

DISSERTATION

WAVE-MEAN FLOW POSITIVE FEEDBACKS ASSOCIATED WITH SUDDEN
STRATOSPHERIC WARMINGS

Submitted by

Jeremiah P Sjoberg

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2014

Doctoral Committee:

Advisor: Thomas Birner

Richard Eykholt
Rolando Garcia
Wayne Schubert
David Thompson

Copyright by Jeremiah P Sjoberg 2014

All Rights Reserved

ABSTRACT

WAVE-MEAN FLOW POSITIVE FEEDBACKS ASSOCIATED WITH SUDDEN STRATOSPHERIC WARMINGS

Sudden stratospheric warmings – most often characterized by zonal mean zonal wind easterlies at 60°N, 10 hPa – represent the largest dynamical perturbations to the wintertime polar stratosphere. Despite this, the predictability of sudden warmings remains low, in part because the forcing of these warming events involves a nonlinear positive feedback between planetary scale waves and the zonal wind of the stratosphere. In the wave-mean flow positive feedback, wave forcing decelerates the mean flow, allowing enhanced upward wave propagation, which then further decelerates the mean flow, etc., until the mean flow no longer supports wave propagation. This positive feedback process is crucial for the initiation of such events. Because the associated low predictability stems from poorly resolving initiation, this dissertation focuses on increasing mechanistic understanding of the wave-mean flow positive feedback associated with sudden stratospheric warmings.

A simple model of wave-mean flow interaction is the first tool utilized here. In the original form of the model, constant bottom boundary wave forcing, set by geopotential height perturbations, results in a zonal wind state that oscillates between positive values (westerlies) and negative values (easterlies). We present a reformulation of the bottom boundary condition which allows for specification of the upward wave activity flux. Unlike with the original bottom boundary condition, we may precisely set the wave amplitudes propagating into the model domain. With this reformulated model, steady incoming wave fluxes lead to a steady zonal wind response. The oscillating state from the original model is found to rely on a representation of the positive feedback that is too strong. Transient

forcing experiments in the reformulated simple model support previous results that there is a preferential wave forcing time scale on the order of 10 days for sudden stratospheric warmings. Forcing the model near this preferential time scale most efficiently drives the positive feedback.

Lower stratospheric wave fields in reanalysis data show supporting evidence for these preferential wave forcing time scales prior to sudden stratospheric warmings. Pulses of wave activity flux are also analyzed in reanalysis data, and a set of pulses which are a novel proxy for strong wave-mean flow positive feedback are found. The zonal wind near these pulses display the expected characteristics of the positive feedback: strong precedent zonal winds and strong subsequent wind decelerations. This proxy is thus a useful diagnostic for the wave-mean flow positive feedback.

A general circulation model forced by idealized planetary scale topography is employed to perform high order experiments. By stepwise increasing the height of the topography, we find that the frequency of sudden stratospheric warmings within the model increases nonlinearly to a maximum at moderate topographic heights and then strongly jumps down to a lower, steady value for still higher topography. Analyzing the proxy for positive feedback here reveals that the positive feedback is strongest in the range of topographic heights associated with the largest occurrence of sudden warmings, and also that preferential wave forcing time scales on the order of 10 days are upheld.

ACKNOWLEDGEMENTS

Many thanks to my Master's and PhD advisor Thomas Birner, who guided me along the path to completion of this dissertation. I also want to thank the members of my committee from the Department of Atmospheric Science, David Thompson and Wayne Schubert. Thanks as well go out to Rolando Garcia who made the journey from Boulder numerous times to participate in committee responsibilities. Last, but not least from the committee, I thank Richard Eykholt who has dutifully represented the outside committee member position on both my Master's and PhD committees.

The research that went into this work would have been considerably more difficult, if not impossible, without the fine help from the support staff in the Department of Atmospheric Science – Jamie Schmidt, Jennifer Armstrong, Cliff Schulenberg, and Karrie Butler in particular. Their constant guidance and assistance truly facilitated a positive work environment.

I want to acknowledge the friends and colleagues I've made and maintained during the course of this dissertation research. These compatriots in science – whom I hope know who they are – were ever the source of mental replenishment through discussions both scientifically productive and not, commiserating, and mutual imbibing on the fruits of life. Members of the Birner group, you're included in this. I specifically wish to acknowledge John Albers, with whom I had bountiful discussions about wave propagation.

Finally, I thank my wife Katie for her constant emotional guidance, indomitable inspiration, and deep patience for the life and quirks of a scientist.

This dissertation is typeset in L^AT_EX using a document class designed by Leif Anderson.

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	iv
List of Figures	vii
Chapter 1. Introduction	1
Chapter 2. Simple modeling of wave-mean flow feedbacks and sudden stratospheric warmings	9
2.1. Model derivation	9
2.1.1. Holton-Mass model	9
2.1.2. Reformulation and parameters	14
2.2. Steady bottom boundary incoming wave activity flux	17
2.3. Time-dependent bottom boundary incoming wave activity flux	28
2.3.1. Zero baseline	28
2.3.2. Nonzero baseline	36
2.4. Tropopause effects	39
2.5. Conclusions	44
Chapter 3. Wave-mean flow feedbacks in reanalyses	49
3.1. Wave forcing time scales of SSWs	49
3.1.1. Lower stratospheric heat flux	49
3.1.2. Upper troposphere to upper stratosphere wave activity fluxes	61
3.2. Pulse duration diagnostic for strong wave-mean flow positive feedbacks	65
3.3. Tropopause Inversion Layer in Reanalysis Data	82

3.4. Conclusions	82
Chapter 4. General circulation model wave-mean flow positive feedbacks	87
4.1. Model introduction	87
4.2. Topography experiments	89
4.2.1. Time mean state	89
4.2.2. Pulse analysis	100
4.3. Tropopause inversion layer experiments	109
4.4. Conclusions	119
Chapter 5. Conclusions	122
5.1. Summary	122
5.2. Outlook	124
References	126
Appendix A. Additional figures	133

LIST OF FIGURES

2.1	Zonal wind evolution from steady geopotential and EP flux integrations.....	18
2.2	Bottom boundary geopotential from two steady EP flux integrations	20
2.3	Slowly ramping incoming EP flux experiment’s zonal wind and stability states ...	22
2.4	Slowly ramping bottom boundary geopotential experiment’s zonal wind and stability states.....	25
2.5	Imposed deceleration experiment in both the EP flux and geopotential models ...	26
2.6	Transient EP flux pulse schematic for time-dependent experiments	29
2.7	EP flux amplitude and total EP flux at the bottom boundary necessary to force a SSW-like transition in the EP flux mode.....	30
2.8	EP flux model feedback activated, deactivated experiment	32
2.9	Pulse duration and minimum zonal wind for 20 day pulses with various amplitudes	35
2.10	EP flux amplitudes necessary to force transitions as a function of baseline EP flux for both models.....	37
2.11	Our three stratification profiles and associated zonal wind equilibrium states	40
2.12	Imposed deceleration experiment with the tropopause, TIL stratification setups ..	43
3.1	CFD of meridional heat fluxes and 40 day averaged heat fluxes.....	51
3.2	Composites of instantaneous, time-averaged heat flux prior to major, minor SSWs	54
3.3	Average duration of heat flux above a given threshold for pre-SSW period, all DJF	57
3.4	Hindcast results of applying wave durations results.....	60
3.5	EP flux pulse durations and maxima in 200 hPa ERAi data and an autoregressive model of order 1	63

3.6	Cumulative frequency distribution of ERAi EP fluxes at all pressure levels	64
3.7	Vertically-coherent pulse algorithm schematic	66
3.8	Mean durations of vertically-coherent pulses in ERAi	67
3.9	Mean maxima of vertically-coherent pulses in ERAi	68
3.10	Mean durations of vertically-coherent pulses of four wavenumbers in ERAi	70
3.11	Cumulative frequency distributions of vertically-coherent pulse depths	72
3.12	Total number of deep pulses and percent of these which are increasing or decreasing	74
3.13	Mean durations of increasing and decreasing pulses	75
3.14	Mean EP flux maxima of increasing and decreasing pulses	76
3.15	Mean total EP flux for increasing and decreasing pulses	77
3.16	Anomalous zonal wind composited about increasing pulses	79
3.17	Anomalous zonal wind composited about decreasing pulses	79
3.18	Changes in zonal wind from the central date for increasing pulses	80
3.19	Changes in zonal wind from the central date for decreasing pulses	81
4.1	Occurrence of SSWs in the GFDL GCM as a function of topography	90
4.2	Zonal wind, temperature, EP flux divergence, and upward EP fluxes as a function of topographic parameter and pressure	91
4.3	Anomalies and per meter of topography changes in the time mean upward EP fluxes and meridional EP fluxes as a function of pressure and topographic parameter	92
4.4	Histograms of upward EP fluxes at 500 hPa	94
4.5	Histograms of upward EP fluxes at 500 hPa for waves 1, 2, 3, and 4+	95

4.6	Percentage of time for which the index of refraction is negative for GCM topography runs	97
4.7	Change in the time mean wave 2 upward EP fluxes and wave 2 meridional EP fluxes per meter change of topographic parameter as a function of pressure and topographic parameter	99
4.8	Durations and upward EP flux maxima for all pulses as a function of pressure and topography	101
4.9	Durations and upward EP flux maxima for vertically-coherent pulses	102
4.10	Number of deep pulses and associated percentage of increasing pulses as a function of topography	103
4.11	Differences in durations between increasing and decreasing pulses	104
4.12	Differences in upward EP flux amplitudes between increasing and decreasing pulses	105
4.13	Differences in total upward EP flux between increasing and decreasing pulses	106
4.14	Means of the anomalous zonal wind prior to the central date for increasing pulses	107
4.15	Differences between increasing pulses and decreasing pulses in the means of the anomalous zonal wind prior to the central date	108
4.16	Differences between increasing pulses and decreasing pulses in the means of the anomalous zonal wind changes following the central date	108
4.17	Zonal winds, temperatures, and EP flux vectors for $p_{pert} = 100, 200$ hPa	111
4.18	Stratification for for $p_{pert} = 100, 200$ hPa	112
4.19	Radiative equilibrium temperature and stratification, and change from the 0 K run, for all λ	113

