The Response of the Lower Stratosphere to Zonally Symmetric Thermal and Mechanical Forcing

ALISON MING, PETER HITCHCOCK, AND PETER HAYNES

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom

(Manuscript received 14 September 2015, in final form 9 February 2016)

ABSTRACT

The response of the atmosphere to zonally symmetric applied heating and mechanical forcing is considered, allowing for the fact that the response may include a change in the wave force (or ''wave drag''). A scaling argument shows that an applied zonally symmetric heating is effective in driving a steady meridional circulation provided that the wave force (required to satisfy angular momentum constraints) is sufficiently sensitive to changes in the mean flow in the sense that the ratio $KN^2D_Q^2/\alpha f^2 L_Q^2$ is large, where K is a measure of the sensitivity of the wave force; α, N, and f are the radiative damping rate, buoyancy frequency, and Coriolis parameter, respectively; and L_Q and D_Q are the horizontal and vertical length scales of the heating, respectively. Furthermore, in the ''narrow heating'' regime where this ratio is large, the structure of the meridional circulation response is only weakly dependent on the details of the wave force. The scaling arguments are verified by experiments in a dry dynamical circulation model. Consistent with the scaling prediction, the regime does not apply when the width of the imposed heating is increased. The narrow-heating regime is demonstrated to be relevant to the double peak in tropical lower-stratospheric upwelling considered in a companion paper, supporting the hypothesis that this feature is radiatively driven. Similar arguments are applied to show that a narrow zonally symmetric applied mechanical forcing is primarily balanced by a change in wave force. This provides an explanation for the recently identified compensation between resolved and parameterized waves in driving modeled trends in the Brewer–Dobson circulation.

1. Introduction

The important role for wave forces (or ''wave drag'') in the dynamics of the mean meridional circulation (e.g., the Brewer–Dobson circulation in the stratosphere) is now well understood. In particular, for a steady circulation that crosses angular momentum contours (which are to good approximation vertical outside of the tropics), there must be a corresponding wave force and the relation between the wave force and the circulation is expressed by the downward control principle (Haynes et al. 1991). In the context of a specified wave force this excludes the possibility of a change in the meridional circulation as a result of a zonally symmetric applied heating. However, in practice, in an atmosphere with an active wave field, such a heating may change the wave force and hence may change the mean meridional circulation. The change in the wave force due to an applied heating needs to be taken into account in seeking a causal explanation for the magnitude and structure of the change in mean meridional circulation. Correspondingly, the change in wave force may need to be taken into account as part of the response to a zonally symmetric mechanical forcing. A specific example has been discussed by Cohen et al. (2014) in considering intermodel differences in the driving of changes in the stratospheric Brewer–Dobson circulation. The imposed mechanical forcing in this example represents a change (or an intermodel difference) in the parameterized gravity wave momentum flux. The wave force is due to the large-scale resolved waves that may respond to the imposed change.

A specific problem, which served as initial motivation for the work reported in this paper, is to explain the double peaks in upwelling and heating in the lower stratosphere in reanalysis datasets such as ERA-Interim and in general circulation models. A preceding companion paper (Ming et al. 2016, hereafter M16) reports a diagnostic study of the angular momentum and radiative

Corresponding author address: Alison Ming, DAMTP, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, United Kingdom.
E-mail: a.ming@damtp.cam.ac.uk

DOI: 10.1175/JAS-D-15-0294.1

balance associated with the double peaks. The conclusion of M16 was that the ERA-Interim estimates of the upwelling and the resolved wave force give a self-consistent angular momentum balance. However, consistency by itself does not establish the cause of the structure in upwelling. Consider a case where external radiative influences are essentially homogeneous in latitude. In the Newtonian cooling approximation, this is equivalent to assuming that the radiative relaxation temperature and damping rates are independent of latitude. A double peak in upwelling could arise through the wave dynamics if that organizes the wave force to be suitably confined away from the equator. This would require the temperature field to have a double-peak structure but this is not observed. The temperatures are in fact almost constant across a broad equatorial region (see M16, their Fig. 1c). An alternative is that the radiative relaxation temperature includes a double-peak structure. Then, as is observed, there can be a corresponding double peak in upwelling with the actual temperature constant latitude. It is unlikely that the required correspondence between the latitudinal structure of the radiative relaxation temperature field and that of the wave force would arise by chance. A more plausible explanation is that the latitudinal structure of the radiative relaxation temperature field, which from purely dynamical considerations at least is externally imposed, determines the latitudinal structure of the wave force.

It was also shown in M16, using offline radiative calculations, that the double peak in the diabatic heating rates observed in ERA-Interim arises primarily from the latitudinal structure in ozone, with contributions from both the longwave and shortwave heating with smaller contributions from the latitudinal variation in the temperature structure below and above the level of the double peak. Specifically, the difference in the clear-sky longwave radiative heating rates between 20°N and the equator at 70 hPa can be attributed to the latitudinal gradient in ozone (~70%) and to the latitudinal difference in the temperature profile (~20%), with the remainder resulting from latitudinal differences in water vapor. The fact that the double peak arises from factors other than the local temperature is analogous to the case with Newtonian cooling where there is externally imposed latitudinal structure in the radiative relaxation temperature field.

The aim of this paper is to investigate further the hypothesis that the double-peak structure in upwelling observed in ERA-Interim and in other reanalyses and models, and the required structure in wave force, is caused by radiative effects that can be regarded as external (in the sense that they do not arise from the temperature structure in the 70-hPa layer). A heuristic dynamical discussion is presented in section 2 and identifies a potentially relevant dynamical regime in which a suitably narrow imposed heating (the ''narrow heating'' regime) is primarily balanced by upwelling rather than by the radiative relaxation associated with a change in temperature. Section 3 describes and discusses an idealized 3D model experiment that verifies this response and section 4 describes an extended set of experiments that, by varying the parameters defining the imposed heating field, clarifies the conditions under which it occurs.

The findings of sections 2–4 are also relevant to understanding the response to an imposed mechanical forcing, rather than an imposed heating, when the response can include an adjustment in the wave force and this problem is discussed in section 5. In this case there is a dynamical regime in which a suitably narrow imposed mechanical forcing (the ''narrow force'' regime) is balanced by a change in wave force rather than by the Coriolis torque associated with a change in meridional circulation. This relates to the recent discussion of the Brewer–Dobson circulation by Cohen et al. (2013, 2014). Finally, section 6 contains a discussion of some of the main findings and gives some conclusions.

2. Dynamical considerations

For a dynamical discussion, we turn to the steady-state transformed Eulerian-mean equations in quasigeostrophic form in spherical and log-pressure coordinates (Plumb 1982):

$$-f\overline{v}^* = \frac{1}{\rho_0 a\cos\phi}\nabla\cdot\mathbf{F} = G[\overline{u}(\phi,\ z);\ \phi, z], \tag{1a}$$

$$f\partial_z\overline{u} + \frac{R}{H}\frac{1}{a}\partial_\phi\overline{T} = 0, \tag{1b}$$

$$\frac{1}{a\cos\phi}\partial_\phi(\overline{v}^*\cos\phi) + \frac{1}{\rho_0}\partial_z(\rho_0\overline{w}^*) = 0, \tag{1c}$$

$$\overline{S}\overline{w}^* = \overline{Q} - \alpha\overline{T}, \tag{1d}$$

where $\overline{v}^*$ and $\overline{w}^*$ are the mean residual meridional and vertical velocities, respectively; $\overline{u}$ is the zonal-mean zonal wind; f is the Coriolis parameter; a is the radius of Earth; ϕ is latitude; z is log-pressure height; $\rho_0 \propto \exp(-z/H)$, where H is a scale height taken to be 7 km; R is the gas constant for dry air; T is the temperature; $\overline{S} = \partial\overline{T}/\partial z + \kappa\overline{T}/H$ is a measure of the static stability; and $\kappa = R/c_p \simeq 2/7$, where c_p is the specific heat at constant pressure. The terms on the right-hand side of (1d) represent radiative heating, with $\overline{Q}$

envisaged as imposed. The notation is consistent with (1) in M16. In the specific context discussed in section 4 of M16, $\overline{Q}$ might be the heating in the lower stratosphere resulting from absorption of longwave and shortwave radiation by ozone and does not depend on local temperatures, whereas the $-\alpha\overline{T}$ term is a simple Newtonian cooling representation of the dependence of the diabatic heating field on local temperatures.

An important aspect of the Eliassen–Palm flux term on the right-hand side of (1a), which we will refer to as the wave force, is that it depends on the mean flow $\overline{u}$ because this affects the propagation and breaking of waves. We choose to emphasize this dependence and use the representation $G[\overline{u}(\phi, z); \phi, z]$. The dependence on $\overline{u}$ is nonlocal and, of course, the explicit form of G for the real atmosphere remains unknown.

We now proceed to analyze the response to a given imposed heating $\overline{Q}$. The closely related problem of the response of an imposed mechanical forcing, which will be included on the right-hand side of (1a) in addition to the wave force G, will be considered in section 5. Note that in the context of the response of a preexisting circulation to an imposed heating or mechanical forcing, the dependent variables must be interpreted as changes in the physical quantities. As a result of the imposed heating, all the dependent variables in (1) will change (including G, because of its dependence on $\overline{u}$). It is convenient to introduce ΔU as a typical magnitude of the response in $\overline{u}$ and ΔG as a typical magnitude of the response in G. For the moment we do not attempt to relate ΔU and ΔG. However, we assume that the horizontal and vertical scales of both these responses are approximately the same as the horizontal and vertical length scales of the imposed heating, taken to be L_Q and D_Q, respectively.

From (1a), it follows that $\overline{v}^* \sim \Delta G/f$ and from (1c) that $\overline{w}^* \sim \Delta G D_Q/(fL_Q)$. Using (1b), it follows that $\overline{T} \sim HL_Q f\Delta U/(RD_Q)$. The sizes of the two terms that can balance $\overline{Q}$ in (1d) are therefore $\overline{S}\Delta G D_Q/(fL_Q)$ (the vertical advection term) versus $\alpha HL_Q f\Delta U/(RD_Q)$ (the Newtonian cooling term).

Noting that $R\overline{S}/H$ is the square of the buoyancy frequency N, the relative sizes of the vertical advection term and the Newtonian cooling term, in (1d), are given by the quantity

$$\frac{\Delta G}{\Delta U}\frac{N^2D_Q^2}{\alpha f^2 L_Q^2} \sim K\frac{N^2D_Q^2}{\alpha f^2 L_Q^2}, \tag{2}$$

where K is a typical value of the ratio $\Delta G/\Delta U$ (i.e., of the sensitivity of the wave force G to the velocity).

In the case where the heating is applied close to the equator then f must be replaced by βL_Q in the above and the relative size is

$$\frac{\Delta G}{\Delta U}\frac{N^2D_Q^2}{\alpha\beta^2 L_Q^4} \sim K\frac{N^2D_Q^2}{\alpha\beta^2 L_Q^4}. \tag{3}$$

If the quantities appearing in (2) and (3) are large, then $\overline{Q}$ is primarily balanced by the term $\overline{S}\overline{w}^*$. They represent "dynamical aspect ratios" that determine whether, in an appropriate sense, the imposed heating is "narrow" or "broad." If the dynamical aspect ratio is large then the heating is narrow and the majority of the imposed heating is balanced by upwelling. (Note the immediate caveat that there is a constraint that the global average value of $\overline{w}^*$ on each z level, implying that the assumed balance cannot be perfect for any arbitrary $\overline{Q}$. We will return to this point in the next section.) If it is small, then the heating is broad, and the majority of the imposed heating is balanced by Newtonian cooling.

For given D_Q, L_Q, α, and K, the heating is "narrower" at low latitudes than at high latitudes in the sense that f is smaller and the aspect ratio is larger. The quantities in the (2) and (3) and the distinction between the broad and narrow response have been identified and discussed in many previous papers including Dickinson (1971), Fels et al. (1980), Garcia (1987), Plumb and Eluszkiewicz (1999), and Haynes (2005). These discussions generally make specific assumptions about the form of G—essentially, that it can be represented by a Rayleigh friction so that $G = -\kappa\overline{u}$. The Rayleigh damping coefficient κ then replaces K in (2) and (3). However, it is generally accepted that Rayleigh friction is a nonphysical and poor representation of the wave forces that operate in the stratosphere. The difference here is that we are taking K to be a rough quantitative description of a more general G that, as emphasized previously, is an unknown, nonlocal, and possibly very complicated function of $\overline{u}$. Evaluating K as the ratio $\Delta G/\Delta U$ in any precise way will therefore be difficult. However, in any given problem, provided that the applied heating perturbation is not too large, there will be some rough proportionality between the typical magnitude of change of $\overline{u}$ and of G and there will be some K that captures that proportionality. In this analysis, it is not necessary to know precisely how G varies with U. Provided K is large enough, the heating or at least the latitudinally varying part of the heating will be balanced by an upwelling.

