are quite similar in the MERRA composite, though the smaller number of events results in weaker statistical significance. The polar

jet reforms higher and stronger than its climatology, with winds peaking near 0.5 hPa from 20 to 40 days following the initial warming. The jet maximum descends with time in both composites, although the vertical extent of the winds in CMAM is larger. The vertically compressed structure in MERRA is likely to be an artifact of the model top; similar biases in operational analyses have been reported by Manney et al. (2008).

Figures 3b,d show anomalous temperatures from the two datasets. The composite anomaly during the initial warming is more than 20 K , which agrees well with the reanalysis; note that anomalies during individual events can be several times larger than this. In the lowermost stratosphere, the warm anomaly persists for nearly 80 days; again, the MERRA composite agrees closely with that of CMAM but with weaker statistical significance. Following the mesospheric cooling that coincides with the sudden warming itself, the stratopause reforms at what are typically mesospheric heights, and descends while the lower-stratospheric anomaly persists. The upper warm anomaly is captured only toward lag day 40 in the MERRA composite, when it has descended into the model domain.

The zonal-mean evolution during these events is quantitatively well captured by this CMAM simulation. Some confidence, therefore, can be placed in the relevance of the CMAM composite dynamics to the real

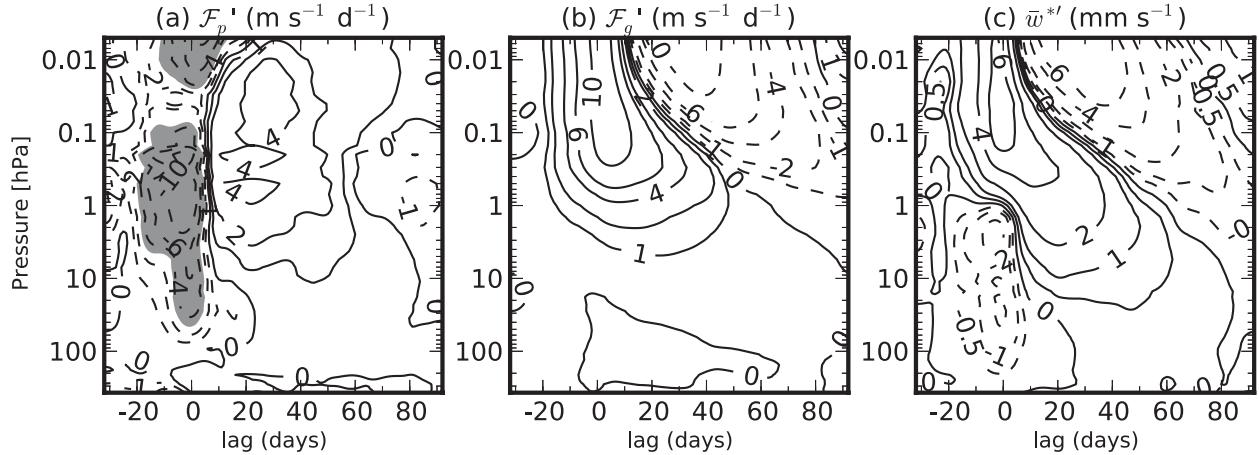


FIG. 4. Composites of PJO events simulated by CMAM. (a) Anomalous acceleration due to planetary-scale (zonal wavenumbers 1–3) EP flux convergence, (b) anomalous acceleration due to parameterized gravity wave drag, and (c) anomalous vertical residual velocity. Quantities in (a) and (b) area averaged from 50° to 70°N , and in (c) from 70° to 90°N . Contours in (a) and (b) have an interval of $2 \text{ m s}^{-1} \text{ day}^{-1}$, with additional contours at $\pm 1 \text{ m s}^{-1} \text{ day}^{-1}$. Contours in (c) are at 0, ± 0.5 , ± 1 , ± 2 , ± 4 , and $\pm 6 \text{ mm s}^{-1}$. Shading in (a) indicates decelerations stronger than $-3 \text{ m s}^{-1} \text{ day}^{-1}$.

atmosphere. We work in the TEM framework, and begin with the zonal-mean accelerations induced by the eddy flux convergences. Figure 4a shows the acceleration due to anomalous EP flux divergence due to planetary-scale fluxes, averaged from 50° to 70°N . The climatological convergence is enhanced prior to the initial wind reversal associated with the sudden warming. Following the wind reversal, the planetary-scale eddies produce an anomalous EP flux divergence. While there are often brief periods of absolute divergences likely produced by the effects of wave transience immediately following the breakdown of the vortex (see, for instance, below), much of this anomalous divergence is indicative of the nearly complete suppression of planetary-scale eddy fluxes noted by HSM. The net gravity wave drag from both orographic and nonorographic parameterizations is shown in Fig. 4b. The climatological drag in the mesosphere is strongly reduced (becoming more westerly) during the onset of the warming, and is reduced at stratospheric heights throughout most of the event. After the initial warming the drag becomes anomalously easterly in a region that descends over time. These forces, in addition to decelerating the winds, contribute to a residual circulation. The anomalous vertical residual velocities are shown in Fig. 4c. The downward circulation in the stratosphere during the warming is anomalously strong, as is the coincident upward circulation in the mesosphere (cf. Ren et al. 2008). These anomalies are consistent with both the enhanced planetary-scale drag in the stratosphere and the gravity wave drag in the mesosphere. During the recovery phase, the stratospheric vertical velocities are anomalously weak (but typically

remain downward in the net). Above, following the descending region of enhanced gravity wave drag, the circulation is anomalously downward.

The parameterized gravity waves are considered in further detail in Fig. 5. The absolute zonal accelerations produced by nonorographic and orographic gravity wave drag are shown in Figs. 5a and b, respectively. The net upward momentum flux from orographic waves is shown in Fig. 5c. The climatological westward drag due to both types of parameterized waves apparent prior to the central date vanishes in the mesosphere coincident with the initial lower-stratospheric wind reversal. The westward winds throughout nearly the entire stratosphere at this point permit eastward phase speed waves to propagate upward; as a result, the net nonorographic drag at this point becomes eastward. The same effect was diagnosed in a version of CMAM constrained by data assimilation following the 2002 Antarctic sudden stratospheric warming (Ren et al. 2008). Note that this contributes to the reformation of the polar jet (Limpasuvan et al. 2012; Tomikawa et al. 2012); the relative importance of this torque will be assessed below. In contrast, the orographic waves are filtered by the lower-stratospheric winds (see the shading in Fig. 5c, indicating composite zonal wind speeds less than 4 m s^{-1}), so their net mesospheric drag vanishes.

As the lower-stratospheric winds recover, the filtering effect switches off, and the westward drag in the mesosphere increases. Over the next 60 days the peak drag descends through the mesosphere. This occurs for both types of parameterized waves, though the nonorographic waves deposit their momentum higher than do

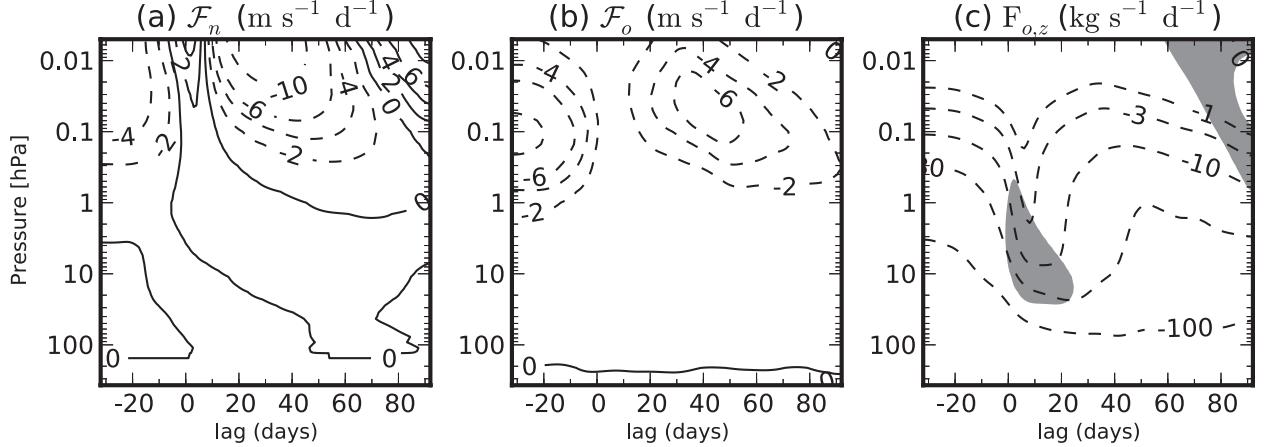


FIG. 5. As in Fig. 4, but for net zonal accelerations due to (a) parameterized nonorographic and (b) parameterized orographic gravity wave drag, and (c) the parameterized net vertical flux of momentum from orographic gravity waves. Contour interval for (a) and (b) is $2 \text{ m s}^{-1} \text{ day}^{-1}$. Contours for (c) are logarithmically spaced as labeled. Shading in (c) indicates zonal-mean winds less than 4 m s^{-1} . Quantities are area averaged from 50° to 70°N .

the orographic waves. This descent is perhaps most obvious in the downward-tilting contours of the orographic wave flux. The reasons for this descent will be discussed in section 7.