4.20	Stratification and differences from 0 K integration as a function of λ	114
4.21	Zonal winds, temperatures, EP flux divergences, and upward EP fluxes for all λ ..	115
4.22	Anomalies of EP flux divergence and upward EP fluxes for all λ	116
4.23	Zonal winds and anomalous EP flux vectors – again scaled by inverse density – for 4 values of λ	117
4.24	Percent of time which the meridional PV gradient is reversed	118
4.25	Total number of major, minor SSWs as a function of λ	119
A.1	Mean durations for vertically-coherent, wavenumber decomposed waves with 500 hPa starting level.....	133
A.2	Mean amplitudes for vertically-coherent, wavenumber decomposed waves with 500 hPa starting level.....	134
A.3	Histograms of upward EP fluxes at 300 hPa for waves 1, 2, 3, and 4+	135
A.4	Durations and upward EP flux maxima for increasing pulses	136
A.5	Composites of zonal winds about increasing pulses with 500 hPa starting level ...	137
A.6	Composite differences in zonal wind between increasing and decreasing pulses	138
A.7	Composite differences in zonal wind changes from the central date between increasing and decreasing pulses	138

CHAPTER 1

INTRODUCTION

The large-scale character of the wintertime polar stratosphere is that of a broad circumpolar vortex. Normally active ultraviolet absorption – primarily by ozone and diatomic oxygen – wanes from autumn until winter as the solar forcing at these polar latitudes likewise wanes. Without the presence of this solar (thermal) forcing, the polar stratosphere cools via longwave radiation. Subtropical and equatorial latitudes meanwhile receive ample solar radiation such that a strong equator-to-pole temperature gradient is established, with the maximum magnitude of this gradient located near 60° (north or south depending respectively on boreal or austral winter). In line with thermal wind balance, a jet of zonal wind – termed the polar night jet – concurrently establishes near this maximum gradient, in one sense marking the edge of the polar vortex. A detailed, historical discussion of this may be found in Hare (1960).

Through balloonsonde measurements and synoptic-scale interpolation, Scherhag (1952) identified an event in which the cold nature of the wintertime polar vortex was greatly disturbed through a sudden warming on the order of 40 K in 2 days. The events that Scherhag first identified, later labeled sudden stratospheric warmings (SSWs), are marked by such order 50 K warmings on order 1 week time scales and by large decelerations of the polar night jet. The warming and zonal wind deceleration of these events are observed to begin near the stratopause and propagate downwards, occasionally reaching the tropopause, over a 2 week time scale. Though the perturbation manifests quickly, the duration over which the vortex state remains disturbed can be much longer, sometimes greater than a month (Schoeberl 1978).

Naturally, Scherhag's discovery was significant, and led to vigorous study to determine the full character and cause of these events. Analysis of the synoptic-scale evolution of atmospheric fields during Scherhag's and later SSWs led to the understanding that stratospheric-level, planetary-scale waves are strongly linked to the evolution of these events (e.g. Teweles and Finger 1958). It was found that these large-scale wave disturbances transported the necessary energy to drive the observed temperature changes via eddy meridional heat fluxes (Mahlman 1969). Furthermore, in part motivated by the association of sudden stratospheric warmings with tropospheric blocks (see Miyakoda et al. 1970, and references within), later studies on the energetics of SSWs revealed that significant eddy energy was supplied from the troposphere to the stratosphere during these events (e.g. Julian and Labitzke 1965). Rocketsonde evidence at that time showed that these planetary-scale waves could propagate up to the stratopause (Finger et al. 1966). Thus the dynamical picture is one where thermal and orographic forcing in the troposphere leads to upward propagating waves that reach high altitudes and supply the necessary energy to drive a SSW.

Theory on the vertical propagation of planetary waves was developed concurrently to the understanding of SSWs. Charney and Drazin (1961) first identified the conditions under which stationary waves may propagate vertically: when the zonal wind is positive (westerly) yet below some critical threshold. Though this theoretical advancement was based on the assumption of latitudinally-invariant zonal wind and Coriolis parameter (i.e. the β -plane assumption), it performed well in observations. Through use of a linearized model, Dickinson (1968) expanded on this advancement by considering a latitudinally-varying zonal wind. He demonstrated that differences in the zonal wind speeds leads to "wave ducts" towards which waves outside of these ducts will propagate. Matsuno (1970) later showed that this ducting of waves is related to regions where a refractive index for planetary waves – which he derived –

is large and positive. In the stratosphere, these ducting regions during midwinter correspond to the equatorial zero wind line and to the edge of the polar vortex (i.e. the polar night jet). The latter of these regions results from large vertical shear and large meridional curvature of the zonal wind there (Chapman and Miles 1981). This polar vortex maximum of the index of refraction is crucial to the forcing of SSWs as it provides a “wave guide” upon which the tropospheric planetary waves may propagate to great depths and subsequently decelerate the zonal wind. Yet, given that the index of refraction approaches $+\infty$ as both the zonal wind approaches zero near to and as the meridional gradient in Coriolis parameter maximizes at the Equator, the equatorial wave guide is considerably more dominant than that of the polar vortex and thus most of the stratospheric planetary waves will refract towards the equatorial zero wind line (Matsuno 1970).

Despite the breadth of knowledge on the relation between vertically propagating planetary waves and SSWs, it was not until Matsuno (1971) that a mechanistic study on the forcing of SSWs was performed. He derived a two-dimensional, quasi-geostrophic, and linearized model of zonal wind and geopotential height. Upward waves in his model were forced through bottom boundary geopotential height perturbations at a nominal tropopause. These perturbations were idealized and set to roughly match observed magnitudes during SSWs. By doing so, Matsuno generated a prototype sudden stratospheric warming, thereby demonstrating that upward-propagating planetary waves were indeed principally responsible for forcing SSWs.

An additional, non-trivial result of Matsuno (1971) relates to the polar vortex wave guide. He found that as the waves force the zonal wind towards easterlies, a critical layer to wave propagation is formed. Like the equatorial zero wind line, the index of refraction becomes larger, focusing more waves towards this critical layer. Because waves can not propagate

past this critical layer, they converge near the critical layer and thus will further decelerate the zonal wind, leading to a downward propagation of easterlies. This interaction between the polar vortex critical layer and the upward waves represents a negative wave-mean flow feedback wherein upward wave fluxes lead to decreased amplitudes of upward wave fluxes.

Matsuno proposed that this critical layer feedback was a necessary condition for the generation of a SSW. Later studies, however, cast doubt on the necessity of a critical layer mechanism leading to SSWs. Notably, Geisler (1974) considered a further simplified form of the equations and mean state, and showed (following Geisler and Dickinson 1974) that critical layer formation leads to wave reflections away from the critical layer. These reflections slow the downward propagation of zonal wind reversal leading to a SSW time scale which is too long.

Furthermore, Matsuno (1971) did not explain the observed wave magnitude enhancements prior to the formation of a SSW. This was problematic as not all instances of SSWs are strongly associated with large bursts of wave fluxes from the troposphere. A recent example is the Antarctic SSW of September 2002, during which tropopause-level Rossby wave fluxes were not so anomalously large that one would consider the event to have been tropospherically-generated (Charlton et al. 2005). Thus without a mechanistic explanation of how the wave fluxes grow nonlinearly prior to events, the field's understanding of forcing SSWs was not complete.

This gap was filled by the Geisler (1974) who, in addition to the above results, showed that a nonlinear feedback between the waves and mean flow exists prior to SSW-like variability, leading to a quasi-resonant amplification of the waves. In this, wave dissipation and absorption decelerate the mean flow which in turn relaxes the Charney-Drazin propagation condition. This represents a positive wave-mean flow feedback wherein upward wave fluxes

lead to enhanced upward wave fluxes. Geisler's findings were later confirmed and expanded upon by Plumb (1981).

Both of the above feedbacks – the negative critical layer feedback and the positive quasi-resonance feedback – aptly describe observed features of a SSW and as such are understood to play a mechanistic role in the generation of warming events (Plumb 2010). The modeling studies discussed so far may have contained aspects of both feedbacks, but the authors focused on only a single one. It was perhaps the simple modeling study of Holton and Mass (1976) that first brought together both feedbacks. In their study, Holton and Mass found that steady bottom boundary wave forcing (introduced through geopotential height perturbations) below a critical threshold led to a steady zonal wind response. Above this threshold, however, both the zonal mean zonal wind and wave fluxes would undergo large vacillations with a semi-regular period.

It is these vacillations which represent both the positive and negative feedbacks discussed above. As noted, the bottom boundary geopotential perturbations must be large to be within this vacillating regime. These large perturbations produce large wave fluxes which sufficiently decelerate the mean flow to allow quasi-resonant amplification of the wave fluxes. These amplified waves are large enough to force the zonal wind to reverse, thereby creating a critical layer to wave propagation. This critical layer descends as upward waves continue to dissipate just below it, reversing the mean flow there. Descent of the easterlies continues until wave reflection off the critical layer dominates the upward wave fluxes down to and including the bottom boundary. During this descent of the easterlies, only small wave forcing above the critical layer occurs due to evanescent waves; radiative damping however is strong at these middle and upper levels, resulting in strong acceleration of the zonal wind. As the zonal wind become westerly once again, the conditions are again suitable for the bottom boundary

geopotential perturbations to force large upward wave fluxes which begin the vacillation anew: the wave-mean flow positive feedback nonlinearly amplifies these wave fluxes, in turn initiating the wave-mean flow negative feedback which dampens the waves and allows the zonal wind to accelerate.

A single full period of a vacillation – i.e. from zonal wind reversal to zonal wind restoration – represents a prototype SSW. Being a prototype, the Holton-Mass model SSW does not fully capture all aspects of real world SSWs, but this does not decrease its usefulness. Rather, the simplicity of the model is an asset to mechanistic study of SSWs and their role in the dynamics of the stratosphere-troposphere system. Evidence of this comes from its wide adaptability to many subjects of study: stratospheric self-determination (Holton and Mass 1976), wave reflections (Harnik 2009), mechanisms for hemispheric differences in wave amplitudes (Plumb 1989), and zonally-averaged ozone distributions (Garcia and Hartmann 1980).