We have noted previously in M16, from their Fig. 1c and sections 3 and 4, that the double-peak structure in heating is not matched by a corresponding structure in temperature. In the light of the scaling above, this

suggests that the double-peak structure is described by the narrow-heating regime. In the next section, we examine the response to a narrow imposed heating in idealized 3D model calculations.

3. Model calculations of response to a double-peak applied heating

a. Model description

We now describe detailed model simulations of the response to an applied heating. Bearing in mind the arguments in the previous section, we expect the change in the wave force will be an important part of the overall response. We use a 3D model in which the wave field, and hence the wave force, are free to vary. In particular we choose the well-known idealized system first defined by Held and Suarez (1994), in which there is a simple thermal relaxation to a specified temperature field T_{HS}. The relaxation state leads, under three-dimensional dynamics that incorporates longitudinal as well as latitudinal and vertical variations, to an active baroclinic eddy field and to a statistical equilibrium state that is maintained away from the thermal relaxation state by the action of the eddies, and that includes, for example, well-defined subtropical jets [see Held and Suarez (1994) for further details]. For numerical simulations we use the University of Reading Intermediate General Circulation Model, version 1 (IGCM 1), which is a hydrostatic primitive-equation model based on the original Hoskins and Simmons (1975) spectral dynamical model. A more detailed description is provided by Forster et al. (2000). The model is set up with the configuration described in Held and Suarez (1994) and with T42 resolution (approximately 4° × 4°) and 40 levels equally spaced in log-pressure coordinates with the model top at 19 hPa. Since there is no imposed inhomogeneity in longitude, there is no stationary planetary wave field. However, there are transient waves on a range of longitudinal scales (e.g., Held and Suarez 1994, their Fig. 4), including the synoptic and planetary scales. These provide the wave force in the subtropical lower stratosphere. The Eliassen–Palm fluxes in the Held–Suarez model integration are generally upward and equatorward and in both hemispheres; there are two main regions of convergence: one in the troposphere and one in the stratosphere as shown in Fig. 1. The broad region of convergence in the stratosphere has a maximum around latitude 40°N/S and 70 hPa with significant contributions from wavenumbers 1–7. Previous studies (e.g., Randel et al. 2008; Shepherd and McLandress 2011) have emphasized the range of wave

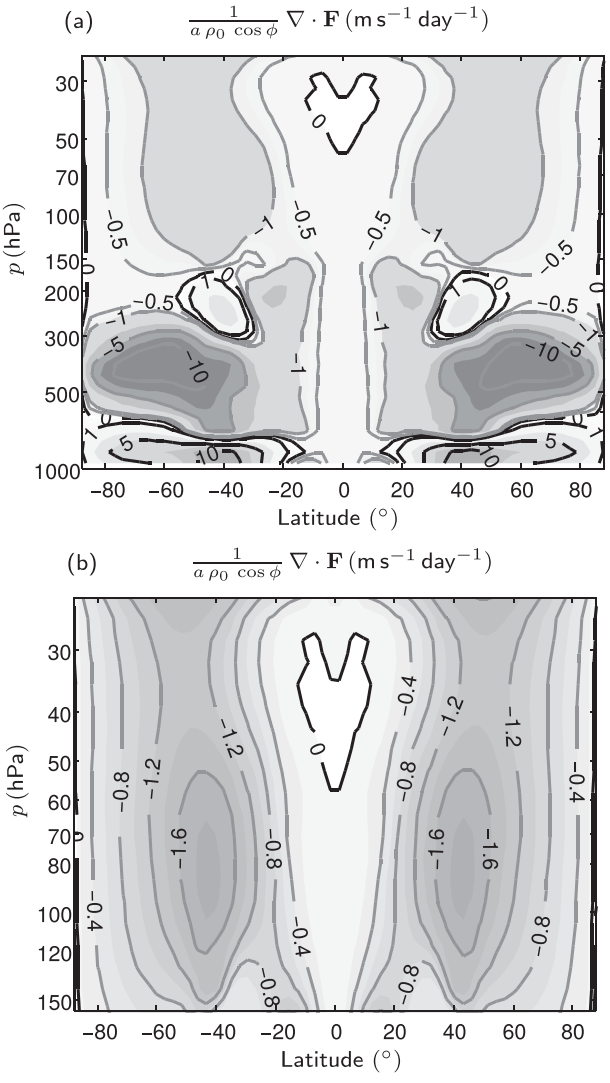

FIG. 1. (a) Wave force due to the divergence of the Eliassen–Palm flux in the idealized 3D model run (T42L40) with the Held and Suarez (1994) restoration state. (b) As in (a), but focusing on the stratosphere, which is of primary interest and shown in later figures.

types that give rise to the wave force in the subtropical lower stratosphere in the real atmosphere and in general circulation models, with no evidence that stationary planetary waves play a dominant role. On this basis, we regard the Held–Suarez configuration as suitable for the study reported here. We consider the sensitivity of the results to the presence of topographically generated stationary planetary waves in a later section.

The response to applied heating is taken as the difference between integrations with and without a specified steady heating perturbation added to the Held–Suarez configuration. In each case the model is

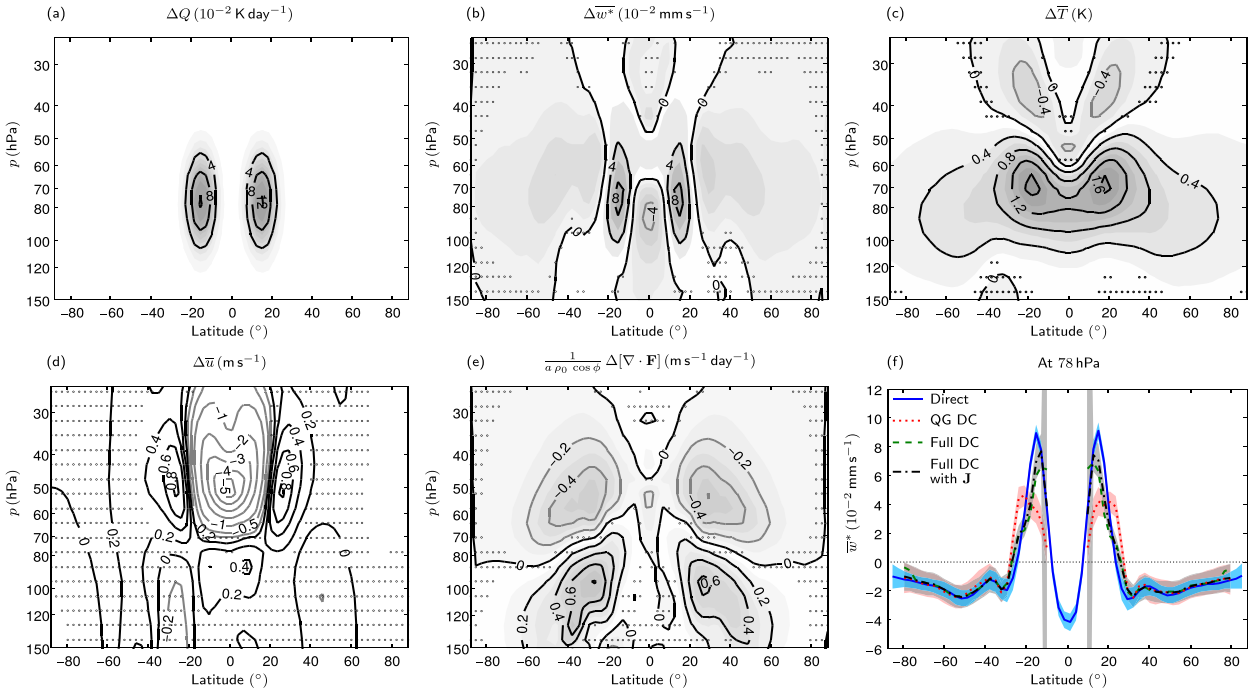

FIG. 2. (a) Imposed change in diabatic heating ΔQ ($10^{-2}\,\mathrm{K\,day^{-1}}$) in the idealized 3D model. (b) Change in $\overline{w}^*$ ($10^{-2}\,\mathrm{mm\,s^{-1}}$). (c) Change in $\overline{T}$ (K). (d) Change in $\overline{u}$ ($\mathrm{m\,s^{-1}}$). (e) Change in $\nabla\cdot\mathbf{F}/(a\rho_0\cos\phi)$ ($\mathrm{m\,s^{-1}\,day^{-1}}$). Differences in the stippled regions are not significant at the 95% confidence level as determined by an adjusted Student's t test. More details of the statistics can be found in the appendix. (f) $\overline{w}^*$ ($\mathrm{mm\,s^{-1}}$) at 78 hPa. The direct and three downward control (DC) calculations are shown. The DC calculations are done with the quasigeostrophic approximation, the full calculation using $\nabla\cdot\mathbf{F}$, and the full calculation using both $\nabla\cdot\mathbf{F}$ and J. The 95% confidence intervals are shown for all lines, except for the full calculation with $\nabla\cdot\mathbf{F}$ in the interest of clarity. The confidence interval for this line is similar to that for $\nabla\cdot\mathbf{F}$ and J together.

integrated from an initial state of rest for a total of 100 000 days, with the first 10 000 days discarded as model spinup. The long integrations are needed in order to allow the response to be clearly distinguishable from the internal dynamical variability. The heating perturbation, which is a function only of latitude and height, is added to the right-hand side of the thermodynamic equation to give

$$\frac{\partial T}{\partial t} = \cdots - \alpha(\phi,\sigma)[T - T_{\mathrm{HS}}(\phi,\sigma)] + \Delta Q(\phi,z), \qquad (4)$$

$$\Delta Q = \Delta Q_d^+ + \Delta Q_d^-, \qquad (5)$$

$$\Delta Q_d^{\pm} = Q_A \exp\left[-\frac{(\phi \pm \phi_{\max})^2}{2\Delta\phi^2} - \frac{(z - z_{\max})^2}{2\Delta z^2}\right],$$

where $z = -H\log\sigma$, $H = 7\,\mathrm{km}$, and $\sigma = p/p_s$ (p is the pressure and p_s is the surface pressure; when ΔQ is computed offline, p_s is set to 1000 hPa).

The first term on the right-hand side of (4) is used in the Held–Suarez configuration with a Newtonian cooling term proportional to the difference between the actual temperature and the temperature specified by the thermal relaxation state. The value of the Newtonian cooling coefficient, outside of a shallow boundary layer, is $0.025\,\mathrm{day^{-1}}$. The term ΔQ is the specified heating perturbation.

In the first integration to be reported, referred to as the standard configuration,

$$Q_A = 0.125\,\mathrm{K\,day^{-1}}, \quad \phi_{\max} = 15°, \quad z_{\max} = 18\,\mathrm{km},$$
$$\Delta\phi = 5°, \quad \Delta z = 1.5\,\mathrm{km}.$$

The amplitude and latitude–height structure of the heating perturbation are here chosen to make it a simple representation of the double-peak structure in the ERA-Interim heating rates.

b. Response to imposed heating

Figures 2a–e show the imposed diabatic heating perturbation ΔQ and the resulting changes in upwelling, temperature, zonal wind, and divergence of the Eliassen–Palm flux, respectively. The stippling shows regions where the change is not significant at the 95% confidence level using an adjusted Student's t test (see appendix). The response in the upwelling has a spatial

structure that broadly resembles the imposed heating. The upwelling term $\overline{S}\,\overline{w}^*$ in the thermodynamic equation balances the applied heating to good approximation as will be demonstrated more quantitatively below. The upwelling in the regions of the applied heating is part of a larger pattern in meridional circulation with a double cell in each hemisphere. The downwelling outside of the regions of applied heating causes a temperature increase. The overall result is that the temperature response at 78 hPa (peak in heating) is much broader than the imposed diabatic heating and to a first approximation flat over the equator. The existence of meridional circulation cells in the response requires a nontrivial balance in the zonal momentum equation. The Coriolis force associated with the meridional flow is balanced by a change in $\nabla \cdot \mathbf{F}$. The heating perturbation causes an upward and equatorward shift in the region of convergence of $\mathbf{F}$. The strongest changes in $\nabla \cdot \mathbf{F}$ are located poleward of the heating region, with negative changes above and positive changes below the imposed heating. There is also a corresponding change in the zonal wind, primarily in the region above the heating, with a large decrease over the equator and two regions on either side of this where $\overline{u}$ increases. Note that the change in the wave force has a significantly different spatial structure to the change in the zonal wind; indeed, there are regions (e.g., 20°N, 100 hPa) where the wave force changes substantially despite a negligible change in $\overline{u}$.