Figure 6a shows composites of the zonal wind acceleration. The deceleration in the stratosphere just prior to the central date is clearly induced by the strong, planetary-scale eddies (Fig. 4a). As is well known (Matsuno 1971), the component of the eddy forcing that does not directly decelerate the winds drives a strong residual meridional circulation. The largest accelerations are in the mesosphere, in contrast to the temperature tendencies (Fig. 6b), which are of similar magnitude throughout the middle atmosphere. The temperature tendencies are driven by the sum of the adiabatic heating (Fig. 6c) and diabatic heating (Fig. 6d). The former correlates strongly with the anomalous vertical velocities (Fig. 4c). The initial warming is adiabatically induced as expected, and the descending stratopause is maintained adiabatically, as well.

However, the adiabatic heating $S\bar{w}^*$ is also determined by the static stability, and in principle anomalies in the latter may also be important in determining the structure of the adiabatic heating. The change in static stability during these events is shown in Fig. 6e. The anomalies have considerable structure, but they are rarely more than 15% of the climatological values. They are therefore quantitatively important to the circulation, as will be discussed below, but are not qualitatively so. We note in particular the strongly enhanced stability just above the tropopause in the composite. This enhanced tropopause inversion layer (TIL) forms in the polar night (when the climatological TIL is weak) and is

clearly associated with the extended recoveries from sudden warmings. The mechanisms by which this transient TIL forms are unlikely to be the same as those responsible for the extratropical TIL outside the polar night (Birner 2006). To our knowledge this enhanced polar TIL has not been previously reported in the literature. Comparing Figs. 6c and e, however, shows that much of the former can be explained by assuming a climatological value of N^2 .

The diabatic heating is very strongly correlated with the anomalous temperatures. The heating rates predicted by a Newtonian cooling term are shown in Fig. 6f. The magnitude is somewhat overpredicted near 5 hPa, which is consistent with the weakly nonlinear effects from the curvature of the Planck function found by Hitchcock et al. (2010).

The picture that emerges from these composites, then, agrees well with the analyses of similar events in other comprehensive middle-atmosphere models (Siskind et al. 2010; Limpasuvan et al. 2012; Tomikawa et al. 2012), and is as follows. The initial stratospheric warming is induced by the enhanced planetary wave drag, while the coincident mesospheric cooling is a result of the strongly reduced residual circulation. This reduction is a combined effect of the anomalous gravity wave drag and of transient effects from the resolved drag; the relative role of each in CMAM will be clarified below. The subsequent upper-stratospheric cooling is a result of the suppressed dynamical driving and strong radiative cooling. Although the composited anomalous adiabatic heating responsible for the initial warming ceases nearly simultaneously throughout the depth of the stratosphere, the anomaly persists longer at lower altitudes in the

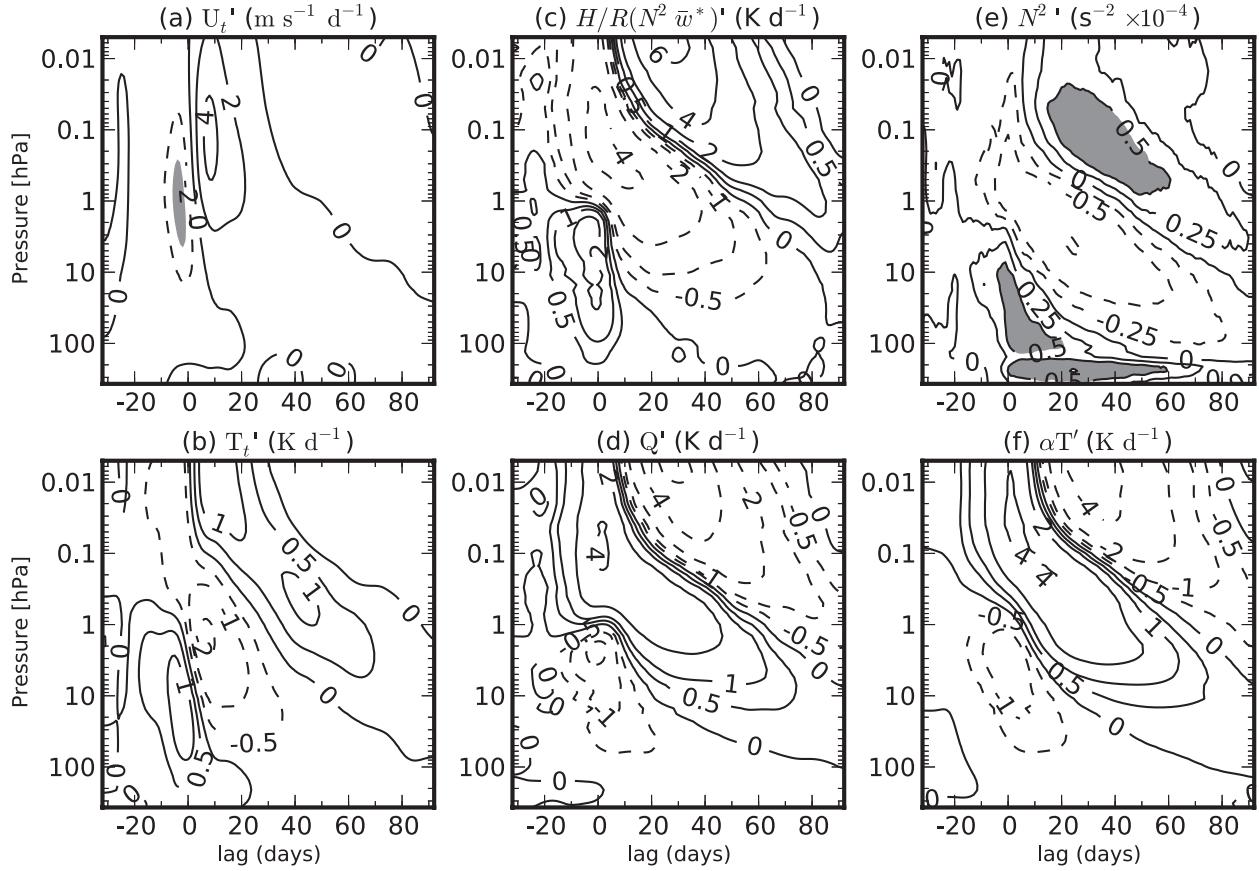


FIG. 6. As in Fig. 4, but for (a) zonal wind acceleration, (b) temperature tendencies, (c) anomalous adiabatic heating, (d) anomalous diabatic heating, (e) anomalous squared Brunt–Väisälä frequency, and (f) linearly predicted anomalous diabatic heating. Contours in (a) are at an interval of $2 \text{ m s}^{-1} \text{ day}^{-1}$. Contours in (b)–(d) and (f) are at $0, \pm 0.5, \pm 1, \pm 2, \pm 4, \text{ and } \pm 6 \text{ K day}^{-1}$. Contour interval in (e) is $2.5 \times 10^{-5} \text{ s}^{-2}$. Gray shading in (e) indicates absolute anomalies greater than $5 \times 10^{-5} \text{ s}^{-2}$. Contours in (c) are at the same numerical values as in (b), but are in K day^{-1} . The zonal wind acceleration is area averaged from 50° to 70°N ; all other quantities are area averaged from 70° to 90°N .

stratosphere. The apparent descent of the cold anomaly is associated, therefore, with the vertical gradient in radiative cooling. The stratopause reforms high in the mesosphere as a result of enhanced and anomalously high gravity wave drag, and the subsequent descent follows that of the region of enhanced drag.

6. Case study

To understand these events more quantitatively, we focus henceforth on a particular case study from the CMAM simulation. The details of this particular event are quite similar to those of the composite just discussed, so much of the interpretation carries over to the general case. There are two advantages to analyzing a single event. First, there is a significant seasonal cycle through the extended winter period when these events occur, so by focusing on individual events we can better

understand the dynamics in an absolute frame of reference. Second, unlike the composite average, this individual event is a solution to the equations of motion integrated by CMAM. Moreover, the eddies depend quite strongly on the zonal-mean state. Focusing on an individual event therefore ensures that features are a result of physical processes, and are not statistical artifacts. It must be stressed that similar analyses of other events yield the same general conclusions (though some particularities of the event chosen are convenient for the purposes of presentation; these are highlighted as they arise). A second example, in this case for a vortex splitting event, is presented once the full decomposition of the transient response has been introduced below.

The particular event discussed here is the PJO event that begins in November of model year 41 (and continues through to March of year 42). It follows a vortex displacement (see highlighted event in top panel of

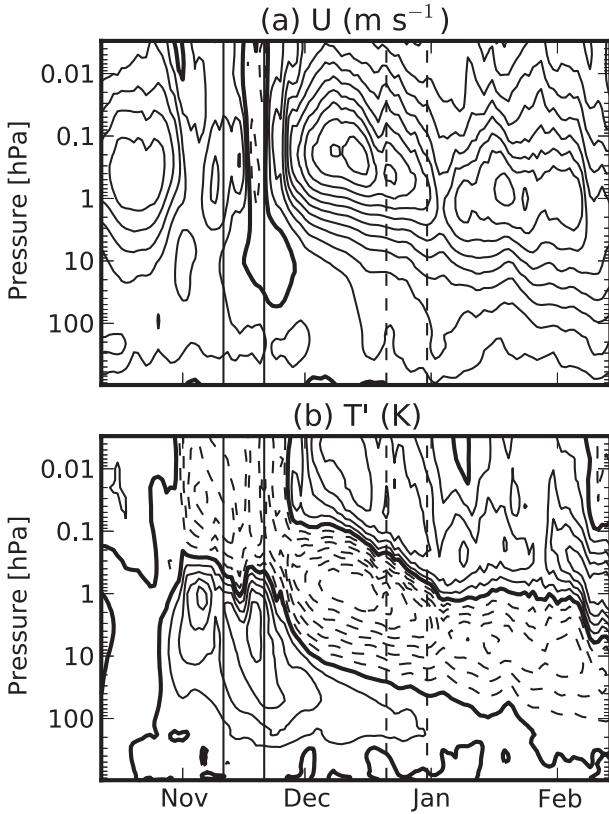


FIG. 7. PJO event from model years 41–42. (a) Zonal wind (50° – 70° N): contour interval of 10 m s^{-1} and zero contour is thick. (b) Temperature anomaly (70° – 90° N): contour interval of 5 K and zero contour is thick.