In part from the simple modeling and observational efforts discussed above, understanding of the SSW problem has been greatly advanced. Despite the field’s breadth of knowledge on the forcing of these events, the overall predictability of SSWs remains quite low relative to the predictability of large-scale circulation in the stratosphere (Mukougawa and Hirooka 2004). This low predictability is problematic as medium-range tropospheric predictability may be enhanced following successful simulation of a sudden warming (Thompson et al. 2002, Baldwin et al. 2003). While events which are forced by the classical large amplitude, tropospheric burst of wave fluxes may be predictable by numerical weather models with time scales of a month (Mukougawa and Hirooka 2004), events related to preconditioning (of the type described by McIntyre 1982, McIntyre and Palmer 1983) and quasi-resonant amplification of the waves appear to have a much shorter time scale of predictability (Hirooka

et al. 2007). Models which are well-suited for capturing this nonlinear wave amplification have been shown to reproduce with fidelity characteristics of the latter types of SSW events (Esler et al. 2006), though sensitivities to initial conditions still exist. Given this, it appears that poor representation in numerical models of the nonlinear, positive feedback processes which grow wave amplitudes may lead to poor prediction of SSWs.

To more fully understand the above nonlinear wave amplification and the wave-mean flow positive feedbacks which lead to it is the goal of this dissertation. The approach to doing so is through what some authors have termed a “model hierarchy” (Hoskins 1983, Held 2005). In this, mechanistic results are simultaneously sought through simple models, observations, and complex models to help eliminate findings of false causality from only considering a single source of information. The simple model used here is the Holton-Mass model and a reformulation of it. The observations will be analyzed through reanalysis data products, which are not strictly observations but are as close as we have for the regions of study. For the complex model, we utilize a dynamical core general circulation model – i.e. the model only resolves the atmospheric dynamical equations.

Rather than seeking to answer all problems related to stratospheric wave-mean flow positive feedbacks – a truly insurmountable task – this dissertation will focus on several aspects that may increase our understanding and qualitative predictability. Four primary results are presented. Firstly, constant incoming wave activity flux leads to a constant zonal wind response. Secondly, the stratosphere has an internal wave forcing time scale. Wave forcing matching or near to this time scale most strongly activates the positive feedback cycle, resulting in large zonal wind decelerations and SSWs. Thirdly, increasing pulse duration with height is a useful diagnostic for analyzing characteristics of stratospheric wave-mean flow positive feedbacks. Finally, the tropopause inversion layer (TIL) may provide a nontrivial

source of wave amplitudes in the lower stratosphere which may then interact with the positive feedbacks higher up.

I will introduce the simple model and show associated results in Chapter 2. Chapter 3 gives the results from the observational analysis. Chapter 4 contains analysis of a general circulation model. I will present my conclusions and future work in Chapter 5.

CHAPTER 2

SIMPLE MODELING OF WAVE-MEAN FLOW FEEDBACKS AND SUDDEN STRATOSPHERIC WARMINGS

Material from Section 2.1.1 is published in Sjöberg (2010). Other material from this chapter is published in Sjöberg and Birner (In Press).

2.1. MODEL DERIVATION

2.1.1. HOLTON-MASS MODEL.

As stated in the Introduction, the simple modeling of this section makes use of the Holton-Mass model with two different bottom boundary conditions. Here we derive the model with the standard geopotential perturbation bottom boundary condition.

The Holton-Mass model is derived from the quasi-geostrophic (QG) primitive equations of dry atmospheric motion on a β -plane. The form of these are presented in Andrews et al. (1987):

$$(2.1) \quad D_g u_g - f_0 v_a - \beta y v_g = X,$$

$$(2.2) \quad D_g v_g + f_0 u_a + \beta y u_g = Y,$$

$$(2.3) \quad \partial_x u_a + \partial_y v_a + \frac{1}{\rho_0} \partial_z (\rho_0 w_a) = 0,$$

$$(2.4) \quad D_g \theta_e + w_a \partial_z \theta_0 = Q,$$

where

$$D_g \equiv \partial_t + u_g \partial_x + v_g \partial_y,$$

subscripts “g” represent geostrophic wind components, and subscripts “a” represent ageostrophic wind components. θ_e represents a small departure from a reference potential temperature $\theta_0(z)$. X and Y here represent friction and other nonconservative processes; for this analysis, these will be assumed to be zero. Q is a radiative heating term which restores the temperature to T_0 , the radiative equilibrium temperature. Q is assumed to be

$$Q = -\alpha c_p (T - T_0),$$

while T_0 is assumed to satisfy thermal balance,

$$\partial_y T_0 = -f_0 \frac{H}{R} \partial_z u_R,$$

for $\partial_z u_R = \Lambda(t)$ such that $u_R(z, t) = u_R(0, t) + \Lambda(t)z$ is the mean radiative zonal wind.

Taking $\partial_x(2.2) - \partial_y(2.1)$ and combining with the thermodynamic equation (2.4) gives the QG potential vorticity equation

$$(2.5) \quad D_g q = -\frac{f_0^2}{\rho_0} \partial_z \left(\frac{\alpha \rho_0}{N^2} \partial_z \psi \right)$$

where

$$q = f_0 + \beta y + \nabla^2 \psi + \frac{1}{\rho_0} \partial_z \left(\rho_0 \frac{f_0^2}{N^2} \partial_z \psi \right)$$

and ψ is the geostrophic streamfunction. Equation (2.5) describes the conservation of QG potential vorticity with the inclusion of a diabatic term that is a function of the radiative heating.

Following Dickinson (1969) we take the zonal mean of the fields in (2.1-2.4) and combine to make the zonal wind equation. With the linearized form of the QG potential vorticity

equation, our set of equations become

$$(2.6) \quad (\partial_t + \bar{u}\partial_x) q' + \partial_y \bar{q} \partial_x \psi' + \frac{f_0^2}{\rho_0} \partial_z \left(\frac{\alpha \rho_0}{N^2} \partial_z \psi' \right) = 0 ,$$

$$(2.7) \quad \partial_t \left[\partial_{yy} \bar{u} + \frac{f_0^2}{\rho_0 N^2} \partial_z (\rho_0 \partial_z \bar{u}) \right] = - \frac{f_0^2}{\rho_0 N^2} \partial_z [\alpha \rho_0 \partial_z (\bar{u} - u_R)] \\ + \frac{f_0^2}{N^2} \partial_{yy} \left[\frac{1}{\rho_0} \partial_z (\rho_0 \overline{\partial_x \psi' \partial_z \psi'}) \right] ,$$

where the perturbation QG potential vorticity is

$$q' = \nabla^2 \psi' + \frac{f_0^2}{\rho_0} \partial_z \left(\frac{\rho_0}{N^2} \partial_z \psi' \right) ,$$

and the basic state potential vorticity gradient is

$$\partial_y \bar{q} = \beta - \partial_{yy} \bar{u} - \frac{f_0^2}{\rho_0} \partial_z \left(\frac{\rho_0}{N^2} \partial_z \bar{u} \right) .$$

This model assumes Newtonian damping to a specified radiative state, where the damping rate is given by

$$(2.8) \quad \alpha(z) = [\alpha_{mid} + \tanh\{(z - z_{mid})/H\}] * 10^{-6} \text{s}^{-1} .$$

Here $\alpha_{mid} = 1.5 \cdot 10^{-6} \text{ s}^{-1}$ is the midlevel radiative damping, $z_{mid} = 35 \text{ km}$ is the midlevel altitude, and $H = 7 \text{ km}$ is the scale height. Expressing the damping in terms of a time scale, the lowest levels of the model damp on a time scale of approximately 23 days, the mid levels on a scale of approximately 8 days, and the upper levels on a scale of approximately 5 days.

This model is confined to a β -plane channel centered at 60°N with a notional channel width of 60° . The model assumes zero normal flow across the latitudinal boundaries of the model such that the geostrophic streamfunction ψ' and the meridional wind $\partial_x \psi'$ are

zero at these boundaries. It is likewise assumed that both the energy density of the wave perturbations and the thermal wind vanish at the upper boundary. At the lower boundary, both the zonal mean zonal wind and wave perturbations are specified.

The wavelike solution assumption is made, given by

$$(2.9) \quad \psi'(x, y, z, t) = \text{Re} [\Psi(z, t) e^{ikx}] e^{z/2H} \sin ly$$

$$(2.10) \quad \bar{u}(y, z, t) = U(z, t) \sin ly,$$

where $k = s/(a \cos \pi/3)$ is the zonal wavenumber, with $s = 1, 2, \dots$, the zonal planetary wavenumber. The meridional wavenumber $l = 3/a$ is specified such that $\sin ly$ is maximized at the channel center and 0 at the boundaries.

Using these wavelike solutions result in the following boundary conditions (described above):

$$\begin{aligned} \Psi(z_B, t) &= \frac{g}{f_0} h(t), & U(z_B, t) &= U_R(0, t), \\ \Psi(z_T, t) &= 0, & \partial_z U|_{z_T} &= \partial_z U_R|_{z_T} = \Lambda(t), \end{aligned}$$

where $h(t)$ is the wave forcing amplitude at the bottom of the model, the radiative equilibrium wind U_R is specified by the vertical shear $\partial_z U_R = \Lambda(t)$, and $\Lambda(t)$ is typically taken to be between 1 and 3 m s⁻¹ km⁻¹. Also from the inclusion of the wavelike solutions will arise terms with meridional dependencies defined by $\sin^2 ly$, which are expanded into a Fourier sine series approximation

$$\sin^2 ly \approx \epsilon \sin ly,$$

where higher order terms have been ignored and $\epsilon \equiv 8/(3\pi)$.