To demonstrate that there is a consistent angular momentum balance in the response, we calculate the upwelling using the downward control integral (see M16 for details). Figure 2f shows that upwelling from the model is in agreement with that inferred from the quasigeostrophic and full downward control equations at 78 hPa. The solid line shows the upwelling calculated from the wind and temperature response in the model. The downward control integral is calculated using the time-averaged changes in $\nabla \cdot \mathbf{F}$ and $\overline{u}$. The upwelling estimated from the wave torque has a double-peak structure that is similar to the one obtained from the direct calculation from the wind fields in the model. In the extratropics, both the quasigeostrophic and the full downward control agree. In the tropical region between 25°S and 25°N, the quasigeostrophic approximation to the integral is unable to capture the precise structure and magnitude of the upwelling and the change in the angular momentum contours must be taken into account. In Fig. 2f, we also show the effect of the J term (Scott 2002), which represents the time-averaged angular momentum advection terms due to the time varying part of the flow (see also M16, their section 2). Unlike the ERA-Interim data, the model runs are sufficiently long to allow the inclusion of this term. While this term does not make a significant difference for this particular heating, it becomes more important the closer to the equator the heating is located.

The response to applied heating found here is in broad agreement with the combination of double peak in upwelling and flat temperature structure observed in ERA-Interim, previously described in section 3 of M16.

Now consider the changes in the various terms in the (time averaged) thermodynamic equation in (1d) when a heating ΔQ is applied, with $\Delta(\cdot)$ indicating the change in the quantity $(\cdot)$:

$$\Delta(\overline{S}\,\overline{w}^*) = \Delta\overline{S}\,\overline{w}^*_0 + \overline{S}_0\Delta\overline{w}^* + \Delta\overline{S}\Delta\overline{w}^* = \Delta Q - \alpha\Delta\overline{T}, \quad (6)$$

where $\overline{w}^*_0$ and $\overline{S}_0$ are the time-averaged mean vertical residual velocity and static stability, respectively, obtained with the control Held–Suarez configuration. Note that unlike in the QG equations (1) discussed in section 2, $\overline{S}$ is allowed to be a function of latitude in (6), and $\overline{w}^*_0$ and $\overline{S}_0$ have been included. Equation (6) indicates that the heating perturbation may be balanced by a combination of change in vertical velocity, change in temperature, and change in static stability, with the latter two quantities being closely related, since $\Delta\overline{S} = \partial_z\Delta\overline{T} + \kappa\Delta\overline{T}/H$.

The impression from Figs. 2a and 2b is the balance comes primarily from the change of vertical velocity; however, this has to be consistent with the constraint that $\langle\overline{w}^*\rangle = 0$, where the angle brackets denotes the global area average on a z level.

Assuming that the nonlinear term in (6) is small, we divide throughout by $\overline{S}_0$ and take the global average to give

$$\langle\Delta\overline{w}^*\rangle = \left\langle\frac{\Delta Q}{\overline{S}_0}\right\rangle - \alpha\left\langle\frac{\Delta\overline{T}}{\overline{S}_0}\right\rangle - \left\langle\frac{\Delta\overline{S}\overline{w}^*_0}{\overline{S}_0}\right\rangle = 0. \quad (7)$$

It follows that if $\langle\Delta Q/\overline{S}_0\rangle \neq 0$. then some combination of $\Delta\overline{T}$ and $\Delta\overline{S}$ must be nonzero. Conversely, a heating can be balanced purely by $\Delta\overline{w}^*$ only if $\langle\Delta Q/\overline{S}_0\rangle = 0$. Therefore, in assessing the balance in (6), it is more appropriate to consider the correspondence between $\overline{S}_0\Delta\overline{w}^*$ and $\Delta\tilde{Q} = \Delta Q - \overline{S}_0\langle\Delta Q/\overline{S}_0\rangle$ rather than ΔQ itself. We therefore choose to partition ΔQ into a component that can potentially be balanced by an upwelling [neglecting nonlinear terms in (6)] $\Delta\tilde{Q}$ and a component that cannot $\overline{S}_0\langle\Delta Q/\overline{S}_0\rangle$. Note that the latter has the latitudinal structure of the control state $\overline{S}_0$ [see Fueglistaler et al. (2011) for further discussion including implications for the annual cycle]. In the Held–Suarez control state, $\overline{S}_0$ has a global-mean value of about 1×10^{-3} K m^{-1} in the

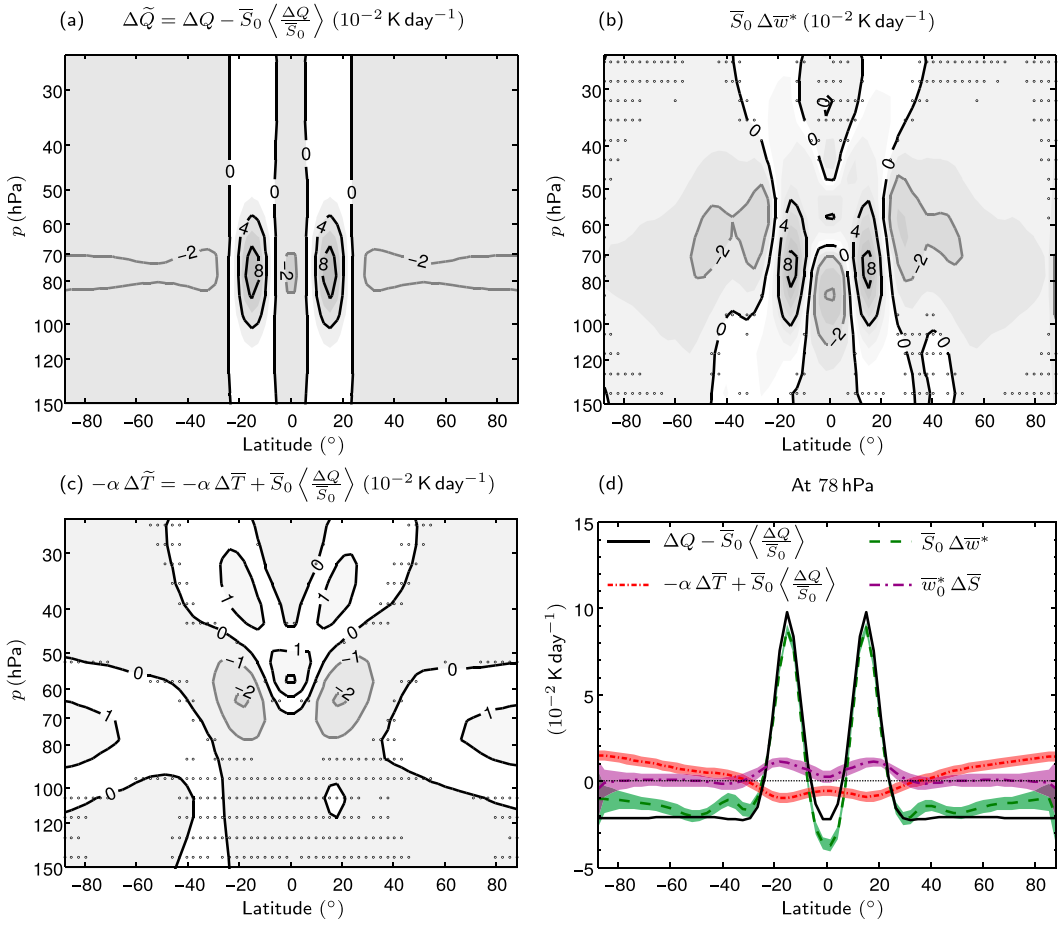

FIG. 3. (a) Heating perturbation with a global mean removed ($\Delta\tilde{Q} = \Delta Q - \overline{S}_0\langle\Delta Q/\overline{S}_0\rangle$) for the case shown in Fig. 2a for the idealized 3D model; (b) $\overline{S}_0\Delta\overline{w}^*$; (c) $-\alpha\Delta\tilde{T} = -\alpha\Delta\overline{T} + \overline{S}_0\langle\Delta Q/\overline{S}_0\rangle$; and (d) the terms in (a)–(c) at 78 hPa. All plots are in units of 10^{-2} K day^{-1}. Differences in the stippled regions are not significant at the 95% confidence level as determined by an adjusted Student's t test. A similar calculation produces the 95% confidence intervals shown in (d).

lower stratosphere with a difference between the tropics and extratropics of about 25%. This latitudinal structure in $\overline{S}_0$ is taken into account in $\Delta\tilde{Q}$. The difference between ΔQ and $\Delta\tilde{Q}$ may be incorporated into the Newtonian cooling term as $-\alpha\Delta\tilde{T} = -\alpha\Delta\overline{T} + \overline{S}_0\langle\Delta Q/\overline{S}_0\rangle$. The resulting budget, with the nonlinear term included, is then

$$\Delta\overline{S}\,\overline{w}_0^* + \overline{S}_0\Delta\overline{w}^* + \Delta\overline{S}\Delta\overline{w}^* = \Delta\tilde{Q} - \alpha\Delta\tilde{T}. \qquad (8)$$

An alternative approach that might have some advantages would be as follows. The changes in temperature and static stability are related. The term $\overline{S}_0\langle\Delta Q/\overline{S}_0\rangle$ is actually balanced by changes in temperature ΔT_e and in static stability ΔS_e and satisfies the equation $\Delta\overline{S}_e[\overline{w}_0^* + \Delta\overline{w}^*] + \alpha\Delta\overline{T}_e = \overline{S}_0\langle\Delta Q/\overline{S}_0\rangle$, where the nonlinear term is included. This can be solved to find the temperature response ΔT_e and static stability ΔS_e that would result from the heating $\overline{S}_0\langle\Delta Q/\overline{S}_0\rangle$. Then $\Delta\tilde{T}$ could be defined by $\Delta\overline{T} - \Delta\overline{T}_e$ and $\Delta\tilde{S}$ by $\Delta\overline{S} - \Delta\overline{S}_e$. Our approach is equivalent to setting $\Delta S_e = 0$.

Figures 3a and 3b show respectively the quantities $\Delta\tilde{Q}$ and $\overline{S}_0\Delta\overline{w}^*$ with evident quantitative agreement between these two terms in the region the heating perturbation is applied. Figure 3c shows the sum $-\alpha\Delta\tilde{T}$, which would be the difference between the quantities in Figs. 3a and 3b if the term $\overline{w}_0^*\Delta\overline{S}$ were negligible. In fact, the latter tends to oppose $-\alpha\Delta\tilde{T}$ in the region where the heating perturbation is applied. This is clear from Fig. 3d, which shows the latitudinal structure at 78 hPa in all of the various quantities and confirms the very close agreement between $\Delta\tilde{Q}$ and $\overline{S}_0\Delta\overline{w}^*$.

The balance in the thermodynamic equation in (6) can be assessed fully only by considering the height–latitude variation of the various quantities shown in Figs. 3a–c. However, it is convenient to find a simple

quantitative measure of the extent to which the applied heating is balanced by a response in upwelling. Here, we will use the ratio $\overline{S}_0\Delta\overline{w}^*/\Delta\tilde{Q}$, evaluated at the center of one of the double peaks in applied heating. This ratio has a value of 1.04 ± 0.04 in the case reported above.

c. *Zonally symmetric simulations with Rayleigh friction*

To illustrate the importance of the sensitivity of the wave field K, we provide some results for the case where the wave force G is represented by Rayleigh friction. This was the basis for many of the earlier studies of the driving of the mean meridional circulation (e.g., Dickinson 1971; Fels et al. 1980; Garcia 1987) and incorporates the u dependence of G, but in a highly simplified manner that is not believed to be realistic. Analytical progress using Hough functions (e.g., Garcia 1987) is possible under certain simplifying assumptions, which essentially require weak departures from a latitudinally independent state. Here it is most convenient simply to calculate the response using a zonally symmetric version (2D) of the full idealized 3D model described previously. This allows incorporation of latitude-dependent temperature structure, for example.