Fig. 1). To demonstrate its similarity to the composited average, the associated zonal wind and temperature anomalies are shown in Fig. 7. The absolute zonal-mean zonal winds from 50° to 70° N (Fig. 7a; compare with Fig. 3c) reverse in the middle stratosphere in mid-November, after which the polar jet reforms higher and stronger than climatology. The lower-stratospheric winds remain weaker than 10 m s^{-1} through mid-December, and weaker than 20 m s^{-1} through mid-January. The polar cap temperature anomalies (Fig. 7b; compare with Fig. 3d) show the characteristic descending tripolar anomaly following the peak warming in November, with an extended lower-stratospheric anomaly, a strong upper-stratospheric cold anomaly, and an elevated stratopause. The descent of the anomalous stratopause is interrupted at the beginning of January when the lower-stratospheric winds briefly accelerate; the occurrence and timing of this feature is particular to this event, though similar behavior is not unusual in other events simulated by CMAM. Similarly, the strong precursor to the warming in early November is present in some but by no means in all events.

The absolute acceleration due to planetary-scale EP flux convergence (Fig. 8a; compare with Fig. 4a) shows two initial pulses of convergence (on the order of $-30 \text{ m s}^{-1} \text{ day}^{-1}$) responsible for the early November precursor and the mid-November warming. Note that logarithmic contour shading is used here to show details of the stratospheric forcing following the warming. Immediately following the warming, the net flux is divergent as was alluded to above. This feature is common to many events in CMAM, and has also been noted by Tomikawa et al. (2012). Following this, however, the amplitude of the forcing remains below about $1 \text{ m s}^{-1} \text{ day}^{-1}$, indicating essentially an absence of planetary-scale wave drag. Note as well the strong region of convergence at mesospheric heights immediately following the initial warming. The net acceleration due to parameterized gravity wave drag (Fig. 8b; compare with Fig. 4b) weakens in the mesosphere during the onset of the warming, and becomes westerly roughly where the winds reverse. Enhanced drag can be seen roughly coincident with the elevated and descending stratopause.

The absolute residual vertical velocity (Fig. 8c; compare with Fig. 4c) shows strong downwelling coincident with the pulses of EP flux convergence that drive the initial warming. As implied by the composite, the circulation in the stratosphere following the warming is much weaker than climatology; in the net, however, it remains weakly downward nearly the entire time. Accounting accurately for the dynamical heating associated with this circulation will be important for understanding the persistence of the lower-stratospheric anomaly. The strong circulation responsible for the descending stratopause is also apparent.

a. Residual circulation

To identify where and to what extent the zonal-mean dynamics are still subject to the transient approach to the downward control limit, the instantaneous and steady-state vertical residual velocity is computed according to (3) and (5), respectively. Since both are essentially linear calculations, they permit the circulation to be decomposed according to the source of mechanical forcing or, in the transient case, diabatic heating.

The results of these decompositions are shown in Fig. 9 for the absolute residual circulation over the polar cap, averaged over two separate phases of the event. The decomposition of the circulation is shown for the transient case in the middle panels, and the downward control case in the right panels. The static stability profile, required by (3), is taken to be the climatological profile shown in Fig. 2. The first phase (Fig. 9a; indicated by the solid vertical lines in Figs. 7 and 8) is during the initial peak warming period, when the resolved wave driving is

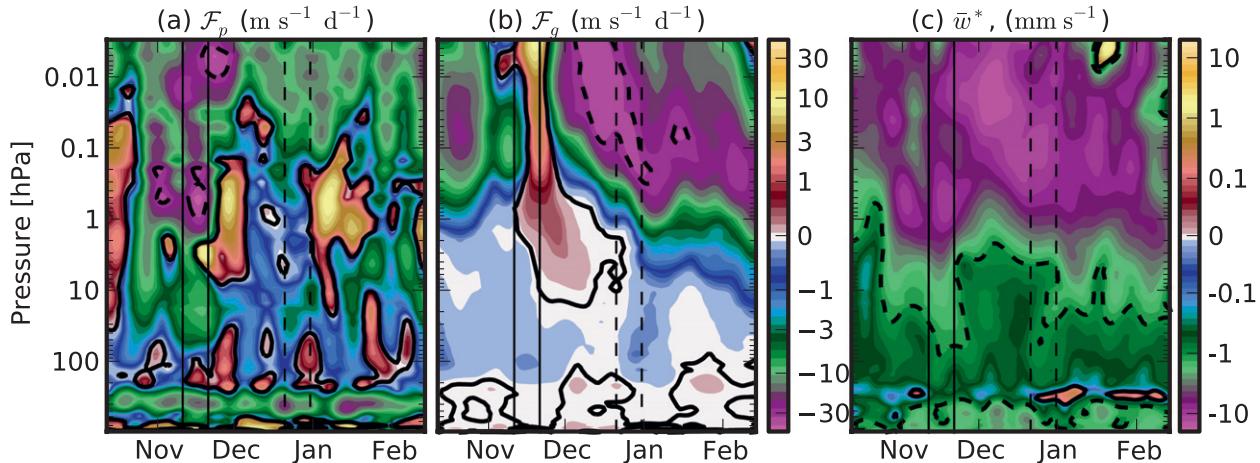


FIG. 8. PJO event from model years 41–42. Acceleration due to (a) EP flux convergence of zonal wavenumbers 1 and 2 and (b) parameterized gravity waves, area averaged from 50° to 70°N , with contour lines at -30 and $0\text{ m s}^{-1}\text{ day}^{-1}$. (c) Residual vertical circulation (70° – 90°N), with contour lines at -1 and 0 mm s^{-1} . Note that the shaded contours are spaced logarithmically, and all quantities have been smoothed with a 5-day low-pass filter.

at its strongest and most transient, and thus where one would expect the largest discrepancy between the transient and equilibrium circulations. Indeed, the downward control circulation overpredicts the magnitude of the circulation everywhere in the stratosphere and mesosphere, particularly so near 1 hPa. The agreement between the transient circulation and the true circulation is substantially better, with only a slight overprediction of the magnitude of the response in the mesosphere. The transient decomposition shows a significant contribution from the diabatic heating, and, in the lower stratosphere in particular, from the near-surface heat flux. The downward control estimate overpredicts the circulation because the temperatures (and thus the diabatic heating) have not had time to adjust to the steady state.

The second phase considered (Fig. 9b; indicated by the dashed vertical lines in Figs. 7 and 8) is nearly a month following the initial warming. At this phase of the PJO event the planetary-wave driving is strongly suppressed in the stratosphere, and the stratopause is in the midst of its descent. This is reflected in the circulation, with strong downward velocities above 1 hPa, and very weak downward velocities from 100 to 10 hPa. At this stage one might expect the downward control estimate to be more accurate than in the first stage considered, given that the imposed torques are varying relatively slowly with time. Indeed, the downward control estimate is considerably more accurate. However, careful inspection of the lowermost stratosphere (inset in Fig. 9b, left) shows that the magnitude of the circulation is underestimated, indicating that steady state has not yet been reached. The transient decomposition indicates three main contributions to the lower-stratospheric circulation: a net upward circulation

arising from the planetary-scale EP flux convergence (this is consistent with the tropospheric convergence and lower-stratospheric divergence seen in Fig. 8c), and downward circulations arising from the diabatic heating and the near-surface heat flux. The importance at these heights of including the latter in the lower boundary condition is evident. In the upper stratosphere, both decompositions show that waves of all scales considered here play a role in the stratopause circulation, with the orographic waves contributing most substantially just above 1 hPa.

To summarize, both resolved and parameterized waves play a role in the reformation of the stratopause following the initial warming, with the net orographic wave drag playing an important role near the base of the mesosphere. The response of the residual circulation to the imposed torques during this PJO event shows clear signs of transience at all heights during the initial warming, and in the lower stratosphere during the recovery phase. The lower-stratospheric residual circulation following the initial warming is, in the net, weakly downward, with a significant contribution from the diabatic heating. Including the near-surface heat flux is essential for the transient decomposition. It remains to be seen, however, how this transience in the response of the overturning circulation influences the persistence of the lower-stratospheric circulation anomaly.

b. Temperature anomalies

Thus, we turn now to the description of the temperature anomalies themselves, focusing initially on the stratospheric component of the circulation. As demonstrated in the previous section, the net adiabatic heating in the stratosphere is weak following the initial warming.