With the wavelike assumptions for the linearized quasi-geostrophic potential vorticity equation and the zonal mean zonal wind equation, together with the associated boundary conditions, we may solve these equations to the form given by Holton and Mass (1976):

$$\begin{aligned}
(2.11) \quad 0 = & - \left[k^2 + l^2 + \frac{f_0^2}{N^2} \frac{1}{4H^2} \right] \partial_t \Psi \\
& + \frac{f_0^2}{N^2} [\partial_t + \bar{u}ik + \alpha] \partial_{zz} \Psi + \frac{f_0^2}{N^2} \partial_z \alpha \partial_z \Psi \\
& + \left[-\bar{u}ik (k^2 + l^2) + \partial_y \bar{q}ik + \frac{f_0^2}{N^2} \left(-\frac{ik}{4H^2} \bar{u} + \frac{1}{2H} \partial_z \alpha - \frac{\alpha}{4H^2} \right) \right] \Psi,
\end{aligned}$$

$$\begin{aligned}
(2.12) \quad 0 = & -l^2 \partial_t U + \frac{f_0^2}{N^2} \left[-\frac{1}{H} \partial_t + \partial_z \alpha - \frac{\alpha}{H} \right] \partial_z U \\
& + \frac{f_0^2}{N^2} [\partial_t + \alpha] \partial_{zz} U + \frac{f_0^2}{N^2} \left[\frac{\alpha}{H} - \partial_z \alpha \right] \partial_z U_R \\
& - \frac{f_0^2}{N^2} \alpha \partial_{zz} U_R - \frac{f_0^2}{N^2} kl^2 \epsilon e^{z/2H} \text{Im} [\Psi \partial_{zz} \Psi^*] ,
\end{aligned}$$

where Im is the imaginary part and Ψ^* is the complex conjugate of Ψ . Here, N^2 ($= 4 \cdot 10^{-4} \text{ s}^{-2}$) is the buoyancy frequency squared; f_0 ($= 1.26 \cdot 10^{-4} \text{ s}^{-1}$) is the Coriolis parameter at 60°N ; β ($= 1.14 \cdot 10^{-11} \text{ s}^{-1} \text{ m}^{-1}$) is the meridional derivative of the Coriolis parameter at 60°N ; and ρ_0 ($= e^{-(z-z_0)/H} \text{ kg m}^{-3}$) is the standard atmospheric density.

The model is integrated using a semi-implicit time-stepping scheme discussed in Holton and Mass (1976), while the vertical derivatives, for any variable M , are defined as

$$\begin{aligned}
\frac{\partial M}{\partial z} &= \frac{M_{j+1} - M_{j-1}}{2\Delta z} , \\
\frac{\partial^2 M}{\partial z^2} &= \frac{M_{j+1} - 2M_j + M_{j-1}}{\Delta z^2} ,
\end{aligned}$$

where $j = 0, 1, \dots, J$, for $J+1$ vertical levels and J corresponds to $z_T = J\Delta z$.

2.1.2. REFORMULATION AND PARAMETERS.

As described above, the standard form of the Holton-Mass model utilizes bottom boundary geopotential perturbations to force waves. Yet, geopotential perturbations are not good proxies for wave activity fluxes. Edmon et al. (1980) showed that these fluxes are more appropriately captured by Eliassen-Palm (EP) flux vectors. From the transformed Eulerian mean (TEM) formulation of Andrews and McIntyre (1976), it follows that the convergence of the EP flux vector is the primary wave forcing on the zonal mean zonal wind. As the waves of interest for forcing of SSWs have tropospheric origins, observational studies on this subject principally analyze either upward EP fluxes or the dominant component of these upward EP fluxes, eddy meridional heat fluxes (e.g. Coy et al. 1997, Waugh et al. 1999, Newman et al. 2001, Limpasuvan et al. 2004). Upward EP flux then is the better proxy for studying wave forcing in the Holton-Mass model. Here we introduce a reformulation of the Holton-Mass model so that the bottom boundary wave forcing becomes a specified upward wave activity flux (EP flux).

Our incoming wave activity specification is achieved by simply prescribing the (time-dependent) geopotential perturbation at the bottom boundary such that the desired upward EP flux is produced. Upward EP flux in this model takes the form

$$(2.13) \quad F^{(z)} = \frac{\rho_0^2 R f_0}{N^2 H} \overline{v'T'}$$

where z_0 is bottom boundary height, R is the gas constant, and the other parameters are as above. The final term is the eddy meridional heat flux, the components of which are eddy

meridional wind and eddy temperature. These two fields may be cast as

$$(2.14) \quad v' = \partial_x \psi'$$

$$(2.15) \quad T' = \frac{H f_0}{R} \partial_z \psi',$$

where ψ' is the perturbation streamfunction. Together with the wave assumption from above, the heat flux may be written as

$$(2.16) \quad \overline{v'T'} = \delta_{vT} e^{z/H} \text{Im} \{ \Psi^* \partial_z \Psi \} ,$$

where $\delta_{vT} = \frac{kHf_0\epsilon}{2R}$ is a constant coefficient. As before, we denote the real and imaginary parts of Ψ as X and Y , respectively. At the bottom boundary (subscript 0), we utilize a coordinate shift such that the imaginary part of Ψ there is 0. Then the discretized form of the bottom boundary heat flux is

$$(2.17) \quad \begin{aligned} \overline{v'T'}_0 &= \delta_{vT} \text{Im} \left\{ \Psi_0^* \frac{\Psi_1 - \Psi_0}{\Delta z} \right\} \\ &= \frac{\delta_{vT}}{\Delta z} \text{Im} \{ X_0 (X_1 + iY_1 - X_0) \} \\ &= \frac{\delta_{vT}}{\Delta z} X_0 Y_1 , \end{aligned}$$

where subscript 1 refers to the first model level above the bottom boundary. Rearranging and recasting in terms of Ψ ,

$$(2.18) \quad \begin{aligned} \Psi_0 &= \frac{\Delta z \overline{v'T'}_0}{\delta_{vT} \text{Im} \{ \Psi_1 \}} \\ &= \frac{2N^2 \Delta z}{k\epsilon f_0^2 \text{Im} \{ \Psi_1 \}} F^{(z)} . \end{aligned}$$

Thus, the functional method by which we set the bottom boundary EP flux is to solve at every time step for the bottom boundary geopotential height perturbation which satisfies the EP flux we specify.

One may readily see in (2.18) that Ψ_0 must be infinitely large when $\text{Im}\{\Psi_1\}$ is exactly zero and the specified incoming EP flux is nonzero. This numerical problem may be overcome through (at least) two methods. The first is to initialize the model streamfunction field with a infinitesimal amplitude. The second method is to inject a small geopotential perturbation in the time step prior to nonzero specified EP flux. Over a large range of experimentation, utilizing either initialization method has succeeded in avoiding unrealizable solutions.

To semi-implicitly integrate the model fields from a given time step, we must know both the bottom boundary geopotential perturbation and the first model level streamfunction at the next model time step. This requires an iterative process which solves for these fields to a specified precision. We use a precision of 0.01 m; results presented above are not found to depend strongly on this precision.

Our model control parameters are as follows. We use wavenumber 2 for the streamfunction perturbations' assumed zonal structure. The radiative equilibrium zonal wind is determined by the constant bottom boundary zonal wind of 15 m s^{-1} and the height-independent radiative equilibrium vertical shear of $1.5 \text{ m s}^{-1} \text{ km}^{-1}$. The model bottom is at 10 km, the model top is at 60 km, and the vertical spacing is 500 m. Time stepping is 900 s. Integrations with deviations from these parameters are noted. All other parameters – such as Newtonian damping – are as in Holton and Mass (1976). The following results are found to be insensitive either to raising the model lid or to inclusion of a Rayleigh friction layer near the model lid.

2.2. STEADY BOTTOM BOUNDARY INCOMING WAVE ACTIVITY FLUX

To begin, we show the model response to both steady bottom boundary geopotential perturbation and upward EP flux. Displayed in Figure 2.1 (a) is the zonal wind (shaded) and EP flux (contoured) for constant bottom boundary geopotential forcing over a range of time far from initialization. We choose the amplitude (300 m) such that the model zonal wind is within the perpetual oscillating (SSW-like) state.

An initial wave propagates into the model (not shown) and dissipates, decelerating the zonal wind. The wave-mean flow positive feedback is then triggered, allowing greater wave propagation. Wind reversal is subsequently forced, producing a critical layer to propagation. This terminates the positive feedback and induces the negative feedback¹. The critical layer – above which the upward EP flux amplitudes decay to zero – extends to roughly 18 km, resulting in wave reflection and net downward wave activity flux to the model bottom boundary. As the zonal wind restores, the model background state again becomes susceptible to wave propagation, and the process continues. The classic theory on the forcing of SSWs holds in these results.

There are both external and internal features of the wave-mean flow positive feedback manifested in Figure 2.1 (a). The external feedback process is associated with the large burst of upward EP fluxes at the bottom boundary. One may think of this as a fluxing upwards, by a favorable mean state, of wave activity from some source existing below the model bottom boundary. The internal part of this feedback behaves in the same way, but relies only on that wave activity which exists or may be generated within the model domain.

¹Hereafter, references to wave-mean flow feedback refers only to the positive wave-mean flow feedback which acts to amplify any initial wave perturbation.