The zonally symmetric model is relaxed toward the Held–Suarez configuration. Rayleigh friction with a constant friction coefficient κ is added throughout the whole domain. All dynamical fields are constrained to be zonally symmetric. This means, for example, that there is no baroclinic instability. The first 1000 days of the model run are discarded as spinup and the model state after this is analyzed.

A set of experiments were performed in the zonally symmetric model with values of $1/\kappa$ of 1, 10, and 30 days. The applied heating field, defined by (4), again with the peaks located at latitudes of 15°N/S, is shown in Fig. 2a. The corresponding changes in temperature, zonal wind, and vertical wind are shown in Fig. 4.

For large values of κ ($1/\kappa = 1$ and 10 days), the system adjusts such that the dominant balance in the thermodynamic equation is between the upwelling and the heating (cf. Fig. 2a and 4c,f). Given the vertical wind response, the continuity equation implies the required change in $\overline{v}$; the balance between the Coriolis torque and the Rayleigh drag in the momentum equation sets the structure of $\overline{u}$ and, finally, the temperature change is related to the wind change by thermal wind balance. Note that within this large κ regime, while the vertical velocity (and hence the latitudinal velocity) change very little as κ varies, at least in the region of the applied heating, there is substantial change in $\overline{u}$, which has a magnitude inversely proportional to κ and a horizontal structure that is broad for large κ and narrows as κ decreases.

We can think of κ as setting the width of some region over which the circulation generated by the heating can spread. As κ is decreased, this region becomes smaller and the temperature and zonal wind changes become increasingly confined to the region of heating. This can also be seen in the location of the regions of downwelling in the right column of Fig. 4. More of the heating is balanced by a temperature change and less of it by the circulation change.

For a weaker Rayleigh drag ($1/\kappa = 30$ days), there are clear quantitative differences between the applied heating Fig. 2a and the vertical upwelling (Fig. 4i). (Note, for example, that the magnitude of the equatorial downwelling relative to the subtropical upwelling is much larger than for smaller values of $1/\kappa$.) It follows that the Newtonian cooling term is an important part of the balance in the thermodynamic equation. For reference, the value of the dynamical aspect ratio (2), setting $K = \kappa$, for the three values of $1/\kappa$: 1, 10, and 30 days is respectively 20, 2, and 0.7, confirming that the scaling arguments given previously are consistent with the calculated response with the ''narrow regime'' no longer applying when κ is sufficiently small.

These results for the zonally symmetric dynamics with Rayleigh friction may be compared with the full three-dimensional dynamical response shown in Figs. 2b–d. The three-dimensional response is very similar to that with large Rayleigh friction in that there is a balance between applied heating and upwelling, but the temperature and zonal wind responses are quite different. In this dynamical regime, the change in zonal wind (and, hence, through thermal wind balance, the change in temperature) is determined by the requirement that there is balance in the momentum equation and this change therefore depends on the details of the function G. The same point has been noted above with respect to different values of the Rayleigh friction coefficient. In other words, provided that G is sufficiently sensitive to $\overline{u}$, the response of the meridional circulation is robust and insensitive to the precise form of G, but the changes in zonal wind and temperature are not.

Note in particular that for the Rayleigh friction case the change in force has, by definition, the same shape in the latitude–height plane as the change in velocity since they are proportional to each other. In contrast, this correspondence does not hold for the three-dimensional case where the change in the acceleration

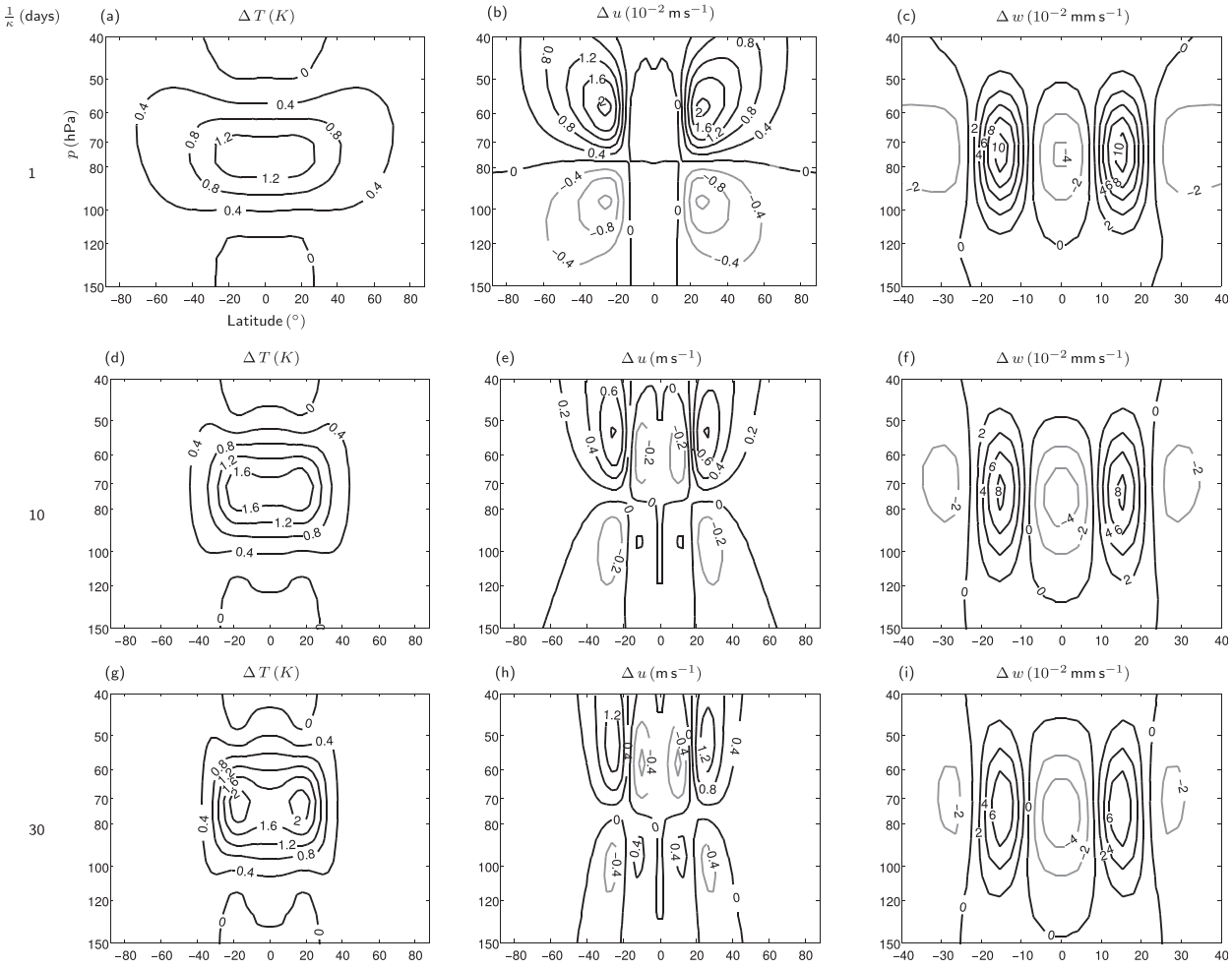

FIG. 4. (a),(d),(g) Change in T (K) for the imposed heating from Fig. 2a in a zonally symmetric 2D model with Rayleigh drag for values of $1/\kappa$ equal to 1, 10, and 30 days, respectively. (b),(e),(h) Change in u (note that the units of u change down the column). (c),(f),(i) Change in w (10^{-2} mm s^{-1}).

due to the waves (Fig. 2e) differs from the zonal wind change (Fig. 2d).

4. Model response to different types of applied heating

Having established a dynamical regime in which the applied heating is balanced by upwelling and the wave force adjusts to balance the corresponding Coriolis torque, we now investigate the implications of variations in the latitude, width, and strength of the heating. The next subsections report details of a set of experiments in the 3D model in which these quantities are varied. The parameters for this set of experiments are listed in Table 1. Whereas the standard case (**marked by an asterisk**) was motivated specifically by the radiative calculations in M16, the wider set of experiments are intended to explore the range of dynamical behavior rather than to model specific aspects of the real atmosphere.

a. Varying the latitude of the maximum heating (group A)

Figure 5 shows the temperature, upwelling, and divergence of the Eliassen–Palm flux response to a heating perturbation located at 10°, 20°, 40°, and 60°N/S (group A in Table 1). It is worth noting that although the experiment is set up such that the response should have hemispheric symmetry in the statistically steady state, there are some asymmetries that are especially prominent in $\overline{u}$ and $\nabla \cdot \mathbf{F}/(a\rho_0 \cos\phi)$. These hemispheric differences could be due to long time scales in the tropospheric jets that are known to occur in some configurations of dynamical cores such as the one we are using. They are unlikely to be a signal directly associated with the imposed diabatic heating in the lower

TABLE 1. Parameter values used in (4) for three sets of experiments that test the response of the idealized 3D model when the latitude (group A), strength (group B), or width (group C) of the applied heating is changed. The standard case (boldface), discussed in section 3, is repeated for completeness.

Group	Q_A (K day^{-1})	$\phi_{\max}$ (°)	$\Delta\phi$ (°)
A	0.125	10	5
	0.125	**15**	**5**
	0.125	20	5
	0.125	30	5
	0.125	40	5
	0.125	50	5
	0.125	60	5
B	0.025	15	5
	0.125	**15**	**5**
	0.250	15	5
	0.500	15	5
	1.000	15	5
C	**0.125**	**15**	**5**
	0.125	15	10
	0.125	15	20
	0.125	60	5
	0.125	60	10
	0.125	60	20
	0.125	60	40

stratosphere. The statistical test used picks out changes that are statistically different to zero at the 95% confidence level, not how different the values are between hemispheres. A further statistical test shows at the 95% confidence level that the response is not different between the hemispheres (not shown).

As the applied heating is moved away from the equator, the dominant structure of the upwelling response continues to match that of the applied heating, even when the latter is located in the extratropics. Further quantitative detail is given by Fig. 6, which shows the ratio $\overline{S}_0\Delta\overline{w}^*/\Delta\tilde{Q}$ evaluated at the center of one of the double peaks in applied heating. For this particular shape of applied heating, the ratio remains close to one for all latitudes plotted. While this result might initially seem surprising, bearing in mind the f dependence implied by (2), it is consistent with (2) provided that K is sufficiently large; that is, the wave force is sufficiently sensitive to the zonal velocity. Indeed, the same behavior is observed in the Rayleigh drag case for large enough values of κ as may be seen from the additional gray curves in Fig. 6. These were obtained through a calculation involving the use of Hough function expansions (e.g., Garcia 1987). The numerical calculation requires a large number of Hough functions to converge close to the equator and adequate convergence was obtained using the first 1000 eigenfunctions. This method is a simplification of the Rayleigh drag calculation performed using the model (described in section 3c) since the buoyancy frequency is set to a constant value typical of the stratosphere in the Held–Suarez case ($N^2 = 3.35 \times 10^{-4}\,\mathrm{s}^{-2}$). However, it provides a useful indication of the behavior of the ratio $\overline{S}_0\Delta\overline{w}^*/\Delta\tilde{Q}$ as κ is increased. The black dotted–dashed curve from the full model further shows that the contribution from changes in static stability that give rise to the term $\overline{w}^*_0\Delta\overline{S}/\Delta\tilde{Q}$ is small but appears to be responsible for the weak departure at low latitudes from the regime where $\overline{S}_0\Delta\overline{w}^*$ balances $\Delta\tilde{Q}$.

b. *Varying the strength of the heating (group B)*

With the peak heating anomalies centered at 15°N/S, the strength of the heating is varied from 0.025 to 1 K day^{-1} (group B in Table 1). Typical heating perturbations that are observed in reanalysis diabatic heating rates are about 0.3 K day^{-1} (see M16, their Fig. 4). The response to the heating in all the cases is similar in structure with a near-linear relationship between the maximum heating Q_A and amplitude of the response, even for large values of the heating perturbation up to 0.5 K day^{-1} (Fig. 7a).