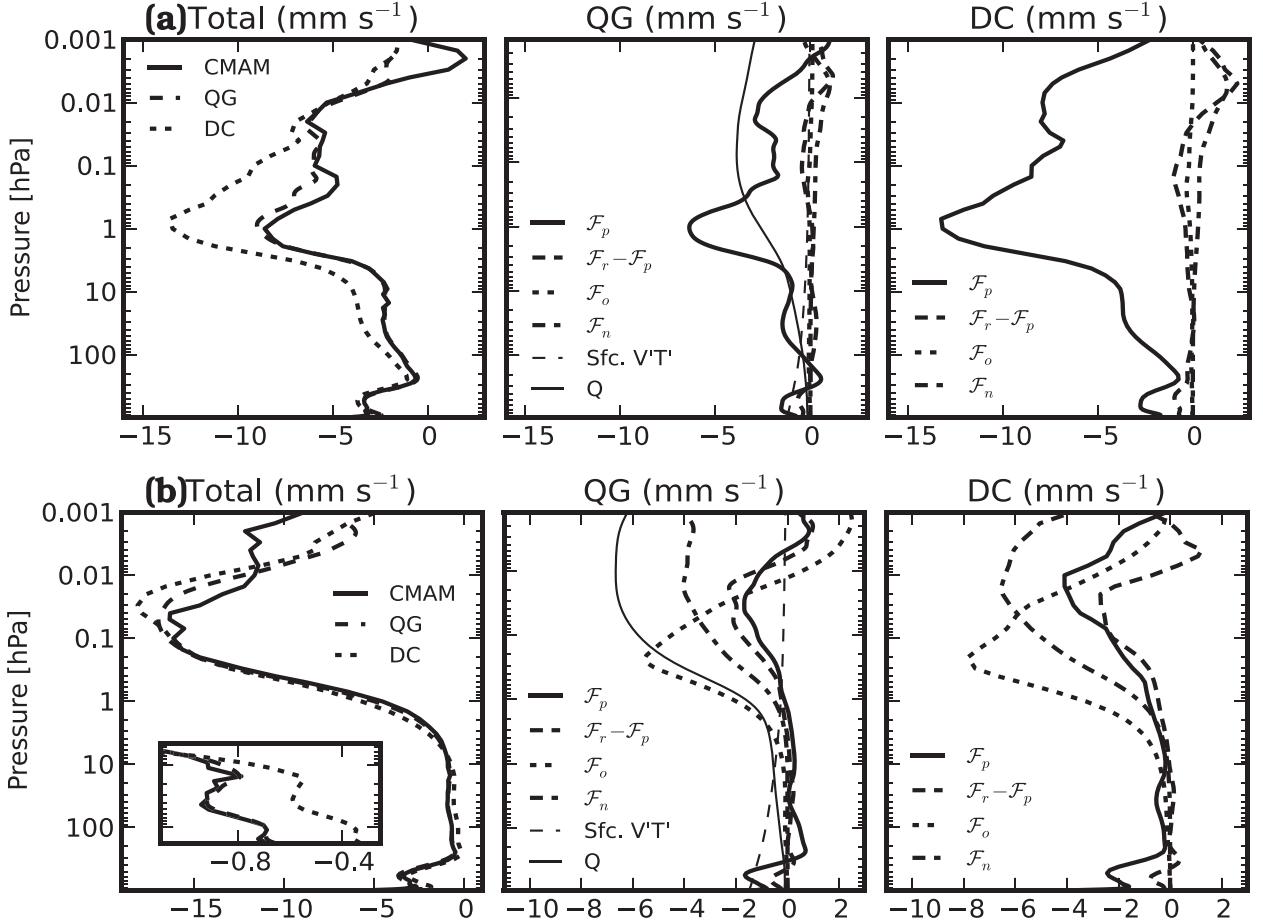


FIG. 9. Decompositions of the polar cap-averaged vertical residual circulation from (a) 20 to 30 Nov and (b) 20 to 30 Dec, model year 41, during the event shown in Fig. 8. (left) The total circulation simulated by CMAM (solid), and as estimated by the transient decomposition (dashed) and by the downward control decomposition (dotted). (bottom left) The inset shows the lower-stratospheric circulation in greater detail (the vertical axis of the inset matches that of the panel). (middle) The transient circulation induced by resolved planetary-scale wave drag, resolved waves of all other wavenumbers, parameterized gravity wave drag, the surface heat flux, and the diabatic heating. (right) The downward control contribution from the four categories of wave drag.

The first approximation, considered in the next subsection, is thus to assume that the temperature anomalies in this phase, while initially produced by the planetary-scale eddy driving, subsequently evolve entirely in response to the diabatic heating. This picture is refined in the second subsection by including the residual circulation induced by the diabatic heating itself. Finally, in the third subsection the full, transient response to the imposed torques will be presented, and the attention broadened to the full middle-atmosphere response.

1) PURE RADIATIVE RELAXATION

In the absence of dynamical heating, longwave radiation will cool the stratosphere toward radiative equilibrium. The net heating rate Q is then given by the climatological cooling rate Q_c , in addition to the linear relaxational component due to the temperature anomaly

$-\alpha T'$. This is most naturally expressed in an absolute framework (i.e., $\partial_t T = Q$); however, since the present focus is on the behavior of the anomalies, it will be more consistent with later arguments to work in an anomaly framework. In transforming to this framework, there arises an additional term due to the climatological temperature tendency (in other words, the absolute temperature should stay constant in the absence of heating of any kind). In equations,

$$\begin{aligned} \frac{\partial T}{\partial t} &= Q, \\ \frac{\partial}{\partial t}(T' + T_c) &= Q_c - \alpha T', \\ \frac{\partial T'}{\partial t} &= Q_c - \frac{\partial T_c}{\partial t} - \alpha T'. \end{aligned} \quad (6)$$

Note that this is equivalent to the fixed dynamic heating approximation in which the absolute dynamic heating is set to zero. This could be equivalently imposed by applying the negative of the climatological dynamic heating, but since the latter is a much noisier quantity than the climatological diabatic heating, the above approach is used.

Two particularities of this case make it especially suitable for this simple analysis. First, the initial warming occurs almost simultaneously throughout the stratosphere. As was pointed out by Zhou et al. (2002), often additional pulses of wave activity are responsible for bringing the warm anomaly down from the middle to the lower stratosphere. That these are absent in this case is convenient for the present analysis, as we can simply take the anomalies on a date near this peak as the initial conditions for integrating the above equation. Second, this particular event occurred very early in the season, so that much of the recovery phase occurs during midwinter. The background seasonal cycle is therefore quite weak, and the corresponding term is essentially negligible (though it is included for completeness). The term plays a more significant role in events that occur later in the season.

The results of three integrations are shown in Fig. 10. The temperature anomalies in each case are initialized at the date indicated by the vertical line, but they differ in the choice of radiative damping rates applied. Figure 10a uses a constant rate of $1/10$ day $^{-1}$. The lower-stratospheric anomaly persists for longer than the upper-stratospheric anomaly because of the weaker climatological cooling. The minimum in climatological cooling near 200 hPa is apparent. Figure 10b uses the profile of Holton and Mass (1976). The weaker damping in the lower stratosphere extends the lower-stratospheric anomaly, while the stronger damping in the upper stratosphere reduces the cooling rate and thus the magnitude of the cooling anomaly. Finally, using the profile fit to the effective rates computed from these runs, the persistence of the lower-stratospheric anomaly is extended substantially.

Two points bear emphasizing here. First, there is no vertical transfer of information in this model, not even numerical diffusion. The downward migration of the stratospheric cold anomaly, which is captured here, can therefore be explained to a large extent as a “phase descent” of a character similar to that discussed by Plumb and Semeniuk (2003). In this case the descent is due in part to the vertical gradient in damping rates, and in part due to the gradient in background cooling rates, which act in the absence of dynamic heating. Second, the lower-stratospheric anomaly simulated by CMAM itself persists for substantially longer than this simple model predicts (cf. the $T' = 0$ K contour in Fig. 10c and Fig. 8a), even given the extended radiative time scales, suggesting a further source of persistence.

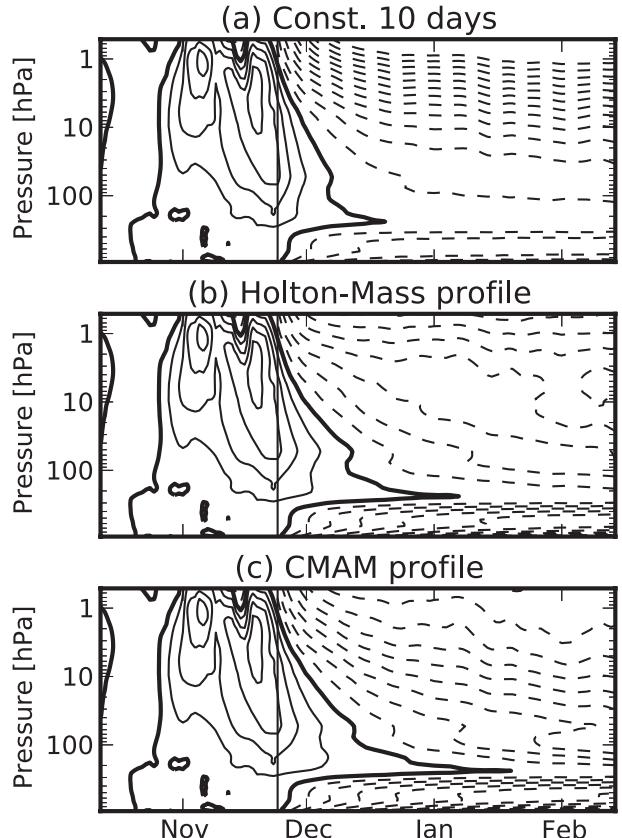


FIG. 10. Polar cap temperature anomalies predicted by radiative relaxation alone (see text), using three profiles of radiative damping time scales: (a) 10 days everywhere, (b) the profile of Holton and Mass (1976), and (c) the analytical approximation to the effective rates computed from CMAM. Contours as in Fig. 7b.