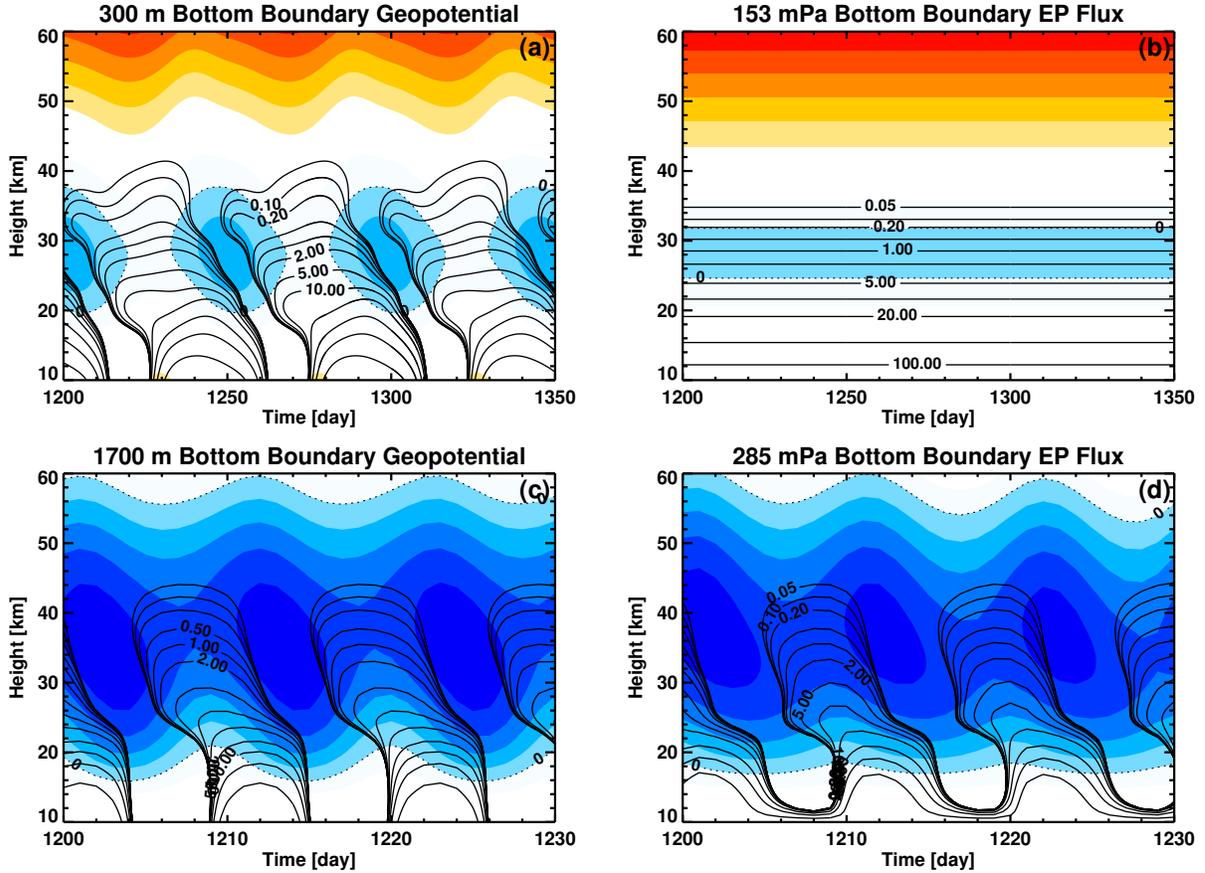


FIGURE 2.1. (a),(c) Zonal wind (shaded) and contours of upward EP flux (solid black) for the geopotential model forced with a constant bottom boundary perturbation of 300 m and 1700 m, respectively. (b),(d) Same as (a) and (c), but the EP flux model forced with constant bottom boundary upward EP flux of amplitude such that the mean bottom boundary geopotential is equal to that from the respective geopotential model integrations. The zonal wind contour spacing is 5 m s^{-1} in all plots, with the 0 m s^{-1} contour labeled and dashed. The upward EP flux contours here and throughout this manuscript are on a logarithmic scale (e.g. 0.1, 0.2, 0.5, 1, 2, 5, 10, etc.). Only positive EP flux contours are shown.

An example of the latter may be seen in Figure 2.1 (a) between days 1310 and 1320, where there is EP flux divergence (wave generation) above 18 km.

The external feedback process here may be thought of as the response of tropopause-level waves to perturbations in the state of the stratosphere by prior waves. While the real atmosphere manifestation of this feedback process will certainly be limited in available tropospheric wave activity, the source of wave activity below the model domain for the standard

Holton-Mass model is potentially limitless. The external feedback process represented in the Holton-Mass model forced with bottom boundary geopotential perturbations is therefore likely unrealistic. Hence, it appears desirable to isolate the internal feedback in models with idealized bottom boundary conditions, such as the Holton-Mass model. As the external feedback mechanism does not operate where the incoming wave activity flux is specified, our EP flux bottom boundary condition does just that. Though we are motivated to recast the Holton-Mass model by the desire to specify upward bottom boundary wave activity flux, isolation of the internal feedback mechanism comes along as a conceptual benefit.

Shown in Figure 2.1 (b) is the zonal wind (shaded) and upward EP flux (log contoured) for constant bottom boundary EP flux forcing. The amplitude of this specified upward EP flux (153 mPa) is chosen such that the mean bottom boundary geopotential is 300 m (i.e. the same as in (a)). Rather than the upward EP flux and zonal wind oscillating in the model interior, both fields are now steady with a maintained layer of easterlies between 24 and 32 km. This should not come as a surprise: steady wave forcing, if based on a conservative wave property (such as the EP flux), should lead to steady zonal wind response even in cases of very large wave forcing. Analysis of the representative index of refraction (see Harnik 2009) shows that the layer between 30 and 35 km represents a critical layer to wave propagation. Waves are no longer able to propagate beyond this critical layer, leaving only the imposed radiative damping and small dissipation from evanescent wave amplitudes to force the zonal wind above.

As discussed above, zonal wind oscillation in the original Holton-Mass model is possible because an upward EP flux oscillation ensues in the lower and midlevels of the model. With a prescribed constant bottom boundary upward wave activity flux, this oscillation is necessarily dampened at lower model levels. The lower level upward wave activity flux is

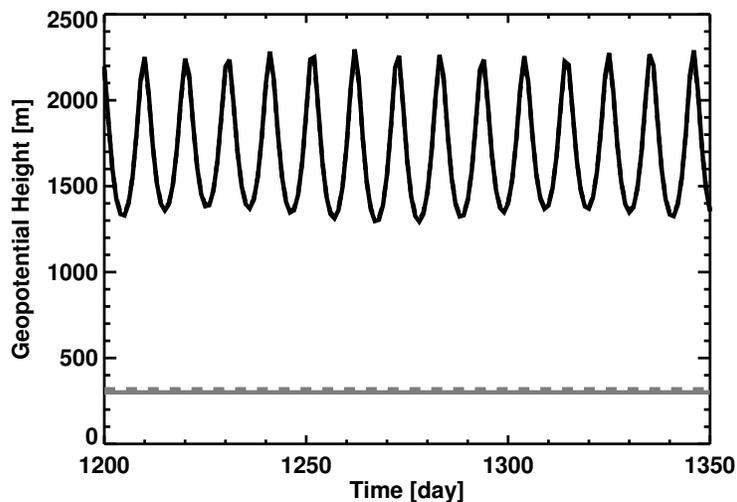


FIGURE 2.2. The geopotential perturbations for the bottom boundary (solid) and for the next highest model level (dashed) plotted for the constant upward EP flux model integrations with 153 mPa EP flux magnitude (gray) and 285 mPa EP flux magnitude (black). Note that the black dashed line is indistinguishable from the full black line.

sufficient to overcome the effect of wave reflection that results from the critical layer starting at 30 km, preventing the system from reaching an oscillatory state.

An oscillating state is still possible in the EP flux model, but only for extremely large upward EP flux amplitudes. Shown in Figure 2.1 (c) and (d) are the same fields as in (a) and (b), but for a constant geopotential perturbation and constant upward EP flux of 1700 m and 285 mPa, respectively. Here, both models oscillate², but about a zonal wind less than 0 m s^{-1} . With the zonal winds through a majority of the model depth consistently easterly, this is far from the observed variability. Yet (d) shows that the internal feedback alone is in principle able to drive the type of perpetual oscillations observed in the geopotential model.

Figure 2.2 shows how the bottom boundary geopotential and next highest model level geopotential vary as a function of time in the two constant EP flux model integrations

²In fact, the zonal wind vacillates in (d), as shown by the slight drift in zonal wind values atop the periodic variability.

(Figure 2.1 (b) and (d)). For the 153 mPa EP flux integration (gray), the geopotential at these two levels is steady by this point in the integration³. This is in line with the steady EP fluxes at all levels between the shown times. For the 285 mPa EP flux integration (black), the geopotential at both levels oscillates with the same periodicity as the upward EP flux. Note that the maximum (minimum) of this oscillation relates to the maximum (minimum) in upward EP flux (see Figure 2.1 (d)). Both levels oscillate in near unison, acting to keep the upward EP flux along the bottom boundary constant.

We further explore the effects of specifying the bottom boundary EP flux by analyzing the stable states of this new setup. A particular characteristic of the original Holton-Mass model is that a bistable steady state with respect to constant bottom boundary geopotential perturbation exists within the zonal wind field⁴. As explored in previous work with this model (Chao 1985, Yoden 1987) and a simpler prototype (Ruzmaikin et al. 2003, Birner and Williams 2008), when the steady geopotential perturbation lies between the two thresholds of the bistable state, the zonal wind stably exists either in a lightly perturbed (near to radiative equilibrium) or in a strongly perturbed (far from radiative equilibrium) state. The bounds of this bistable regime represent pitchfork bifurcation points where one of the stable solutions ceases to exist.

Such a behavior may be found by very slowly ramping up and subsequently ramping down the bifurcation parameter (i.e. bottom boundary geopotential or EP flux) so that the zonal wind remains in a quasi-steady state (Birner and Williams 2008). Once the bifurcation point is passed, a sudden transition between stability states will occur. Shown in Figure 2.3 (a) is the resultant zonal wind from one such integration with the EP flux model, where the

³We note that there is some transience during initialization of the model.

⁴Note that one branch of this zonal wind steady state is not strictly steady, but rather oscillating. However, because this state is stable and the oscillations regular, it has been referred to as steady.