The ratio of the imposed heating to the upwelling term (Fig. 7b) is close to 1, which is consistent with the fact that the dynamical aspect ratio (2) does not include a dependence on magnitude of the heating. For small Q_A, the statistical uncertainty in the ratio is large and, while there is a hint that the ratio may be smaller for very small Q_A, it is not clear that this would stand up to further scrutiny. For larger values of the heating perturbation (amplitude of 1 K day^{-1}), the behavior starts to become nonlinear, because the term $\Delta\overline{S}\Delta\overline{w}^*$ increases as shown in Fig. 7b.

The conclusion from this set of experiments is that for heating amplitudes less than 0.5 K day^{-1}, the response is essentially linear and the upwelling consistently provides the dominant balance to the applied heating (for this particular heating structure). Larger amplitudes of applied heating give rise to nonlinear effects but are not likely to be relevant to the real tropical lower stratosphere.

c. *Varying the width of the heating (group C)*

The experiments in group C (Table 1) address the change in response as the width of the double peaks is increased. As shown by Fig. 8, for cases with $\phi_{\max}$ = 15°N/S and $\phi_{\max}$ = 60°N/S, as the width of the perturbation is increased, the upwelling term no longer provides the dominant balance to the applied heating and other terms make comparable contributions. For $\phi_{\max}$ = 15°N/S, the change in the static stability term $\overline{w}^*_0\Delta\overline{S}$ becomes increasingly important as the width increases. For $\phi_{\max}$ = 60°N/S both $\overline{w}^*_0\Delta\overline{S}$ and $-\alpha\Delta\tilde{T}$ become important. The fact that these two terms tend to cancel

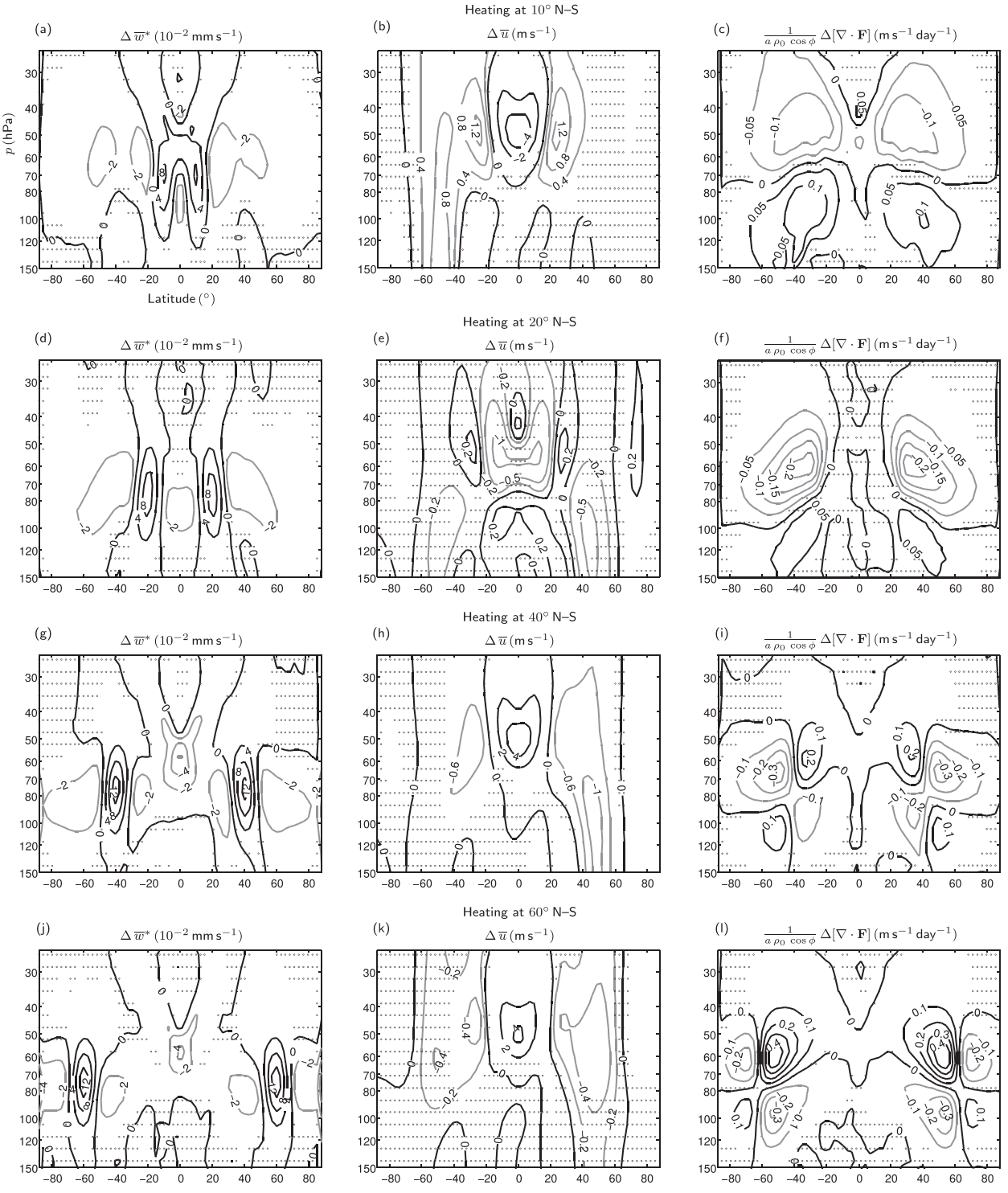

FIG. 5. Changes in (left) $\overline{w}^*$ (10^{-2} mm s^{-1}), (center) $\overline{u}$ (m s^{-1}), and (right) $\nabla \cdot \mathbf{F}/(a\rho_0 \cos\phi)$ (m s^{-1} day^{-1}) due to an imposed change in diabatic heating with peaks at (a)–(c) 10°, (d)–(f) 20°, (g)–(i) 40°, and (j)–(l) 60°N/S in the idealized 3D model (group A in Table 1). Differences in the stippled regions are not significant at the 95% confidence level as determined by an adjusted Student's t test.

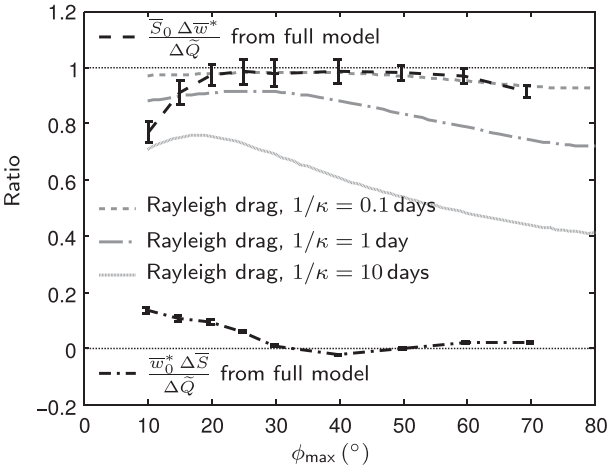

FIG. 6. Plots of the ratios $\overline{S}_0\Delta\overline{w}^*/\Delta\tilde{Q}$ and $\overline{w}_0^*\Delta\overline{S}/\Delta\tilde{Q}$ against the latitude of the maximum heating perturbation from the idealized 3D model runs with $Q_A = 0.125\,\mathrm{K\,day^{-1}}$. The ratios are calculated at the location of the maximum in heating in the Northern Hemisphere. The 95% confidence intervals are calculated using an adjusted Student's t test. The ratio $\overline{S}_0\Delta\overline{w}^*/\Delta\tilde{Q}$ is also plotted for the Rayleigh drag cases with κ equal to 10, 1, and 0.1 $\mathrm{day^{-1}}$ from the Hough function calculation.

at the center of the applied heating means that $\overline{S}_0\Delta\overline{w}^*/\Delta\tilde{Q}$ evaluated at the center of the applied heating decreases relatively little with width. However, inspection of the height–latitude structure of the response in various fields confirms that, as width increases, there is a clear transition from a response in which upwelling essentially matches heating, as seen in the cases in Fig. 5 for narrow heating, to a more complicated response.

This is illustrated by the response to a broad heating, for the case ($\phi_{\max} = 15°$, $Q_A = 0.125\,\mathrm{K\,day^{-1}}$, $\Delta\phi = 20°$), which is shown in Fig. 9. Note in particular that the region of upwelling (Fig. 9b) is much narrower than the applied heating (Fig. 9a) and furthermore that the maximum in upwelling is shifted significantly upward relative to the region of heating. Figure 9f summarizes the role of the different terms in the thermodynamic equation at 78 hPa. While some of the heating is balanced by an upwelling, the static stability change becomes equally as important in the tropical region with the term $\overline{w}_0^*\Delta\overline{S}$ becoming significant as shown in Fig. 9d. The $\Delta\overline{w}^*\Delta\overline{S}$ term remains small (Fig. 9e).

d. *Idealized orography*

The standard Held–Suarez configuration is a convenient vehicle for a first exploration of the response of a system with synoptic- and planetary-scale eddies to applied heating in the tropical lower stratosphere. However, it is not defensible as an accurate quantitative model of the troposphere–stratosphere system and it is important to establish whether the results reported so far are robust to changes in this idealized configuration. As a first step, a crude representation of orography is added to the Northern Hemisphere in the standard model run. A wave-1 perturbation is added to the surface geopotential height as a sine wave in longitude and a half sine wave in latitude between 25° and 45°N with an amplitude of 500 m. Figure 10 shows the same quantities as Figs. 2b–e but for the case with orography. Comparing Figs. 10a and 5b, we find that the upwelling response is qualitatively similar in the case with and without orography, again confirming that in this regime of "narrow heating," the response is not sensitive to the details of the wave field. The ratio $\overline{S}_0\Delta\overline{w}^*/\Delta\tilde{Q}$ in the case

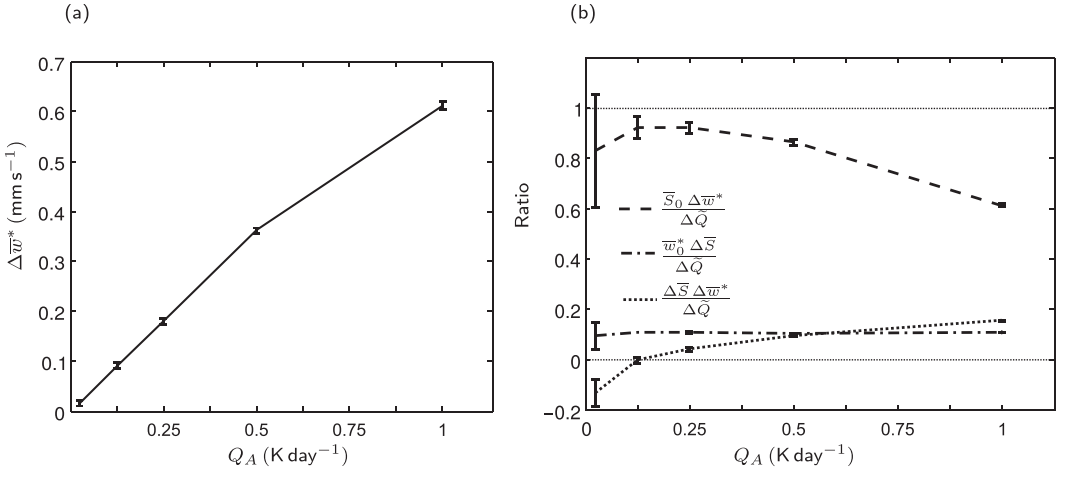

FIG. 7. (a) Change in upwelling against the magnitude of the heating perturbation Q_A for the idealized 3D model (group B in Table 1). (b) Ratio of various quantities in the thermodynamic equation in (6) to the heating plotted against the magnitude of the heating perturbation when the double peaks in heating are centered at 15°N/S. The 95% confidence intervals are shown although the intervals are too small to be clearly seen in some cases.