2) RADIATIVE RELAXATION WITH ELIASSEN ADJUSTMENT

The underprediction of the persistence of the lower-stratospheric anomaly from radiative considerations alone is consistent with the weak downward circulation found in Fig. 9b. We therefore extend the simple radiative relaxation used in the previous section by adding in the effects of the Eliassen adjustment induced by the diabatic heating rates themselves:

$$\frac{\partial T'}{\partial t} = Q_c - \frac{\partial T_c}{\partial t} - \alpha T' - S \bar{w}_T^*. \quad (7)$$

Here the circulation is predicted using (3) from the absolute diabatic heating rates predicted by Newtonian cooling, leaving \mathcal{F} still set to zero. In this section a constant profile of static stability is used ($N^2 = 4 \times 10^{-4}$ s $^{-2}$), to clarify the Eliassen adjustment to the diabatic heating in as simple a context as possible. The analytical fit to the CMAM damping rates is retained, and the result is shown

in Fig. 11a. The effect is to extend the lifetime of the lower-stratospheric anomaly; essentially, some of the diabatic cooling is going into accelerating the winds rather than into lowering the temperatures.

Figures 11b,c compare the predictions of these conceptual models in greater detail to the event itself, showing temperature time series at 1 and 200 hPa, respectively. At 1 hPa the temperature anomaly associated with the two pulses of planetary waves (thick solid line) has nearly decayed by the beginning of the integration. The temperatures drop rapidly to nearly 30 K below climatological values (thick dashed line), before returning to climatological values by the beginning of January. The initial drop in temperature is reasonably well described by all three integrations: the radiative relaxation with a 10-day time scale (thin dotted line) and with the CMAM profile (thin dashed line), and by the radiative relaxation with Eliassen adjustment (thin solid line). All three integrations cool toward their respective radiative equilibrium temperatures; since the subsequent warming in the real event requires further wave-driven heating, these integrations do not capture it. At 200 hPa the assumption that there is no torque-induced heating is a better approximation. Here, the warm anomaly in the true event persists significantly longer even than the straight radiative relaxation with the CMAM profile, better matching the integration with Eliassen adjustment. The oscillations with periods of roughly 2 weeks that are absent in the conceptual models are due to the weak pulses of planetary-scale wave driving that can be seen in the lower stratosphere in Fig. 8a.

Note that if the climatological heating rates were neglected in the anomaly framework, one would over-predict the persistence of the anomaly, even in the absence of Eliassen adjustment. This would, however, implicitly include the climatological dynamic heating, which is absent in the aftermath of the warming. While it is true, therefore, that the time scale of PJO events exceeds that of radiative relaxation, one must be careful about the reference frame in which this is meant.

Finally, it is stressed that although the vertically synchronous nature of the initial warming in this event is not particularly universal, the ultimate suppression of planetary wave driving in the vortex is robust. The conclusion that the downward migration of the cold anomaly is ultimately a result of the gradient in both the radiative damping rates and the climatological cooling rates thus holds more generally.

3) FULL TRANSIENT ADJUSTMENT

Having established the radiative origins of the persistence of the lower-stratospheric anomaly, we now consider the full transient response to each type of zonal-mean

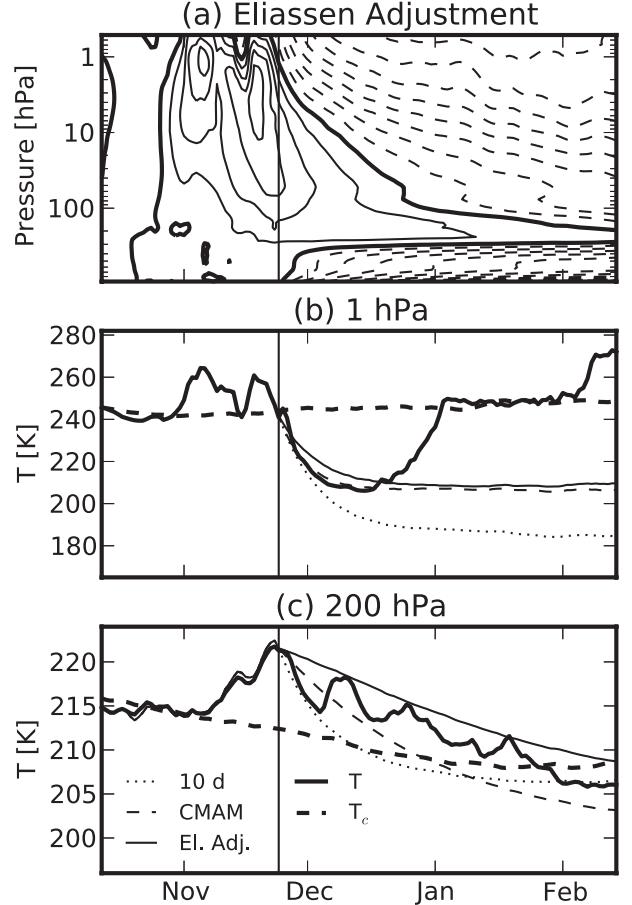


FIG. 11. (a) Polar cap temperature anomalies predicted by radiative relaxation with Eliassen adjustment (see text); contours as in Fig. 7b. Also, comparison of temperature time series at (b) 1 and (c) 200 hPa. Shown are the temperatures from the case study (solid lines), climatological temperatures (dashed lines), temperatures predicted by radiative relaxation alone with a constant damping time scale of 10 days (thin dotted lines) and the analytical profile (thin dashed lines), and by radiative relaxation with Eliassen adjustment (thin solid lines).

torque. We work explicitly with anomalies from the climatology, integrating

$$\frac{\partial T'}{\partial t} = -\alpha T' - S\bar{w}_T^{*'}, \quad (8)$$

where the anomalous circulation $\bar{w}_T^{*'}$ is given by solving (3) for the imposed anomalous torques \mathcal{F}' and the anomalous diabatic heating predicted by the Newtonian cooling term. This is equivalent to the model solved analytically by Haynes et al. (1991), but with \mathcal{F} taken from a fully self-consistent integration of CMAM, and a realistic profile of radiative damping rates. The climatological radiative cooling (the terms Q_c and $-\partial_t T_c$) is no longer present since it is balanced (implicitly) by

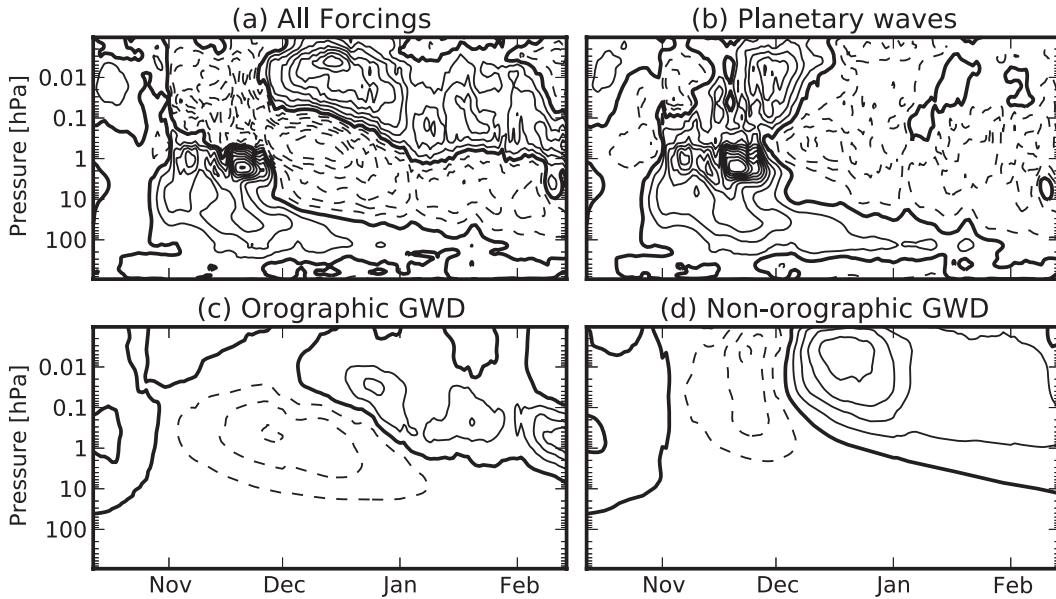


FIG. 12. Polar cap temperature anomalies predicted by the transient adjustment of the circulation to (a) all torques, resolved and parameterized, (b) planetary-scale eddy driving, (c) parameterized orographic gravity wave drag, and (d) parameterized nonorographic gravity wave drag. Contours as in Fig. 7b.

the climatological forcing. As described in section 4, the conceptual framework is that, since the diabatic heating is relaxational, one can ultimately associate the temperature anomalies and residual circulations induced by the heating with the torque that drove the radiative imbalance in the first place. The anomaly framework is adopted to retain the connection to the composite picture. Since the goal here is to reproduce the polar cap temperature field produced by CMAM, the polar cap averaged static stability profile of Fig. 2b is used here.