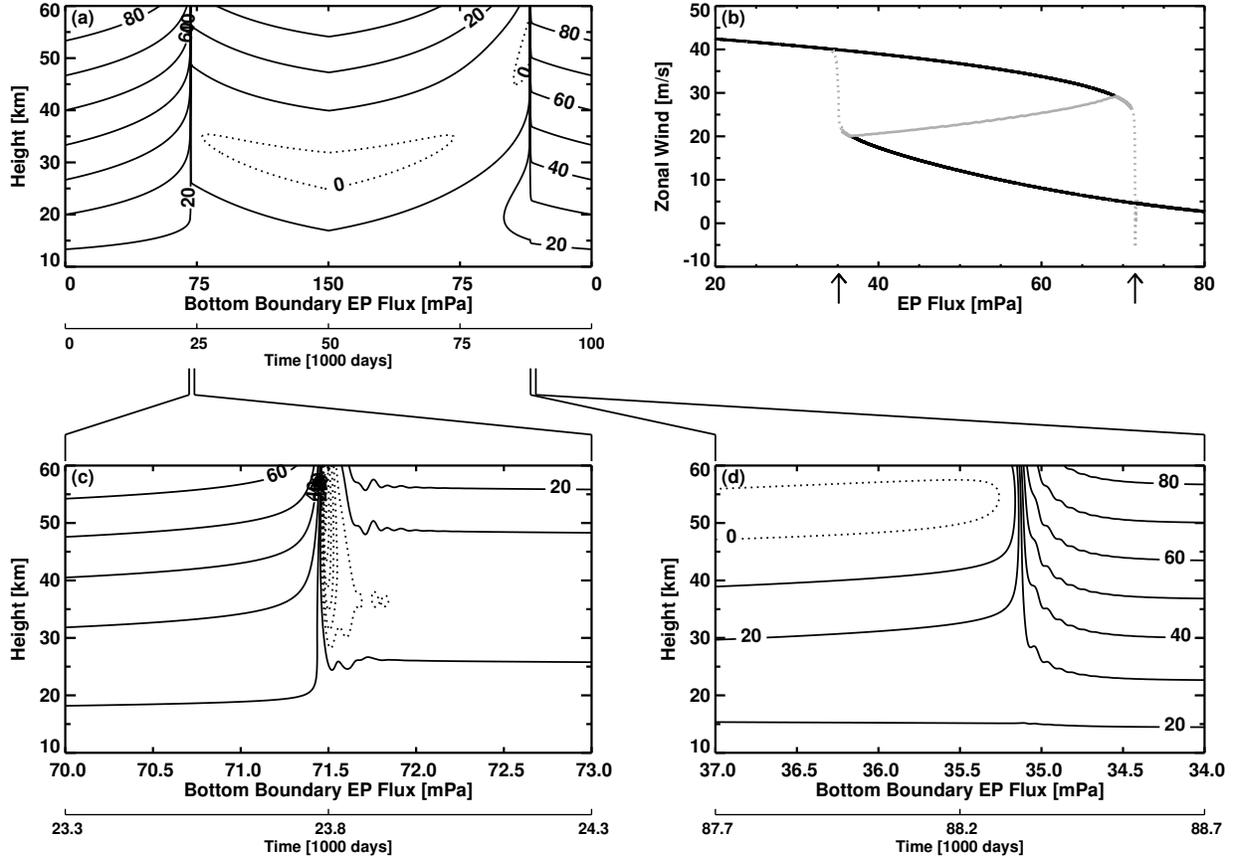


FIGURE 2.3. (a) EP flux model zonal mean zonal wind resulting from linearly varying EP flux of magnitude plotted along the abscissa. Contour intervals are 10 m s^{-1} . (b) The zonal wind at the 30 km level. Solid black curves are stable equilibrium solutions while solid gray curves are unstable equilibrium solutions. Transient solutions are shown as gray dots. Arrows indicate the EP flux values at which the bifurcations between stable states lie. Zonal wind initialized within the bistable regime and above the unstable solution will stabilize to the upper steady state, while zonal wind initialized within the bistable regime and below the unstable solution will stabilize to the lower steady state. (c) The same as in (a), but zoomed in on the bifurcation point near day 23800. (d) As in (a), but zoomed in on the bifurcation point near day 88200.

durations of both the ramp up and the ramp down are 50000 days. For this experiment, the maximum incoming upward EP flux is 150 mPa. This is below the values of steady upward EP flux which result in oscillating zonal wind.

At the start of the integration, the zonal wind is initialized at the radiative equilibrium state. As the bottom boundary EP flux increases, the zonal wind is increasingly, yet

smoothly, pushed away from this radiative state. Near day 23800, the incoming EP flux becomes sufficiently large so as to force a transition from this lightly perturbed zonal wind state to a different, strongly perturbed state. Shown in Figure 2.3 (c) is a zoomed in view of this transition, which appears to be a prototype SSW: wind reversal starting in the upper stratosphere propagates downward to the middle stratosphere. Beyond this transition, the zonal wind evolution again becomes smooth, implying that the model again resides within a stable state. In this (quasi-)steady state, the zonal wind is near to (or less than) zero in the layer between 25 and 35 km – corresponding to a critical layer for wave propagation – and increases above this layer due to the imposed radiative damping and very weak wave forcing.

As the EP flux is linearly decreased beyond 50000 days, the critical layer shifts upwards along with the height of the minimum zonal wind. Near day 88200, the zonal wind state rapidly transitions from the strongly perturbed to the weakly perturbed state. As shown in Figure 2.3 (d), this transition back to the lightly perturbed steady state is analogous and opposite to the transition in (c). Beyond this point, the zonal wind gradually relaxes towards the radiative equilibrium zonal wind as the wave forcing wanes.

That the two transitions occur at different values of bottom boundary EP fluxes shows that the model has a bistable regime. In order to better visualize this, Figure 2.3 (b) displays the zonal wind at 30 km from this integration. Given that the bistable regime exists within the EP flux model, the external wave-mean flow feedback is not a necessary component of this bistability. Remarkably, two stable zonal wind solutions exist for the same amount of incoming wave activity flux.

We test the stability of the system by initializing the zonal wind at points between the two zonal wind states and determining to which state the model equilibrates. Through this,

we identify where multiple equilibrium solutions exist in the model and whether these are stable (solid black) or unstable (solid gray). Between incoming upward EP flux ranges of $\sim 35\text{-}36$ mPa and $\sim 69\text{-}71$ mPa, only two equilibrium solutions exist. In the former two-solution regime, we find that the upper branch is a stable solution while the flipped branch is an unstable solution, and vice versa for the latter two-solution regime. Between ~ 36 and ~ 69 mPa, three equilibrium solutions exist: two stable solutions and one unstable solution. In this three-solution regime, model initializations with zonal wind greater than (less than) the unstable solution will equilibrate to the upper (lower) steady state solution. Given that the parameter space where solutions stabilize to the strongly perturbed state is larger than the parameter space where solutions stabilize to the weakly perturbed state, it is apparent that the strongly perturbed state is the more stable solution for this model setup.

Figure 2.4 shows the bistability of the geopotential model under the same experiment, but with maximum bottom boundary geopotential of 500 m. As in Figure 2.3, the model remains steadily within the lightly-perturbed equilibrium solution until the bottom boundary forcing becomes sufficiently strong so that this equilibrium solution no longer exists. The zonal wind then equalizes to and remains at the strongly-perturbed equilibrium solution until the bottom boundary geopotential is again small. Note that as previously discussed, the strongly-perturbed equilibrium solution is actually a vacillating zonal wind (and EP flux) state.

Just like in the EP flux model, the two equilibrium solutions in the geopotential model exist at overlapping parameter space, indicating that the bistability exists in both model formulations. Note that we do not perform the prior stability analysis for Figure 2.4 because the external feedback dominates the evolution, nearly always leading to stabilization to the strongly-perturbed regime. Following Yoden (1987), this should not be the case, suggesting

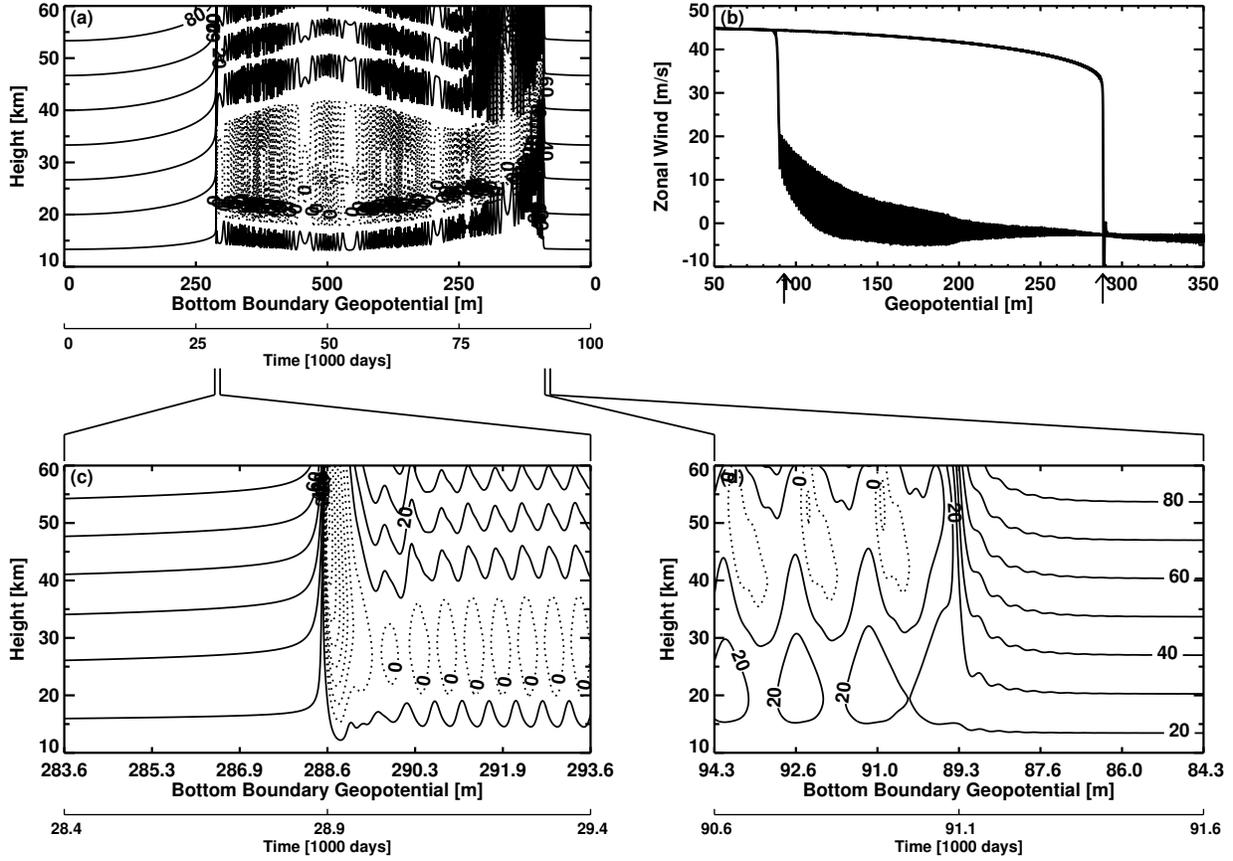


FIGURE 2.4. As in Figure 2.3, but for the geopotential model. Stability has not been assessed for (b).

that the exact form of the stability analysis we performed for the EP flux model is not suitable for the geopotential model. Rather, a linear stability analysis such as that in Yoden (1987) would be required but is not performed here.