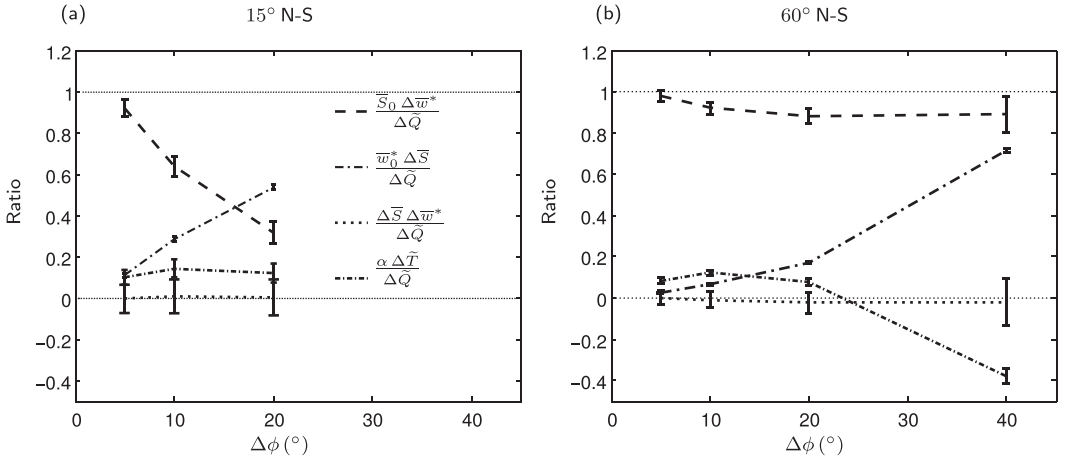

FIG. 8. Plots of the ratios of the various terms in (6) to the diabatic heating against the width of the double peaks in heating when they are centered at (a) 15° and (b) 60°N/S for the idealized 3D model (group C in Table 1). The 95% confidence intervals are shown. The legend for the two plots is the same.

with orography is still close to 1 and has a value of 1.2 ± 0.05.

Since both hemispheres are in the appropriate dynamical regime where the wave force is sufficiently sensitive to the zonal velocity that this applied heating is narrow according to (2), the dominant balance in the thermodynamic equation is that the upwelling balances the applied heating. Therefore, the upwelling response, in the region of the applied heating, is relatively insensitive to the details of the wave force and, as predicted by the scaling arguments given above, is similar between the two hemispheres. On the other hand, the

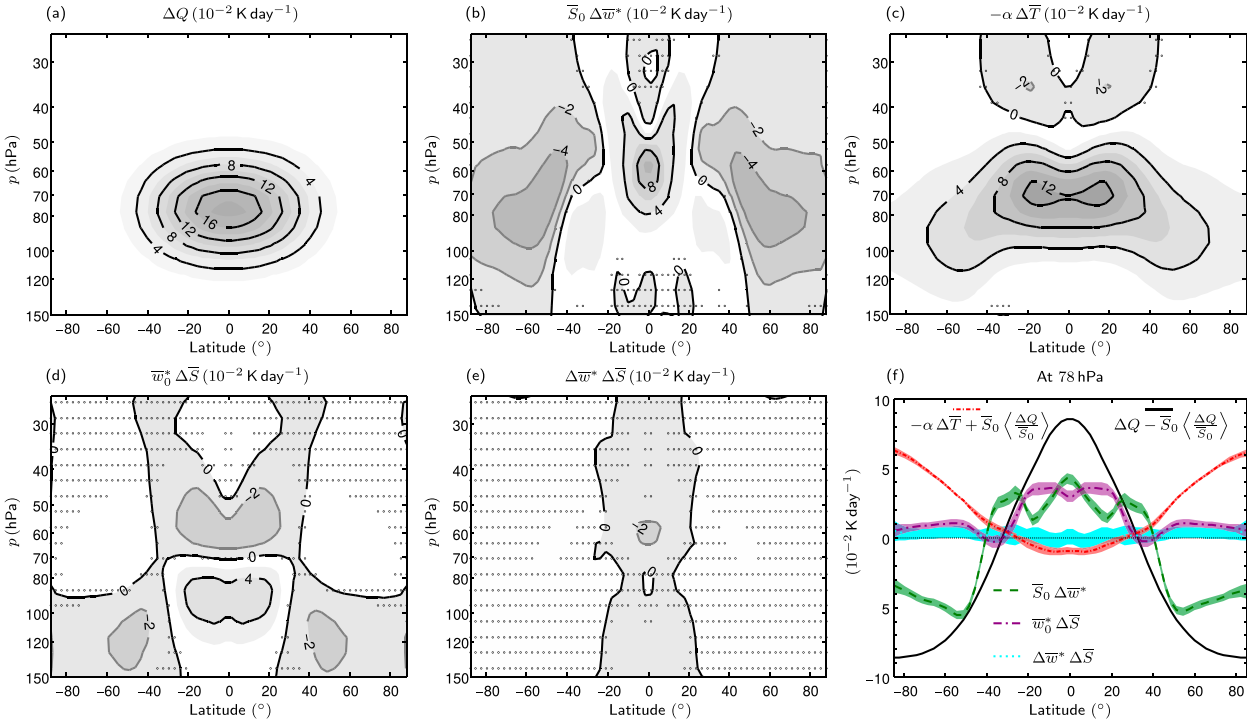

FIG. 9. Plots of the various terms in the thermodynamic equation for the idealized 3D model experiment where the heating is at 15° with amplitude 0.125 K day^{-1} and width 20°. (a) Change in imposed heating. (b) Change in $\overline{S}_0\Delta\overline{w}^*$. (c) Change in $-\alpha\Delta\overline{T}$. (d) Change in $\overline{w}^*_0\Delta\overline{S}$. (e) Change in $\Delta\overline{w}^*\Delta\overline{S}$. Stippled regions are not statistically significant at the 95% level. (f) The same terms plotted at 78 hPa. The global-mean heating term $\overline{S}_0\langle\Delta Q/\overline{S}_0\rangle$ is removed from ΔQ. All terms are in units of 10^{-2} K day^{-1}. Differences in the stippled regions are not significant at the 95% confidence level as determined by an adjusted Student's t test.

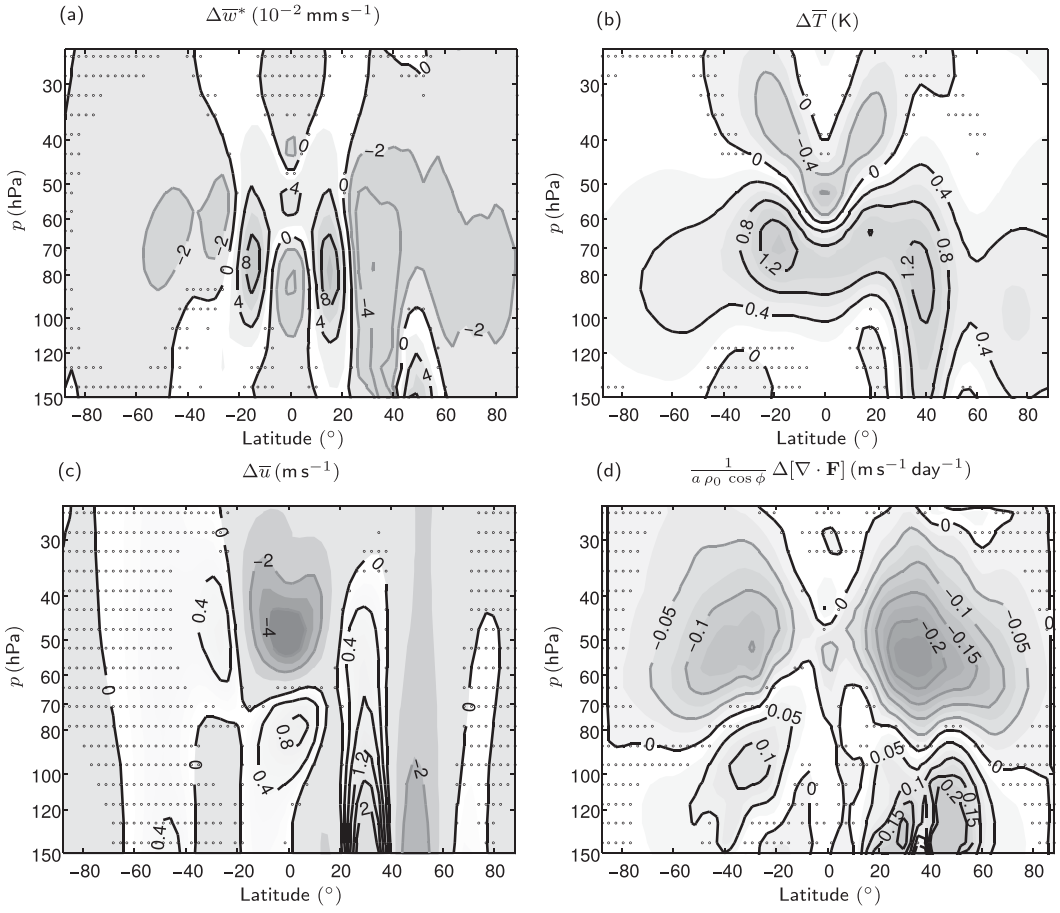

FIG. 10. Full idealized 3D model run with heating at 15°N and amplitude 5 K but with a wave-1 surface geopotential height perturbation in the Northern Hemisphere. (a) Change in $\overline{w}^*$ (10^{-2} mm s^{-1}). (b) Change in $\overline{T}$ (K). (c) Change in $\overline{u}$ (m s^{-1}). (d) Change in $\nabla\cdot\mathbf{F}/(a\rho_0\cos\phi)$ (m s^{-1} day^{-1}). Differences in the stippled regions are not significant at the 95% confidence level as determined by an adjusted Student's t test.

response in the zonal wind, temperature, and divergence of the Eliassen–Palm flux are more sensitive to the details of the wave force, and these therefore differ between the two hemispheres. In particular, the response in $\nabla\cdot\mathbf{F}$ in the Northern Hemisphere is also stronger and has a smaller meridional length scale, and this implies differences in the pattern of extratropical downwelling response between the two hemispheres.

5. Response to an imposed force

Having observed a dynamical regime where there is significant compensation between an imposed narrow heating by the upwelling response, we explore the response to an imposed mechanical forcing. This is motivated by recent discussion (Cohen et al. 2013, 2014) of the dynamics of model-predicted increases in the strength of the Brewer–Dobson circulation due to increases in the concentrations of long-lived greenhouse gases (e.g., Butchart 2014). It has been previously noted (e.g., Cohen et al. 2013) that while the predicted rate of increase in the strength of the circulation is broadly consistent across many models at about 2% decade^{-1}, there are significant disagreements among the models regarding the quantitative contributions of changes in wave forces from different wave types (Butchart et al. 2011). In some models the change is primarily from parameterized gravity waves, while in others it is primarily from synoptic- and planetary-scale Rossby waves that are resolved by the model dynamics. Cohen et al. (2013) have characterized this as a "compensation" by which the Brewer–Dobson circulation response to changes to the parameterized gravity waves, for example, is compensated by the Brewer–Dobson circulation response to consequential changes in the resolved waves. Cohen et al. (2013, 2014) have discussed possible mechanisms for this compensation, including a role for

barotropic instability; however, the relevance of the latter mechanism has been questioned by Sigmond and Shepherd (2014), who studied compensation in a general circulation model, and by Watson and Gray (2015), who studied it in a stratosphere–mesosphere model.

Returning to the arguments of section 2, consider the wave force G appearing in (1a) to be a function not only of the zonal-mean state but also of some set of external parameters $\mathbf{h} = h_1, h_2, h_3, \ldots$. These parameters could, for example, represent orography on different scales or parameters in a gravity wave scheme. A key point, noted by Cohen et al. (2013) and others, is that if one or more of these external parameters is changed, then the resulting change in G will be due in part to the change in the zonal mean state—that is, the zonal-mean zonal wind and temperature field. We could express this formally by writing the change in G as a part that involves partial derivatives with respect to $\mathbf{h}$ and a part involving partial derivatives (or functional derivatives) with respect to flow variables. A convenient simplification would be to approximate the change ΔG as the sum of two parts: an imposed force $\Delta G_{\mathbf{h}}$ and a force that has to be determined as part of the response ΔG_u.

Just as previously we have asked whether an applied heating perturbation is balanced by upwelling or by a change in Newtonian cooling (or more generally longwave radiative heating) due to a change in temperature, here we ask whether the change in force $\Delta G_{\mathbf{h}}$ arising directly from changing external parameters is balanced by the Coriolis force due to a change in meridional circulation or by a compensating change ΔG_u in the wave force. The horizontal and vertical scales of $\Delta G_{\mathbf{h}}$ are assumed, respectively, to be L_G and D_G.