The results of this integration are shown in Fig. 12. The total response to imposing all torques present in the full CMAM integration (i.e., resolved wave driving at all wavenumbers including the near-surface heat flux, and the parameterized gravity wave drag of both types) is shown in Fig. 12a. This can be compared with Fig. 8a. The details are reasonably well captured: the two initial pulses of planetary-scale waves that warm the stratosphere in November are apparent, following which the warm anomaly in the lower stratosphere persists nearly to the end of January. The amplitude of the cold anomaly in the upper stratosphere following the warming is somewhat underpredicted and the amplitude of the mesospheric anomalies are somewhat overpredicted, but the details of the descending stratopause, including the arrest in its descent at the beginning of January, are well captured. The response can thus be decomposed into the response to each of the dominant forcings.

The initial warming, as expected, is nearly entirely explained by the planetary-scale EP flux convergence

(Fig. 12b; note that the planetary-scale near-surface heat fluxes are also included). The mesospheric response to the initial pulse is to cool; that this is caused by the transient upward-closing cell of the residual circulation can be inferred by noting that the anomalous EP flux divergence is everywhere negative over the pole (not shown); the cooling predicted by the transient response is not, therefore, a result of reduced drag. The secondary pulse, however, generates considerable mesospheric warming as well, contributing to the initial formation of the high stratopause. This is a relatively common feature of the events simulated by CMAM in this run; whether this is true of the atmosphere is not clear from reanalyses but has been noted by Siskind et al. (2010) and Tomikawa et al. (2012). The lower-stratospheric response is in fact more persistent, while the cold anomaly in the upper stratosphere is considerably weaker than the full response.

The reduced orographic gravity wave drag in the upper stratosphere and lower mesosphere following the initial two pulses of planetary waves (and resulting deceleration of the polar jet) results in significant radiative cooling (Fig. 12c)—that this should be thought of as a radiative response is clear from the weak net orographic drag shown in Fig. 5b. This plays a role both in the mesospheric cold anomaly during the initial warming and to a significant extent in the stratospheric cold anomaly as well. By mid-December the response becomes anomalously warm near the top of the model domain. This warm anomaly descends through December

and persists through January, contributing somewhat less than half of the total anomaly seen in the full response.

The remainder of the anomalous mesospheric cooling during the initial stratospheric warming, and the anomalous mesospheric warming during the PJO event, is explained by the nonorographic parameterized gravity waves (Fig. 12d). Unlike the orographic wave drag, some of the cold anomaly is in fact a result of a net eastward torque during the initial warming as the westward waves are filtered by the wind reversal in the stratosphere. The thermal response to the eastward torque imposed by these waves near the base of the reforming stratospheric jet is negligible.

To investigate the possible dynamical relevance of the anomalous static stability structures seen in Fig. 6e to the zonal-mean adjustment, we have repeated the calculation using the time-dependent polar cap profile of N^2 produced by CMAM during this event, which exhibits the same features seen in the composite (not shown). This could affect adjustment in two ways: by modifying the calculation of the transient circulation by (3), or by modifying the adiabatic heating rates directly. The difference in the response to all forcings is shown in Fig. 13a. The correction is clearly a minor one, with notable differences arising only near the base of the descending stratopause. In particular, the enhanced polar TIL produced by the lower-stratospheric warming does not play a significant role in extending the time scales of the lower-stratospheric anomaly. Note that it may still play a role in the suppression of planetary waves entering the vortex; such an effect was suggested by Chen and Robinson (1992). A detailed investigation of this effect is left for future study.

The effect of the extended radiative time scale in the lowermost stratosphere can also be estimated by computing the response assuming the Holton–Mass profile, which does not include this feature (Fig. 13b). The use of the more realistic profile does lead to a more persistently warm lower stratosphere, which provides additional confirmation that the radiative time scales in this region are significantly longer than assumed by the Holton–Mass profile. Differences are also seen at higher altitudes where the radiative damping rates (the inverse of the radiative time scales) differ more markedly.

To demonstrate that the basic dynamics just described apply to other events, Fig. 14 shows the results from a similar calculation for the PJO event that occurs in the winter of model years 93–94 (see Fig. 1). The sudden warming that initiates this event was a vortex-splitting event, as opposed to the displacement that initiated the previously considered event. The temperature anomalies during the event are shown in Fig. 14a, and the anomalies predicted by the transient quasigeostrophic

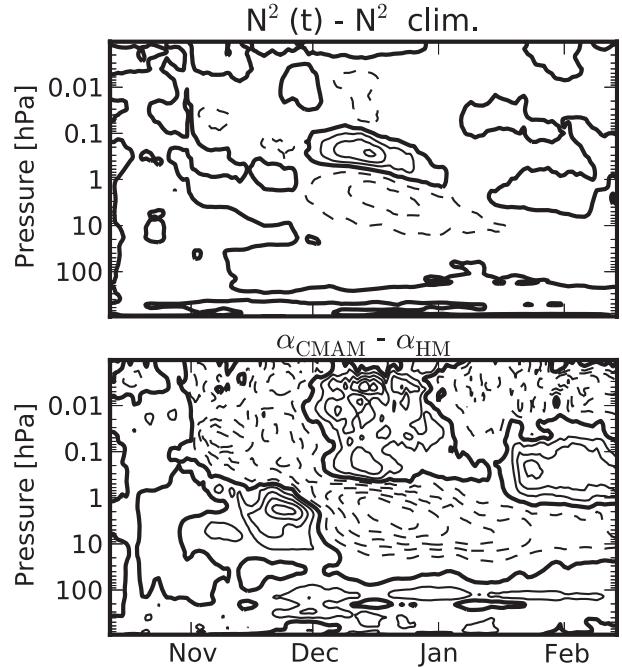


FIG. 13. Difference in the total predicted temperature response as a result of using (a) the time-dependent polar cap profile of N^2 and (b) the Holton–Mass profile of radiative damping time scales instead of the profiles used in Fig. 12. Contour intervals are 3 K and zero contours are thick.

adjustment are shown in Figs. 14c,d. As in the vortex-displacement event, the transient adjustment captures the full model behavior with some fidelity. The initial warming is produced as expected by planetary-scale wave drag (Fig. 14c), as is the initial mesospheric warming, though in this event the drag at mesospheric heights persists for longer than in the displacement event. Nonetheless, the majority of the upper warm anomaly (and its descent) is produced by anomalous gravity wave drag (Fig. 14). Nearly all of the stratospheric cooling in this case is in fact due to reduced gravity wave drag. It is worth noting that in both cases the stratospheric cooling is quite substantial already during the sudden warming itself, suggesting that these events would be even more dramatic were it not for the filtering effects of the weakening vortex on upward-propagating gravity waves.

7. Stratopause descent

We consider finally the descent of the stratopause. It is clear from Fig. 12 that the anomalous temperatures at the top of the stratospheric jet are maintained by torques of a variety of origins; initially, the planetary-scale waves provide some heating, followed by the parameterized nonorographic gravity waves, and finally the parameterized

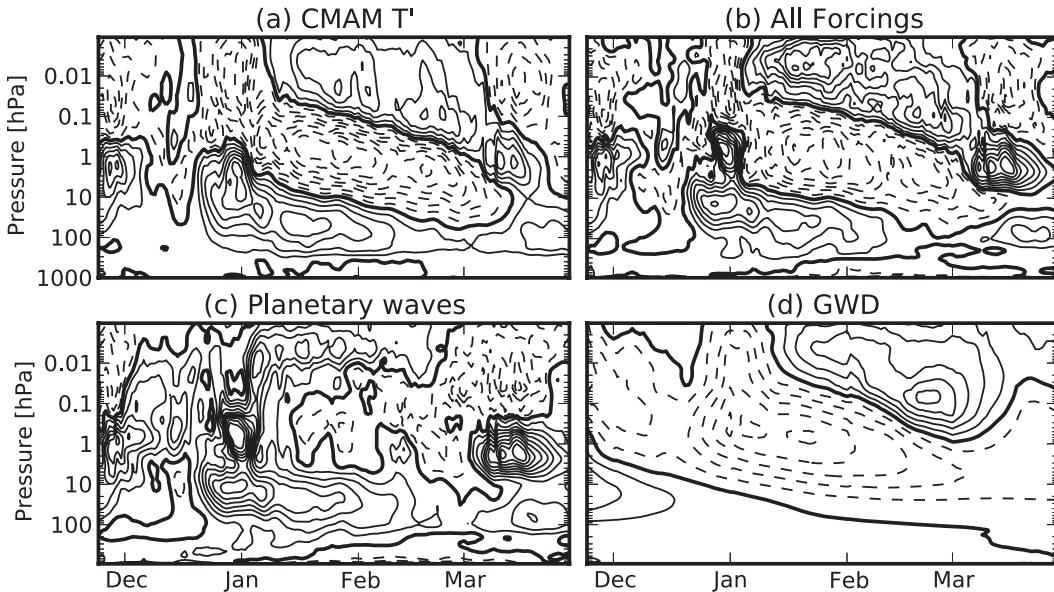


FIG. 14. (a) Polar cap temperature anomalies during the PJO event in model year 94. (b)–(d) Polar cap temperature anomalies predicted by the transient adjustment of the circulation to (b) all torques, resolved and parameterized; (c) planetary-scale eddy driving; and (d) parameterized gravity wave drag of both types. Contours as in Fig. 7b.

orographic gravity waves dominate as the stratopause descends to more climatological heights. The relevance of a variety of scales of forcing is consistent with existing studies of these elevated stratopause events in other models (Siskind et al. 2010; Limpasuvan et al. 2012; Tomikawa et al. 2012). The strong radiative cooling at these heights implies that the temperature anomalies must be actively maintained by the dynamic heating. We consider here the behavior of the orographic wave drag in some detail, partly because it descends most clearly along with the stratopause, and partly because the anomaly picture emphasized in Fig. 12 belies the greater role that the orographic waves play in the climatological polar winter stratopause.