With the above bistability in mind, we attempt to force a transition between the stable states of the system, similar to Birner and Williams (2008) and Hardiman and Haynes (2008). Here, we integrate both models with equal amplitudes of constant bottom boundary wave forcing, and we impose a zonal wind deceleration in the model stratosphere independent of the waves. This deceleration was chosen such that the wave-mean flow feedback would allow wave propagation sufficient to drive a transition between stable states, and may be thought

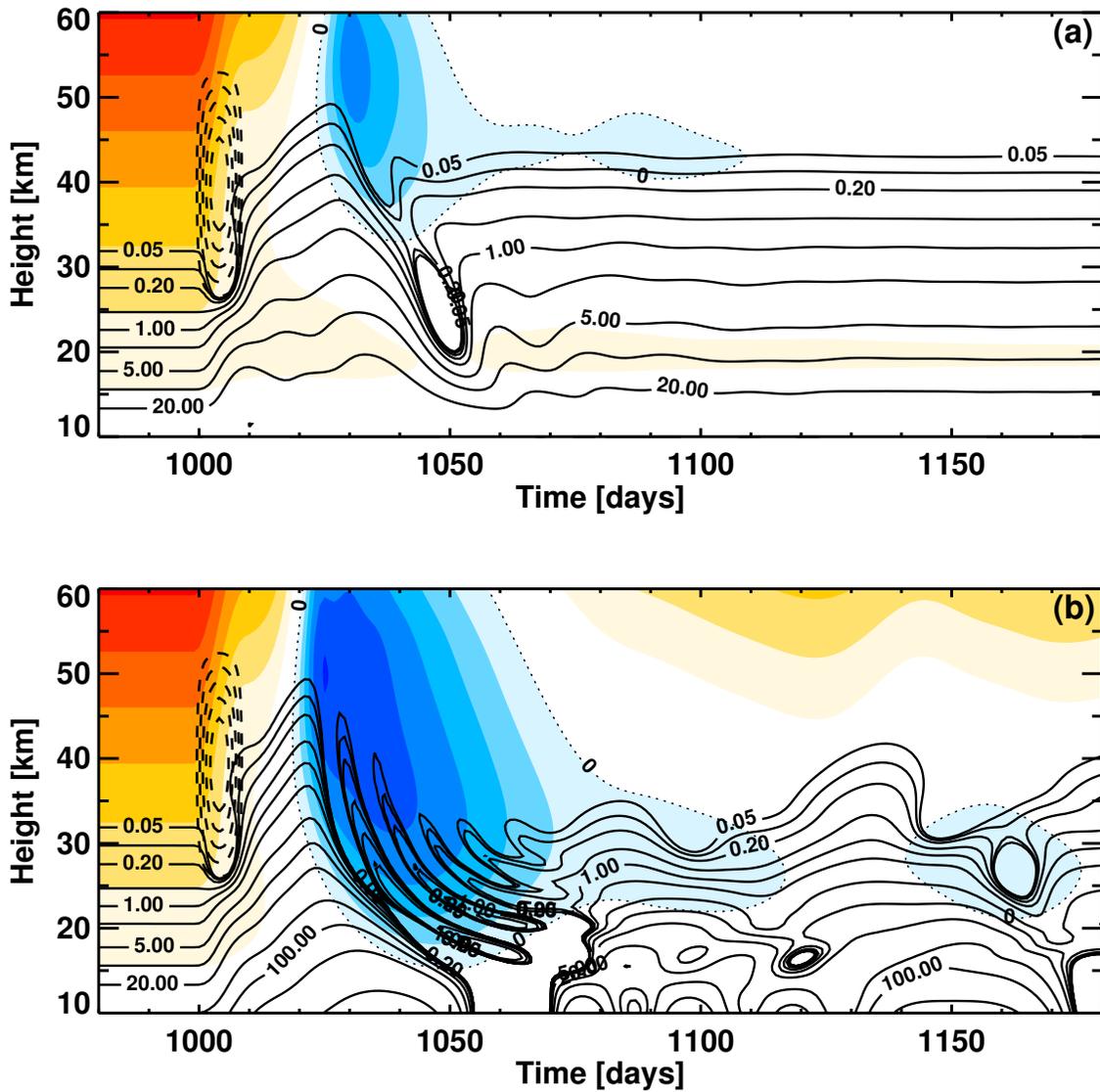


FIGURE 2.5. (a) Zonal mean zonal wind from the EP flux model forced with constant bottom boundary EP flux of magnitude 50 mPa. The dashed black contours after 1000 days shows imposed zonal wind deceleration; this deceleration is a sinusoid in height and time, and applied between 25 and 55 km, over a period of 10 days, and at a maximum amplitude of $30 \text{ m s}^{-1} \text{ day}^{-1}$. Upward EP flux values are log contoured in solid black, imposed deceleration contour spacing is $5 \text{ m s}^{-1} \text{ day}^{-1}$, and zonal wind contour spacing is 10 m s^{-1} . The 0 m s^{-1} wind contour is dotted and labeled. (b) Same as in (a), but from the standard Holton-Mass model with constant bottom boundary geopotential forcing of 273 m.

of as resulting from unresolved waves – e.g., gravity waves – or from prior planetary wave breaking.

Shown in Figure 2.5 is the zonal wind evolution following the imposed deceleration and the upward EP flux, from both (a) the EP flux model and (b) the geopotential model. Incoming upward EP flux is held constant at 50 mPa, while incoming geopotential perturbation is held constant at 273 m. For reference, the bistable regime in the geopotential model exists between bottom boundary geopotential perturbations of ~ 92 m and ~ 287 m. Note that the initial zonal winds and upward EP fluxes are identical in both integrations.

Up to day 1000 in both integrations, the models are steady and within the lightly perturbed (near-radiative equilibrium) regime. After the onset of the imposed external deceleration at day 1000, the upward EP flux increases in the lower model levels. This anomalous upward EP flux continues to grow at all levels and forces the zonal wind to transition into the strongly perturbed state. Following the transition in both models, the zonal wind stabilizes to the strongly perturbed regime.

Examining the SSW-like transition in the EP flux model shows how the internal feedback causes enhanced wave amplitudes in the middle and upper model levels. Following the imposed deceleration, EP flux convergence decreases on average by $\sim 50\%$ in the lowermost 10 km of the model, allowing greater vertical penetration of wave amplitudes from the bottom boundary. Furthermore, following the imposed deceleration and prior to the zonal wind reversal, the index of refraction is positive up to 27 km – 6 km deeper than prior to the imposed deceleration – indicating considerably deeper wave propagation.

In the geopotential model, the index of refraction evolves in a nearly identical manner as in the EP flux model. The main difference between the integrations is that the bottom boundary upward EP flux increases by over an order of magnitude from the initial, constant value. The primary effect of the external feedback is to provide a substantial source of additional wave activity to the model domain. In further experimentation, it was found that

this source is strong enough that the imposed deceleration in the geopotential model need only be a factor of one-third as strong as that for the EP flux model.

A question then arises: is the induced, transient upward EP flux along the bottom boundary of the geopotential model sufficient to drive the SSW-like transition without the imposed zonal wind deceleration which originally induced it? By injecting the transient bottom boundary upward EP flux from the geopotential model integration into the EP flux model, we find that this is indeed the case (not shown). Given that such transient wave forcing is imperative to understanding wave-mean flow processes in the real world, we now focus attention on the two feedback processes under time-dependent forcing.

2.3. TIME-DEPENDENT BOTTOM BOUNDARY INCOMING WAVE ACTIVITY FLUX

2.3.1. ZERO BASELINE.

In this section we examine SSW-like transitions in our versions of the Holton-Mass model using time-dependent forcing, with particular emphasis on the wave-mean flow positive feedback. Following Sjoberg and Birner (2012), we determine such a transition to have occurred when the zonal wind at all model levels above and including 30 km is equal to or less than a threshold wind speed for each respective level. We take this threshold zonal wind profile to be the zonal wind at day 24000 in Figure 2.3 (c), i.e. where the zonal wind is within the strongly perturbed steady state. While 30 km is an arbitrary choice of lower level for our threshold, it is near to the 10 hPa pressure surface. This is a commonly used level for identification of major SSWs (e.g. McInturff 1978, Charlton and Polvani 2007). We found that the qualitative nature of the following results remains intact across a range of lower levels for this definition.

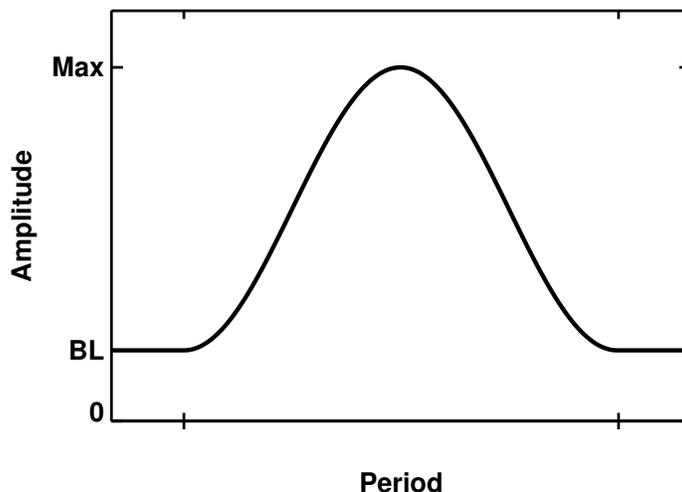


FIGURE 2.6. Schematic of the sine-squared EP flux forcing utilized in the time-dependent forcing experiments. The time scale and baseline EP flux (labeled “BL”) are specified, while the maximum amplitude (“Max”) is stepwise increased until a stability state transition is forced. With amplitude A and time scale P known, total incoming EP flux is given by $F_{int} = \int_P A \sin^2\left(\frac{t}{P}\pi\right) dt = \frac{1}{2}AP$.

We use sine-squared pulses of bottom boundary upward EP flux or geopotential perturbation which are shown schematically in Figure 2.6. For our idealized pulses, the baseline (initial and final) bottom boundary upward EP flux and the time scale of the pulse are specified. The amplitude is increased in small steps until a SSW-like transition has occurred for the chosen pulse time scale.

Shown in Figure 2.7, as a function of bottom boundary upward EP flux pulse time scale, are the minimum amplitudes (purple) and associated total incoming EP flux (green) necessary to force a transition for zero baseline EP flux. To aid in the analysis of these results, the dashed purple line marks the incoming upward EP flux amplitude for the bifurcation point at which only the strongly perturbed solution exists. Note that this is a steady state amplitude and is thus independent of time scale.