Again, dynamical scaling arguments may be developed by assuming typical changes Δu in the zonal velocity and ΔG_u in the wave force and then exploiting (1b)–(1d) to compare the relative sizes of the terms in (1a) that may balance $\Delta G_{\mathbf{h}}$. In a method analogous to the arguments in section 2, it is assumed that the responses in all the dependent variables have characteristic horizontal and vertical length scales L_G and D_G, respectively. Using (1b), (1d), and (1c) in sequence, it follows that the ratio of the relative sizes of the ΔG_u term and the Coriolis force term $f\overline{v}^*$ in the zonal momentum equation is

$$\frac{\Delta G_u}{\Delta U}\frac{N^2 D_G^2}{\alpha f^2 L_G^2} \sim K\frac{N^2 D_G^2}{\alpha f^2 L_G^2}, \tag{9}$$

where again K is a typical value of the ratio $\Delta G_u/\Delta U$—that is, of the sensitivity of the wave force G to the velocity. This is simply (2) with the length scales L_Q and D_Q of the heating replaced by the length scales L_G and D_G of the externally determined part of the change in wave force $\Delta G_{\mathbf{h}}$.

When the dynamical aspect ratio (9) is large, $\Delta G_{\mathbf{h}}$ is balanced primarily by the adjustment ΔG_u in the wave force rather than by a Coriolis force. In other words, a narrow mechanical forcing is ineffective in driving a meridional circulation. (Recall from section 2 that, in contrast, a narrow heating is effective in driving a meridional circulation, since it is primarily balanced by upwelling.)

To test this scaling argument, we carry out a series of model experiments similar to those described in section 4c but rather than adding an imposed heating to the right-hand side of the thermodynamic equation in (1d), we add an imposed force to the right-hand side of the zonal momentum equation in (1a). The zonally symmetric force is imposed artificially, directly representing the imposed force $\Delta G_{\mathbf{h}}$ and is given by

$$\Delta G_{\mathbf{h}} = \frac{G_A}{(a\rho_0\cos\phi)}\exp\left[-\frac{(\phi+\phi_{\max})^2}{2\Delta\phi^2} - \frac{(z-z_{\max})^2}{2\Delta z^2}\right].$$

The imposed force has the form of a single Gaussian peak centered at $\phi_{\max} = 30°$S. All other corresponding parameters describing the height and width of the forcing are set to those used in the heating case in (4). The amplitude G_A is set to $1\,\mathrm{kg\,m^{-1}\,s^{-2}}$. The imposed force for the $\Delta\phi = 5°$ and $\Delta\phi = 10°$ cases are shown in Fig. 11a. We explore a range of widths, $\Delta\phi = 5°$, 6°, 8°, 10°, and 12°, and plot the ratio of the response ΔG_u to the imposed forcing $\Delta G_{\mathbf{h}}$ at the location of the maximum in the imposed Gaussian forcing (Fig. 11b). The amount of compensation is strongly dependent on the width of $\Delta G_{\mathbf{h}}$. For a narrow force regime, $\Delta\phi = 5°$, there is significant compensation of the imposed wave force by the wave response (~70%). The change in wave force in the narrow force case ($\Delta\phi = 5°$), which is shown in Fig. 11c, has a similar spatial structure to the imposed forcing. The compensation is not perfect and the remaining (uncompensated) wave torque drives a meridional circulation as shown by the plot of the residual-mean meridional velocity in Fig. 11e, ~30% of what would be produced in the absence of an active wave field. Corresponding plots to Figs. 11c–e for a wider forcing case ($\Delta\phi = 10°$) are shown in Figs. 11d–f. The amplitude of the wave response is smaller than in the narrow force regime and the response extends into the other hemisphere. Correspondingly, the Coriolis term is larger as can be seen from the plot of $\Delta\overline{v}^*$ in Fig. 11f. The compensation in

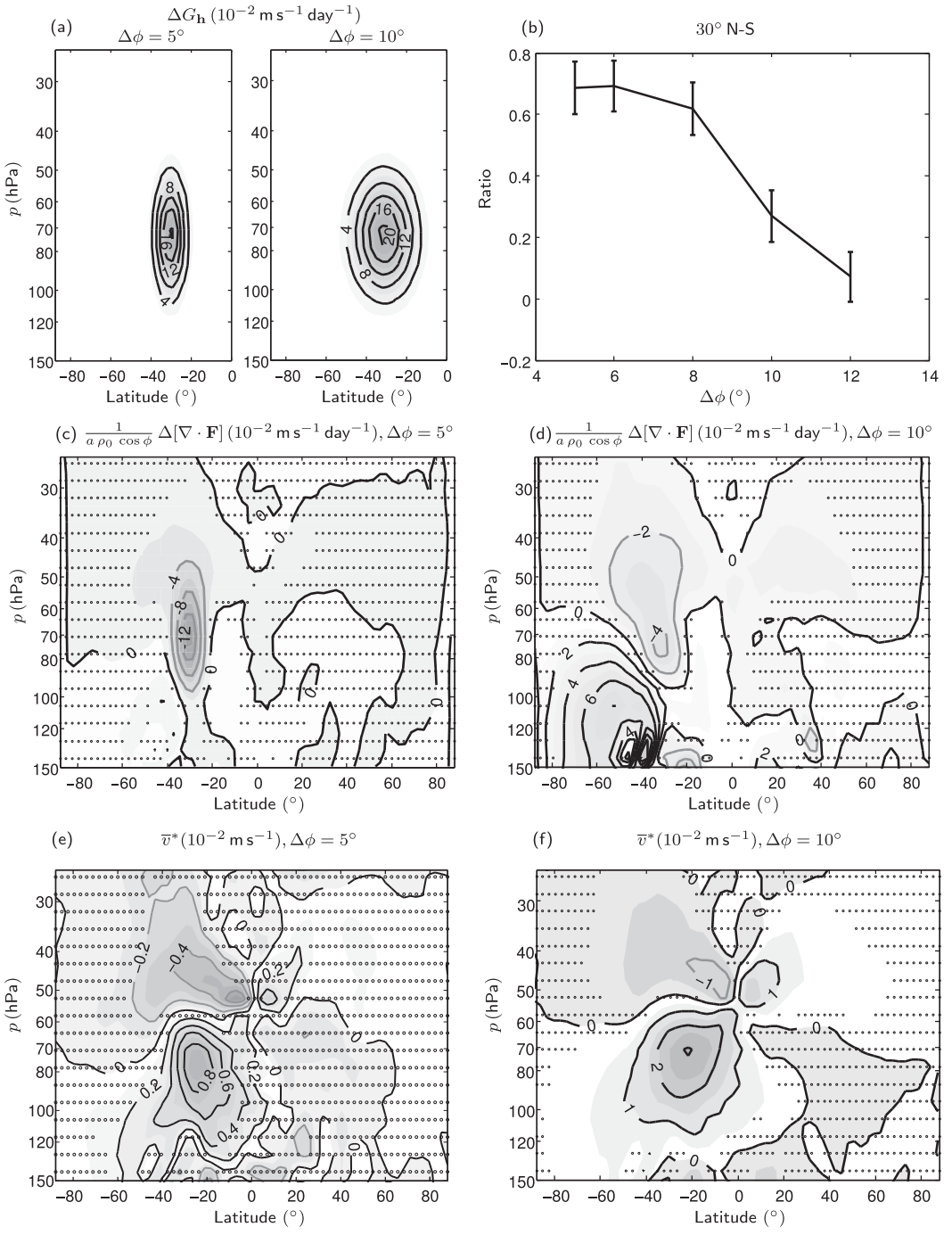

FIG. 11. (a) Imposed Gaussian perturbation in the idealized 3D model $\Delta G_{\mathbf{h}}$ located at 30°S only. Two different widths ($\Delta\phi = 5°$ and 10°) are shown. Note that only the Southern Hemisphere is plotted as the perturbation is only in this hemisphere. (b) Ratio of $\Delta G_u/\Delta G_{\mathbf{h}}$ at the maximum of the perturbation (30°S, 78 hPa) plotted against the width of the perturbation $\Delta\phi$. (c) Change in $\nabla\cdot\mathbf{F}/(a\rho_0\cos\phi)$ for a perturbation of width $\Delta\phi = 5°$. (d) As in (c), but for a width of $\Delta\phi = 10°$. (e) Change in $\overline{v}^*$ for $\Delta\phi = 5°$. (f) As in (e), but for $\Delta\phi = 10°$. Note the different contour intervals in (e) and (f). Differences in the stippled regions are not significant at the 95% confidence level as determined by an adjusted Student's t test.

this case is ~25%, decreasing to ~10% for the case where $\Delta\phi = 12°$.

Our findings are consistent with the results reported by Cohen et al. (2013), who, in seeking to explain the compensation in the driving of the meridional circulation between the resolved waves and the parameterized waves, carried out numerical experiments where a given force (corresponding to our $\Delta G_{\mathbf{h}}$) was applied and the resolved waves, and hence the wave force due to those waves (corresponding to our ΔG_u), were allowed to change as part of the response. They found that a narrower $\Delta G_{\mathbf{h}}$ resulted in more compensation that a broader narrower $\Delta G_{\mathbf{h}}$. However, in discussing relevant mechanisms, Cohen et al. (2013) argued that a narrower force would tend to give rise to change in sign of potential vorticity gradients and that the resulting dynamical instability would give rise to a redistribution of the wave force. This would suggest some kind of nonlinear threshold behavior where a certain amplitude of an applied forcing was required for instability but we found, in the 3D simulations with applied heating, that the amplitude of the meridional circulation response is linear in the amplitude of the applied heating, suggesting that no particular instability threshold has to be exceeded to give the required dynamical effect. The same conclusion might be drawn from the results of Watson and Gray (2015), who found the linear behavior that changing the sign of the applied force simply changed the sign of the response. Our argument presented above, on the other hand, simply asserts that there will be, as part of the response to an applied heating ΔQ (emphasized in sections 2–4) or an applied perturbation to part of the wave force ($\Delta G_{\mathbf{h}}$ in our notation), a change of the total wave force because the propagation and dissipation of waves will change as the zonal flow changes. Compensation then occurs provided that the sensitivity of the wave force to changes in the mean state (i.e., $\Delta G_u/\Delta U$) is sufficiently large in a sense made precise by the dynamical aspect ratio in (9).

6. Discussion

In this paper, we have analyzed the response of the circulation to imposed zonally symmetric heating and mechanical forcing when the wave force can change as part of the response. A specific motivation for considering the response to the heating was to understand the double peak in upwelling in the tropical lower stratosphere. A specific motivation for considering the response to mechanical forcing was the relevance to recent discussions of dynamical compensation in the trends in the Brewer–Dobson circulation. In M16, we looked at diagnostic studies of the angular momentum balance and radiative heating and argued that the double peak in upwelling near 70 hPa and 20°N/S is likely to be caused by latitudinal structure in the radiative heating rather than being a response to latitudinal structure in the wave force alone in the absence of any externally imposed structure in the radiative heating. This hypothesis implicitly requires a mechanism by which a long-term change in the meridional circulation can be caused by a change in radiative heating. For such a change to be maintained, there has to be a self-consistent angular momentum balance and hence also a change in the wave force. In this paper we have investigated this hypothesis further in a simple three-dimensional dynamical model, set up in the Held and Suarez (1994) configuration, which we argue captures the essential wave dynamics relevant to the subtropical lower stratosphere. A radiative heating perturbation was imposed by adding two localized regions of heating to the Held–Suarez configuration. For a latitudinally confined diabatic heating perturbation, the dominant balance in the thermodynamic equation in the region of the heating perturbation is between the heating and the upwelling terms. The temperature change makes a relatively small contribution (through the Newtonian cooling term) to the thermodynamic equation in this region and the latitudinal scale of the overall temperature change is much broader than the scale of the heating perturbation, with weak temperature gradients across the tropics and subtropics. Angular momentum balance is maintained by a change in the Eliassen–Palm flux, so that the change in wave force balances the Coriolis force associated with the change in meridional circulation.

We set out scaling arguments to provide some dynamical insight into this circulation response. These arguments assume that the typical magnitude of the change in wave force is K times the typical magnitude of the change in zonal velocity, with K having the dimensions of inverse time. According to these arguments, an applied heating would be primarily balanced by an upwelling provided that the dynamical aspect ratio $KN^2D_Q^2/\alpha f^2 L_Q^2 \gg 1$; that is, the heating is relatively deep and narrow in a sense that depends on the various parameters K, N, f, and α. In particular, large values of the parameter K and/or small values of the Coriolis parameter f (i.e., low latitudes) make it more likely that the condition is satisfied. The scaling arguments are similar to those applied by previous authors (e.g., Fels et al. 1980; Garcia 1987) in considering the zonally symmetric response to heating when the wave force is represented by Rayleigh friction (K is then simply the Rayleigh friction coefficient) but potentially have wider applicability.