Following the initial warming, the radiative cooling of the stratosphere leads to a high and strong stratospheric jet, evident in Fig. 8b. Three profiles of the zonal-mean zonal wind, averaged over consecutive 10-day periods following the reformation of the jet, are shown in Fig. 15a. The jet maximum is at its highest in the first profile (solid black line), and it subsequently descends and weakens over each of the next two periods (dashed and dotted lines, respectively) as the lower-stratospheric winds slowly accelerate. The upward flux of westward momentum carried by the parameterized waves in the same periods are shown (on a log scale) in Fig. 15b. In all three periods the flux is near constant through the stratosphere until it reaches the height of the jet maximum; above this the flux decreases rapidly. This is a consequence of the strong positive shear in the stratosphere, which

prevents the amplitudes of the waves from growing large enough to saturate. After the jet peaks the combined effects of decreasing density and background wind result in strong breaking because of saturation. The gravity waves then deposit their momentum on the upper flank of the jet (Fig. 15c), eroding the maximum from above and leading to the descent of the jet maximum.

There is also an indication that the source of the waves is increasing throughout the event, since the fluxes throughout the depth of the atmosphere are increasing. This is consistent with the strengthening of near-surface winds at these latitudes.

The descent of the region of wave drag, though similar in some ways to the descending shear zones of the quasi-biennial oscillation (QBO), differs in several key respects. Unlike the tropics, the acceleration of the winds is due to the radiatively generated strong thermal gradient, not to the dissipation of waves of eastward phase speed. Moreover, the waves break as a result of saturation processes, not as a result of critical lines (or the strongly reduced vertical group velocities encountered near them). The region of wave drag is thus above the jet maximum, not below. The parameterized waves dissipate, rather than accelerate, the winds, confirming that the high wind speeds are radiatively rather than dynamically produced. In these events this is permitted by the strong filtering of the waves by the stratospheric wind reversal during the warming itself (Holton 1983). The role of the lower-stratospheric winds in this filtering bears reemphasizing; in some events simulated by CMAM, the recovery of the

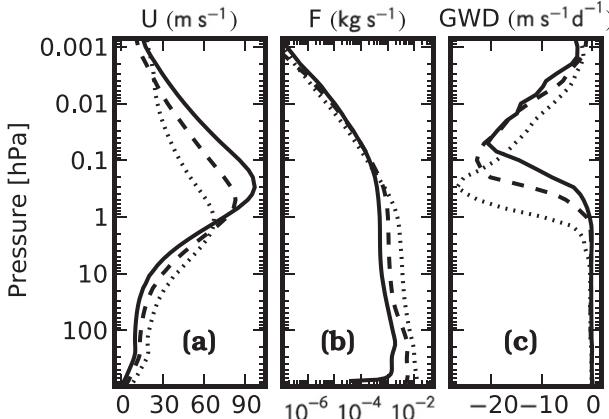


FIG. 15. Vertical profiles of (a) zonal wind, (b) upward flux of westward momentum carried by parameterized orographic gravity waves, and (c) acceleration due to orographic gravity wave drag. All quantities are averaged from 50° to 70°N. The solid, dashed, and dotted lines are averages over consecutive 10-day periods following 22 Dec, model year 41.

lower-stratospheric winds is interrupted by a further pulse of planetary waves. If this decelerates the winds sufficiently, the parameterized waves are filtered strongly, and the descent of the stratopause above is arrested. The lower-stratospheric anomaly therefore has a great deal of influence on the circulation above as well as on the tropospheric response below.

8. Conclusions

The purpose of this paper is to clarify the dynamics that lead to the characteristic anomalous zonal-mean circulation that occurs during the extended time-scale recoveries from sudden warmings, termed here PJO events. Particular emphasis has been placed on the structure of the temperature anomalies. This evolution is robust enough that details of individual events closely resemble those of the composite average, sufficiently so, that many conclusions from the study of an individual event carry over to the general case. One implication of this robustness is that these circulation anomalies (or at least their zonal-mean component) should be quite predictable in the stratosphere and mesosphere during these events.

The residual circulation and zonal-mean temperatures simulated by CMAM during these events are found to be well described by a zonal-mean quasigeostrophic model on the sphere, given the zonal-mean torques produced by CMAM and an appropriate linearized description of the radiative heating rates. This provides a means of decomposing the transient zonal-mean adjustment toward the steady-state, downward

control limit attributable to the various types of wave-driven forcings represented in CMAM.

The zonal-mean dynamics of the initial sudden warming are well known and are corroborated by the present analysis; the rapid warming of the stratosphere is a result of strong EP flux convergence associated with the first two zonal wavenumbers. PJO events are distinguished from other sudden warmings in that this initial warming descends to the lowermost stratosphere; subsequently, the composite mean-resolved wave driving in the stratosphere is strongly suppressed, confirming the result of HSM. The persistence of this lower-stratospheric anomaly in the prolonged absence of climatological wave driving is understood here as a transient Eliassen adjustment toward the steady-state downward control limit. The time scale of the anomaly is shown here to be a result not only of the weak radiative cooling in this region (a result both of long radiative time scales and weak climatological cooling) but also of the dynamic heating induced by the Eliassen adjustment to the resulting diabatic cooling. The apparent descent of the cold anomaly through the stratosphere can be explained by the vertical gradient in both the radiative time scales and in the climatological rates of cooling. Suppression of parameterized gravity wave drag in the stratosphere following the warming was found to be required to permit sufficient radiative cooling to explain the cold anomaly in the stratosphere.

The cooling in the mesosphere during the initial warming phase is induced in the present run by a combination of the radiative cooling permitted in the absence of parameterized gravity wave drag and by a transient circulation induced by the resolved wave drag in the stratosphere below. This results in the disappearance of the stratopause at its climatological height. The stratopause subsequently reforms well above this height, initially as a result of planetary waves that reach the mesosphere, followed by nonorographic and subsequently orographic parameterized gravity waves. This sequence of events matches that described by other studies of elevated stratopause events (Siskind et al. 2010), though the transient calculations performed here allow us to clearly attribute the temperature anomalies to the different sources of wave driving. The descent of the stratopause is somewhat analogous to the descent of shear zones in the QBO in that the region of wave drag descends with the peak of the jet; there are, however, significant differences in details of the processes responsible. There appears to be some correspondence between the apparent rate of descent of the cold anomaly in the stratosphere and the warm anomaly in the mesosphere. The lower-stratospheric circulation evidently plays a role in determining the gravity wave flux

responsible for the mesospheric anomaly, but whether the roughly equal rates of descent are anything more than a coincidence remains unclear.

The anomaly in the lowermost stratosphere following the sudden warming that triggers the PJO event is, therefore, essential for understanding the evolution of the extratropical circulation response from the surface through to the mesosphere. The results of the present study, therefore, suggest that an a priori understanding of the time scale of these events would require an understanding of the vertical structure of radiative time scales and of the climatological diabatic heating. This latter is equivalent to the climatological structure of planetary-scale wave driving, the absence of which is felt during PJO events. To be fully satisfying, however, this perspective must also be complemented with an improved understanding of just why the wave driving collapses in the aftermath of these particular sudden warmings.

Acknowledgments. We thank Isla Simpson and Michael Sigmond for providing the CMAM data and Alan Plumb and Peter Haynes for helpful conversations. We also acknowledge the support of the Natural Sciences and Engineering Research Council and the Canadian Foundation for Climate and Atmospheric Sciences.

APPENDIX

Calculation of Eddy Flux Terms

The resolved EP flux is computed following Andrews et al. (1987):

$$\begin{aligned} F_{r,y} &= -\rho_0 \cos\phi \left(\overline{u'v'} - \partial_z \bar{u} \frac{\overline{v'\theta'}}{\partial_z \bar{\theta}} \right), \\ F_{r,z} &= \rho_0 \cos\phi \left[\left(f - \frac{1}{a \cos\phi} \frac{\partial \cos\phi \bar{u}}{\partial \phi} \right) \frac{\overline{v'\theta'}}{\partial_z \bar{\theta}} - \overline{u'w'} \right], \end{aligned}$$

where θ is the potential temperature. The fluxes are computed from 6-hourly output of CMAM data; daily averages of the fluxes are then computed. The wave-number decomposition to compute the planetary-scale fluxes is performed on the anomaly fields prior to multiplication and zonal averaging.

Parameterized gravity wave fluxes and tendencies are output directly by their respective parameterizations.

Zonal-mean quasi geostrophy on the sphere

Solving the set (1) for \bar{w}^* yields the differential operators of (3):

$$\begin{aligned} \mathcal{L}(\cdot) &= \frac{\partial}{\partial z} \left[\frac{1}{\rho_0} \frac{\partial(\rho_0 \cdot)}{\partial z} \right] + \frac{N^2}{(2\Omega a)^2} \frac{\partial}{\partial \mu} \left[\frac{1 - \mu^2}{\mu^2} \frac{\partial(\cdot)}{\partial \mu} \right], \\ \mathcal{L}_F(\cdot) &= \frac{1}{2\Omega a} \frac{\partial}{\partial \mu} \left[\frac{(1 - \mu^2)^{1/2}}{\mu} \frac{\partial(\cdot)}{\partial z} \right], \quad \text{and} \\ \mathcal{L}_Q(\cdot) &= \frac{1}{(2\Omega a)^2} \frac{\partial}{\partial \mu} \left[\frac{1 - \mu^2}{\mu^2} \frac{\partial(\cdot)}{\partial \mu} \right], \end{aligned}$$

where $\mu = \sin\phi$.

The lower boundary condition used is that of Haynes and Shepherd (1989) modified for the TEM. The assumption is that geometric velocity normal to the lower surface vanishes ($D\Phi/Dt = 0$ at $z = 0$). This yields the following condition for the residual circulation in log-pressure coordinates at the lower boundary

$$\begin{aligned} \rho_0 \frac{\partial \rho_0 \bar{w}^*}{\partial z} - \frac{g}{4\Omega^2 a^2} \frac{\partial}{\partial \mu} \left(\frac{1 - \mu^2}{\mu^2} \frac{\partial \bar{w}^*}{\partial \mu} \right) \\ = \frac{g}{4\Omega^2 a^3} \frac{\partial}{\partial \mu} \left[(1 - \mu^2)^{1/2} \frac{\overline{v'T'}}{S} \right]. \end{aligned}$$

Here g is the acceleration due to gravity. Since the lower boundary is not in fact a pressure surface in CMAM, we use meridional heat fluxes interpolated on to the 1000-hPa surface.

The method of solving this equation is as follows. The equation is separated following Plumb (1982), yielding the zonal-mean Laplace equation in the meridional direction, and restricting N^2 and α to be independent of latitude. This defines two Hermitian operators of interest, $\mathcal{M}^\dagger \mathcal{M}$ and $\mathcal{M} \mathcal{M}^\dagger$, where $\mathcal{M}(\cdot) = [(1 - \mu^2)/\mu^2]^{1/2} \partial_\mu(\cdot)$, each of which is associated with a family of eigenfunctions [μB_n and Θ_n in the notation of Plumb (1982)]. The meridional dependence of all quantities can then be expanded in terms of these two orthogonal sets. This expansion is carried out numerically and therefore requires that the computed eigenfunctions be orthonormal under the numerical integration operations; while this is straightforward to obtain for each set independently, finding two sets that are both simultaneously orthogonal proved to be difficult. Several initial attempts to compute these functions based on finite-difference approximations yielded sets of functions unusable for the required expansions. Computing these functions based on expansions in terms of Legendre polynomials yields a set that is adequately orthogonal through roughly the first half of the eigenfunctions. Since these describe the larger length scales expected to satisfy quasigeostrophic scaling, truncating the expansions is equivalent to filtering out some of the unbalanced component of the circulation. For all the calculations carried out here, the sums are truncated at mode 14.

The resulting vertical equation is then discretized on a grid with 0.5-km resolution. The differential operators become tridiagonal matrices that can be applied and inverted explicitly for every order n .

The temperature equation is integrated forward in time using a third-order Adams–Bashforth method. For the full adjustment problem, (3) is inverted at every time step.

REFERENCES

- Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics*. Academic Press, 489 pp.
- Baldwin, M. P., and T. J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science*, **294**, 581–584, doi:10.1126/science.1063315.
- Birner, T., 2006: Fine-scale structure of the extratropical tropopause region. *J. Geophys. Res.*, **111**, D04104, doi:10.1029/2005JD006301.
- Butchart, N., and Coauthors, 2011: Multimodel climate and variability of the stratosphere. *J. Geophys. Res.*, **116**, D05102, doi:10.1029/2010JD014995.
- Charlton, A. J., and L. M. Polvani, 2007: A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *J. Climate*, **20**, 449–469.
- Chen, P., and W. A. Robinson, 1992: Propagation of planetary waves between the troposphere and stratosphere. *J. Atmos. Sci.*, **49**, 2533–2545.
- Eliassen, A., 1951: Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astrophys. Norv.*, **5**, 19–60.
- Haynes, P. H., and T. G. Shepherd, 1989: The importance of surface pressure changes in the response of the atmosphere to zonally-symmetric thermal and mechanical forcing. *Quart. J. Roy. Meteor. Soc.*, **115**, 1181–1208.
- , C. J. Marks, M. E. McIntyre, T. G. Shepherd, and K. P. Shine, 1991: On the “downward control” of extratropical diabatic circulations by eddy-induced mean zonal forces. *J. Atmos. Sci.*, **48**, 651–678.
- Hitchcock, P., T. G. Shepherd, and S. Yoden, 2010: On the approximation of local and linear radiative damping in the middle atmosphere. *J. Atmos. Sci.*, **67**, 2070–2085.
- , —, and G. L. Manney, 2013: Statistical characterization of Arctic polar-night jet oscillation events. *J. Climate*, in press.
- Holton, J. R., 1983: The influence of gravity wave breaking on the general circulation of the middle atmosphere. *J. Atmos. Sci.*, **40**, 2497–2507.
- , and C. Mass, 1976: Stratospheric vacillation cycles. *J. Atmos. Sci.*, **33**, 2218–2225.
- Kuroda, Y., and K. Kodera, 2001: Variability of the polar night jet in the Northern and Southern Hemispheres. *J. Geophys. Res.*, **106** (D18), 20 703–20 713.
- , and —, 2004: Role of the Polar-night Jet Oscillation on the formation of the Arctic Oscillation in the Northern Hemisphere in winter. *J. Geophys. Res.*, **109**, D11112, doi:10.1029/2003JD004123.
- Limasuvan, V., J. H. Richter, Y. J. Orsolini, F. Stordal, and O.-K. Kvissel, 2012: The roles of planetary and gravity waves during a major stratospheric sudden warming as characterized in WACCM. *J. Atmos. Sol.-Terr. Phys.*, **78–79**, 84–98, doi:10.1016/j.jastp.2011.03.004.
- Manney, G. L., and Coauthors, 2008: The evolution of the stratospause during the 2006 major warming: Satellite data and assimilated meteorological analyses. *J. Geophys. Res.*, **113**, D11115, doi:10.1029/2007JD009097.
- Matsuno, T., 1971: A dynamical model of the stratospheric sudden warming. *J. Atmos. Sci.*, **28**, 1479–1494.
- McIntyre, M. E., 1980: An introduction to the generalized Lagrangian-mean description of wave, mean-flow interaction. *Pure Appl. Geophys.*, **118**, 152–176.
- Mukougawa, H., T. Hirooka, and Y. Kuroda, 2009: Influence of stratospheric circulation on the predictability of the tropospheric Northern Annular Mode. *Geophys. Res. Lett.*, **36**, L08814, doi:10.1029/2008GL037127.
- Plumb, R. A., 1982: Zonally symmetric Hough modes and meridional circulations in the middle atmosphere. *J. Atmos. Sci.*, **39**, 983–991.
- , and K. Semeniuk, 2003: Downward migration of extratropical zonal wind anomalies. *J. Geophys. Res.*, **108**, 4223, doi:10.1029/2002JD002773.
- Ren, S., S. M. Polavarapu, and T. G. Shepherd, 2008: Vertical propagation of information in a middle atmosphere data assimilation system by gravity-wave drag feedbacks. *Geophys. Res. Lett.*, **35**, L06804, doi:10.1029/2007GL032699.
- Richter, J. H., F. Sassi, and R. R. Garcia, 2010: Towards a physically based gravity wave source parameterization in a general circulation model. *J. Atmos. Sci.*, **67**, 136–156.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648.
- Rodgers, C. D., and C. D. Walsh, 1966: The computation of infra-red cooling rate in planetary atmospheres. *Quart. J. Roy. Meteor. Soc.*, **92**, 67–92, doi:10.1002/qj.49709239107.
- Scinocca, J. F., N. A. McFarlane, M. Lazare, J. Li, and D. Plummer, 2008: The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmos. Chem. Phys.*, **8**, 7055–7074, doi:10.5194/acp-8-7055-2008.
- Simpson, I. R., P. Hitchcock, T. G. Shepherd, and J. F. Scinocca, 2011: Stratospheric variability and tropospheric annular-mode timescales. *Geophys. Res. Lett.*, **38**, L20806, doi:10.1029/2011GL049304.
- Siskind, D. E., S. D. Eckermann, J. P. McCormack, L. Coy, K. W. Hoppel, and N. L. Baker, 2010: Case studies of the mesospheric response to recent minor, major, and extended stratospheric warmings. *J. Geophys. Res.*, **115**, D00N03, doi:10.1029/2010JD014114.
- Tomikawa, Y., K. Sato, S. Watanabe, Y. Kawatani, K. Miyazaki, T. Kinoshita, and M. Takahashi, 2012: Growth of planetary waves and the formation of an elevated stratosopause after a major stratospheric sudden warming in a T213L256 GCM. *J. Geophys. Res.*, **117**, D16101, doi:10.1029/2011JD017243.
- Yoden, S., T. Yamaga, S. Pawson, and U. Langematz, 1999: A composite analysis of the stratospheric sudden warmings simulated in a perpetual January integration of the Berlin TSM GCM. *J. Meteor. Soc. Japan*, **77**, 431–445.
- Zhou, S., A. J. Miller, J. Wang, and J. K. Angell, 2002: Downward-propagating temperature anomalies in the preconditioned polar stratosphere. *J. Climate*, **15**, 781–792.