We presented explicit zonally symmetric calculations with Rayleigh friction to explore in a crude way the dependence of the response on K. These calculations capture the balance between applied heating and upwelling seen in the 3D simulations when the heating is deep and narrow (cf. Fig. 2b and the third column in Fig. 4). However, they do not capture, for example, the response in zonal velocity (cf. Fig. 2b and the second column in Fig. 4). This response is determined by the details of the dependence of the wave force on the zonal velocity in the 3D simulations. This dependence is poorly represented by Rayleigh friction.

We have demonstrated by varying the width L_Q of the region of applied heating that the scaling arguments give useful insight. As the width is increased, the change in upwelling no longer provides the dominant balance in the thermodynamic equation for the heating perturbation ΔQ and other terms in the thermodynamic equation became more important. In addition to the change in upwelling, both the Newtonian cooling term and changes in static stability must be taken into account to understand how the heating is balanced in this case.

Perhaps surprisingly, for an applied heating with a width of $\Delta\phi = 5°$, the response continues to be dominated by upwelling as the heating is moved from low to high latitudes, suggesting that the value of K was at each latitude sufficiently large to ensure large values of (2). Whether such ''narrow'' higher-latitude cases are an appropriate model of any specifically realized process in the real atmosphere is less clear, although one might consider their relevance to the diabatic effects of trends in extratropical ozone, particularly associated with the ozone hole.

From the radiative calculations in section 4 of M16 and the dynamical calculations in section 2, we deduce that the ozone distribution (and its radiative implications) is an important part of the cause of the double peak in upwelling. In reality, of course, dynamics, radiation, and chemistry are fully coupled and the ozone distribution is determined by transport processes. This is not captured by our dynamical calculations, in which the structure of the applied heating is simply imposed, but for reasons explained in section 5 of M16, these calculations nonetheless seem to give significant insight to the double-peak structure of the low-latitude upwelling.

Note the additional important point that a fixed dynamical heating calculation, which is often used to infer temperature changes that result from changes in constituents such as ozone, would not be relevant here—it is precisely the dynamical heating that is the main response to the structure in the ozone field.

We have noted with respect to the response to an applied mechanical forcing, added to the wave force, that the dynamical discussion leading to (2) can be extended to consider this response. In this case when the applied force is narrow in the sense that the aspect ratio in (9) is large, the applied force is primarily balanced by an adjustment to the flow-dependent wave force rather than by the Coriolis torque. The response in the meridional circulation is therefore small. If it is accepted that differences in parameterized waves correspond to narrow applied forces, which is suggested by the results of Cohen et al. (2013, their Figs. 4b and 5d) and Sigmond and Shepherd (2014, their Fig. 2), then this potentially offers an overall dynamical principle that explains the compensation (between changes in the Brewer–Dobson circulation driven by resolved waves and changes driven by parameterized waves) observed in climate model simulations. Our analysis suggests that the compensation is, to leading order, independent of the details of the background wave force, similar to the case with an imposed heating, and does not rely on a specific mechanism for the dependence of the wave force on the mean state. Of course, if the aspect ratio, in (9) is to be of quantitative use, then the sensitivity K must be estimated. As we have emphasized previously, this is by no means straightforward, because it essentially requires a parameterization of wave force for an arbitrary mean flow.

Acknowledgments. The authors thank Amanda Maycock for help with the radiation code and for helpful discussions. AM and PHi acknowledge funding support from the European Research Council through the ACCI project (Grant 267760) lead by John Pyle. PHi also acknowledges support from an NSERC postdoctoral fellowship. The authors are grateful to Stephan Fueglistaler and Tom Flannaghan for conversations that stimulated some of this work. We received detailed and helpful comments from three anonymous reviewers that improved this manuscript.

APPENDIX

Statistical Methods

The method used to estimate the statistical significance of the difference is closely related to that described in the appendix of M16. The effective number of degrees of freedom n_e is estimated in the same way. In the model experiments, we showed calculations of the difference of the mean between two time series X and Y

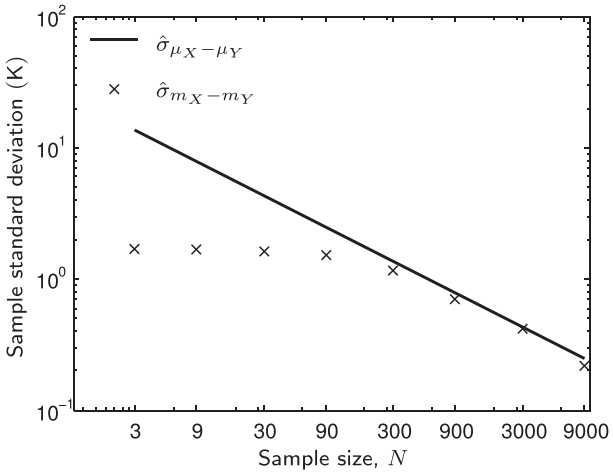

FIG. A1. Sample standard deviations for the temperature difference calculated at 78 hPa and 15°N for the heating located at latitude 15° (case described in section 3). The crosses are the sample standard deviations calculated by scaling the effective number of degrees of freedom with sample size. The line is obtained by estimating the standard deviation of the sample.

and tested the null hypothesis H_0: $\mu_X = \mu_Y$. One of the two sets of formulas in von Storch and Zwiers (2001, 111–116) is used to calculated the test statistic

$$t = \frac{\hat{\mu}_X - \hat{\mu}_Y}{\sqrt{\hat{s}_X^2/\hat{n}_{eX} + \hat{s}_Y^2/\hat{n}_{eY}}} \tag{A1}$$

and the approximate degrees of freedom given by

$$df = \frac{(\hat{s}_X^2/\hat{n}_{eX} + \hat{s}_Y^2/\hat{n}_{eY})^2}{\dfrac{(\hat{s}_X^2/\hat{n}_{eX})^2}{\hat{n}_{eX}-1} + \dfrac{(\hat{s}_Y^2/\hat{n}_{eY})^2}{\hat{n}_{eY}-1}}. \tag{A2}$$

We check this calculation, in features likely to be relevant to the conclusions in this work, by dividing the time series into subsamples of size N and calculating $\hat{\sigma}_{\mu_X-\mu_Y}(N)$ and $\hat{\sigma}_{m_X-m_Y}(N)$. As described in M16, these two quantities should converge as N tends to n.

An example of this calculation is shown in Fig. A1 for the temperature change when the heating is applied at latitude 15° as described in section 3. Let X be a time series for the unperturbed Held–Suarez run (with mean μ_X) and Y be a time series with a heating perturbation (with mean μ_Y). Each time series is 90 000 days long and has $n_{eX} = 470$ and $n_{eY} = 490$. Figure A1 shows the two quantities can be seen to converge. For changes in the zonal wind near the top of the model over the equator, long time scales in the data means that the calculation of $\hat{n}_e$ does not always converge and it is necessary to sum to where the lag (see the appendix of M16) falls below e^{-3}.

REFERENCES

Butchart, N., 2014: The Brewer-Dobson circulation. *Rev. Geophys.*, **52**, 157–184, doi:10.1002/2013RG000448.

——, and Coauthors, 2011: Multimodel climate and variability of the stratosphere. *J. Geophys. Res.*, **116**, D05102, doi:10.1029/2010JD014995.

Cohen, N. Y., E. P. Gerber, and O. Bühler, 2013: Compensation between resolved and unresolved wave driving in the stratosphere: Implications for downward control. *J. Atmos. Sci.*, **70**, 3780–3798, doi:10.1175/JAS-D-12-0346.1.

——, ——, and ——, 2014: What drives the Brewer–Dobson circulation? *J. Atmos. Sci.*, **71**, 3837–3855, doi:10.1175/JAS-D-14-0021.1.

Dickinson, R. E., 1971: Analytic model for zonal winds in the tropics. *Mon. Wea. Rev.*, **99**, 501–510, doi:10.1175/1520-0493(1971)099<0501:AMFZWI>2.3.CO;2.

Fels, S. B., J. D. Mahlman, M. D. Schwarzkopf, and R. W. Sinclair, 1980: Stratospheric sensitivity to perturbations in ozone and carbon dioxide: Radiative and dynamical response. *J. Atmos. Sci.*, **37**, 2265–2297, doi:10.1175/1520-0469(1980)037<2265:SSTPIO>2.0.CO;2.

Forster, P. M., M. Blackburn, R. Glover, and K. P. Shine, 2000: An examination of climate sensitivity for idealised climate change experiments in an intermediate general circulation model. *Climate Dyn.*, **16**, 833–849, doi:10.1007/s003820000083.

Fueglistaler, S., P. H. Haynes, and P. M. Forster, 2011: The annual cycle in lower stratospheric temperatures revisited. *Atmos. Chem. Phys.*, **11**, 3701–3711, doi:10.5194/acp-11-3701-2011.

Garcia, R. R., 1987: On the mean meridional circulation of the middle atmosphere. *J. Atmos. Sci.*, **44**, 3599–3609, doi:10.1175/1520-0469(1987)044<3599:OTMMCO>2.0.CO;2.

Haynes, P., 2005: Stratospheric dynamics. *Annu. Rev. Fluid Mech.*, **37**, 263–293, doi:10.1146/annurev.fluid.37.061903.175710.

——, M. E. McIntyre, T. G. Shepherd, C. J. Marks, and K. P. Shine, 1991: On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces. *J. Atmos. Sci.*, **48**, 651–680, doi:10.1175/1520-0469(1991)048<0651:OTCOED>2.0.CO;2.

Held, I. M., and M. J. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Amer. Meteor. Soc.*, **75**, 1825–1830, doi:10.1175/1520-0477(1994)075<1825:APFTIO>2.0.CO;2.

Hoskins, B. J., and A. J. Simmons, 1975: A multi-layer spectral model and the semi-implicit method. *Quart. J. Roy. Meteor. Soc.*, **101**, 637–655, doi:10.1002/qj.49710142918.

Ming, A., P. Hitchcock, and P. Haynes, 2016: The double peak in upwelling and heating in the tropical lower stratosphere. *J. Atmos. Sci.*, **73**, 1889–1901, doi:10.1175/JAS-D-15-0293.1.

Plumb, R. A., 1982: Zonally symmetric Hough modes and meridional circulations in the middle atmosphere. *J. Atmos. Sci.*, **39**, 983–991, doi:10.1175/1520-0469(1982)039<0983:ZSHMAM>2.0.CO;2.

——, and J. Eluszkiewicz, 1999: The Brewer–Dobson circulation: Dynamics of the tropical upwelling. *J. Atmos. Sci.*, **56**, 868–890, doi:10.1175/1520-0469(1999)056<0868:TBDCDO>2.0.CO;2.

Randel, W. J., R. Garcia, and F. Wu, 2008: Dynamical balances and tropical stratospheric upwelling. *J. Atmos. Sci.*, **65**, 3584–3595, doi:10.1175/2008JAS2756.1.

Scott, R. K., 2002: Wave-driven mean tropical upwelling in the lower stratosphere. *J. Atmos. Sci.*, **59**, 2745–2759, doi:10.1175/1520-0469(2002)059<2745:WDMTUI>2.0.CO;2.

Shepherd, T. G., and C. McLandress, 2011: A robust mechanism for strengthening of the Brewer–Dobson circulation in response to climate change: Critical-layer control of subtropical wave breaking. *J. Atmos. Sci.*, **68**, 784–797, doi:10.1175/2010JAS3608.1.

Sigmond, M., and T. G. Shepherd, 2014: Compensation between resolved wave driving and parameterized orographic gravity wave driving of the Brewer–Dobson circulation and its response to climate change. *J. Climate*, **27**, 5601–5610, doi:10.1175/JCLI-D-13-00644.1.

von Storch, H., and F. W. Zwiers, 2001: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.

Watson, P. A. G., and L. J. Gray, 2015: The stratospheric wintertime response to applied extratropical torques and its relationship with the annular mode. *Climate Dyn.*, **44**, 2513–2537, doi:10.1007/s00382-014-2359-2.