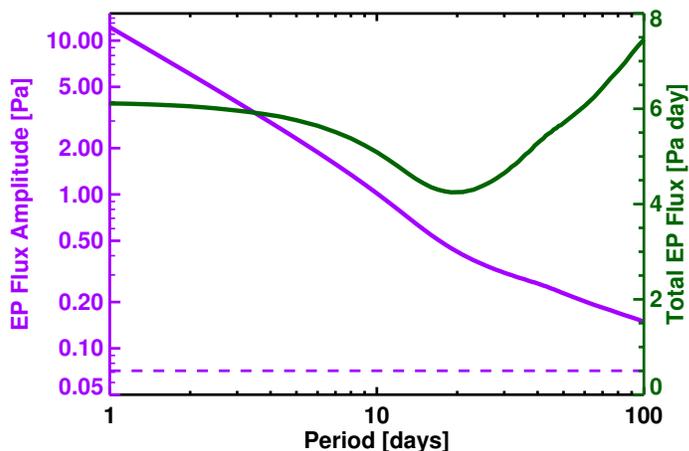


FIGURE 2.7. EP flux amplitude (purple) and total EP flux (green) at the bottom boundary necessary to force a SSW-like transition in the EP flux model, as a function of the time scale of sine-squared pulses. The dashed purple curve gives the bifurcation point amplitude where only the strongly perturbed state exists. The amplitude ordinate and the abscissa are both plotted logarithmically.

For pulses which have time scales such that the forcing is quasi-steady (much longer than 100 days; not shown), the maximum amplitude is near to the steady state solution. As these pulses become more transient (going from right to left in the figure), the amplitudes increase while the total incoming EP flux decreases, both at roughly linear rates proportional to the pulse time scale. This behavior continues down to ~ 20 days, at which point the total incoming EP flux minimizes. At still shorter time scales, the EP flux amplitudes of the pulses continue to increase with decreasing time scale, though at a rate more quickly than for time scales greater than 20 days. This manifests as an asymptotic approach in the total incoming EP flux to a value larger than that near 20 days.

If the prescribed radiative equilibrium zonal wind (which is also the initialized zonal wind state) is changed, the pulse time scale of the total incoming EP flux minimum likewise changes. For instance, by adding 5 m s^{-1} to the entire radiative equilibrium zonal wind profile, the minimum moves to ~ 15 days. By subtracting 5 m s^{-1} from the entire profile,

the minimum moves to ~ 30 days.⁵ No such changes in the location of this minimum are found by varying the time scales of radiative damping. This suggests that the minimum is a manifestation of the wave-mean flow positive feedback and not due to externally prescribed (e.g. radiative) time scales.

To further understand how the wave-mean flow feedback affects the time scale of minimum total incoming EP flux, we remove the time-dependency of zonal wind terms affecting the waves. That is, for the streamfunction Ψ evolution equation (see Holton and Mass 1976),

$$(2.19) \quad (\partial_t + ik\epsilon U) \left[- (k^2 + l^2) \Psi + \frac{f_0^2}{N^2} \left(\partial_{zz} \Psi - \frac{1}{4H^2} \Psi - 2 \frac{\partial_z N}{N} \left\{ \partial_z \Psi + \frac{1}{2H} \Psi \right\} \right) \right] \\ + \alpha \frac{f_0^2}{N^2} \left[\partial_{zz} \Psi - \frac{1}{4H^2} \Psi + \left(\frac{\partial_z \alpha}{\alpha} - 2 \frac{\partial_z N}{N} \right) \left(\partial_z \Psi + \frac{1}{2H} \Psi \right) \right] \\ + \partial_y \bar{q} i k \Psi = 0,$$

we set the zonal wind U and the meridional gradient of mean potential vorticity $\partial_y \bar{q}$ to their initial values at all time steps (the form of $\partial_y \bar{q}$ is given below in Eq. (2.20)). Figure 2.8 (a) shows the zonal wind anomalies in shading and the upward EP flux in log contouring for an integration of the standard EP flux model. Here the model is forced by a pulse of 10 days with amplitude 1.17 Pa, sufficient to drive a SSW-like transition. (b) shows the same fields, but for the EP flux model with no time-dependent internal feedback of the wind on the waves.

For the experiment shown in (a), the upward EP flux strongly dissipates, driving zonal wind reversal. In contrast, the experiment shown in (b) results in only weak EP flux dissipation of evanescent waves above a reflecting layer located near 21 km. This reflecting layer

⁵The magnitude of the minimum also varies with the radiative equilibrium zonal wind profile, being 5.10, 4.25, and 4.16 Pa day for the $+5 \text{ ms}^{-1}$ case, the control case, and the -5 ms^{-1} case, respectively.

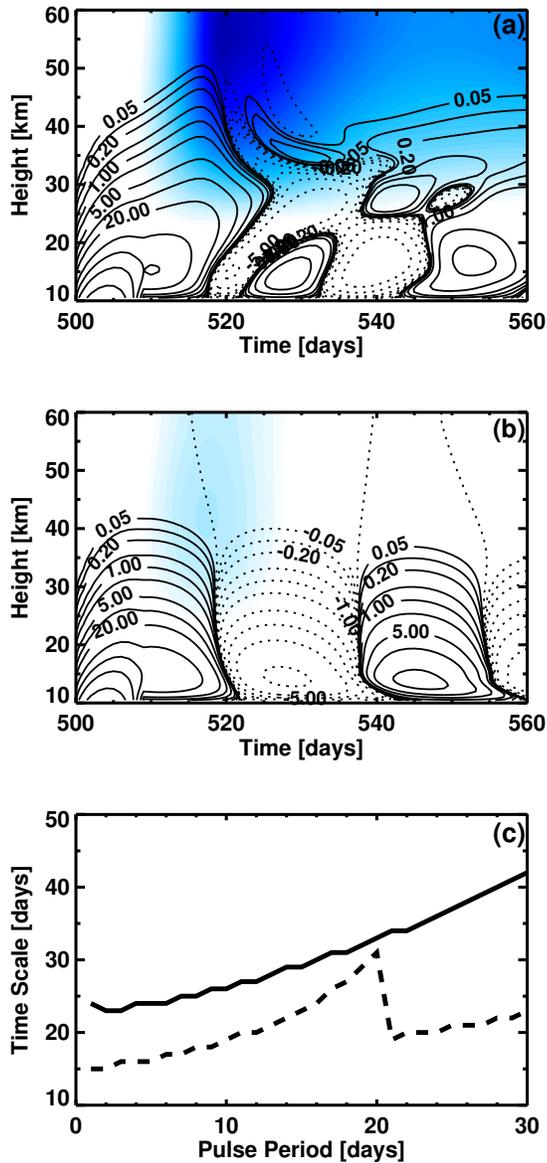


FIGURE 2.8. Zonal wind anomalies (shading) and log contoured upward EP fluxes for (a) the standard EP flux model and (b) the EP flux model with mean flow feedbacks on the wave turned off (see Section 3). Both models are forced by a 10 day pulse of bottom boundary EP flux with amplitude 1.17 Pa, sufficient to drive a SSW transition in (a). Positive EP flux contours are solid while negative EP flux contours are dotted. (c) Duration of upward EP flux pulse at 30 km as a function of bottom boundary pulse time scale for the standard EP flux model (solid) and for the EP flux model with zonal wind feedbacks on the waves turned off (dashed). The maximum amplitudes of the bottom boundary upward EP flux pulse are those which force a SSW-like transition for the given pulse time scale (see Figure 2.7).

results not from slow but from fast zonal winds. Because the internal feedback is fully active in (a), the upward EP flux amplitudes of the initial wave pulse are enhanced (relative to those in (b)) and the duration over which the initial upward EP flux pulse is greater than zero becomes longer with height, especially up to 30 km. In contrast, the duration over which the initial upward EP flux pulse is greater than zero in (b) decreases with height.

Despite the zonal wind feedback on the waves being independent of time in (b), the 10 day pulse of bottom boundary wave activity flux is organized onto an almost 20 day time scale within the model domain. To test what happens when the bottom boundary pulse time scale approaches the time scale onto which the model state organizes the waves (here, 20 days), we repeat the experiments shown in Figure 2.8 (a) and (b), but across a range of bottom boundary pulse time scales. The duration where the initial upward EP flux pulse at 30 km is greater than zero is shown in (c) as a function of bottom boundary pulse time scale for the standard EP flux model and for the EP flux model with no time-dependent feedback.

In the fully active feedback case, the duration of the initial wave activity pulse increases as a function of bottom boundary pulse time scale. In the time-independent feedback case, the duration of the initial wave activity flux pulse increases for bottom boundary pulse time scales up to 20 days. A sudden decrease in durations occurs for a 21 day pulse, indicating that the pulse durations aloft are actually shorter than the initial bottom boundary pulse in the time-independent feedback case. This results from internal EP flux reflection (e.g. between days ~ 520 -540 in (b)) suppressing positive upward EP fluxes.

The pulse duration behavior found here – increasing for bottom boundary pulse time scales up to 20 days and then suddenly decreasing – further validates that the ~ 20 day bottom boundary pulse time scale is preferred, given our particular model initial state. Increasing the radiative equilibrium (initial) zonal wind profile by 5 m s^{-1} results in this

jump instead occurring at a 15 day pulse period, while decreasing the radiative equilibrium zonal wind profile by 5 m s^{-1} moves this jump to a 30 day pulse period. As we previously found, these are exactly the periods at which the total incoming wave activity flux minimum occurs for the respectively altered radiative equilibrium zonal wind.

The above changes to the radiative equilibrium zonal wind produce concomitant changes to both the height of the model internal reflecting layer and the wave group velocity, where larger wind speeds produce a lower reflecting layer (in line with Charney and Drazin 1961) and larger group velocity. Thus with faster (slower) background flow, we expect internal propagation to be faster (slower) and over a shorter (longer) distance within the model domain. Note that the simulated waves here have vertical wavelengths of $\sim 50 \text{ km}$. The aforementioned reflecting level (near 21 km in our control setup, i.e. 11 km above the model's bottom boundary) therefore severely limits the vertical propagation depth of the waves relative to their vertical wavelengths, with the bulk of the vertical domain leading to evanescence. We note again that these results are insensitive to the height of the model lid or to inclusion of upper level Rayleigh friction – i.e. insensitive to wave reflection from the model lid.

The minimum in total incoming wave activity flux seems to occur because we force the model with a pulse of duration at which the wave-mean flow positive feedback acts most strongly. We therefore hypothesize that pulses shorter than ~ 20 days are not as efficient at forcing the zonal wind and thus require increasingly larger maximum amplitudes of upward EP flux (and larger associated total incoming EP flux). Pulses longer than ~ 20 days are also not as efficient at forcing the zonal wind, but have the added effect of larger integrated radiative damping. Compounding the above functional behaviors of the necessary amplitudes and total incoming EP flux leads us to conclude, as in Sjoberg and Birner (2012), that there is a preferred time scale of the order of 10 days for wave forcing associated with SSWs.

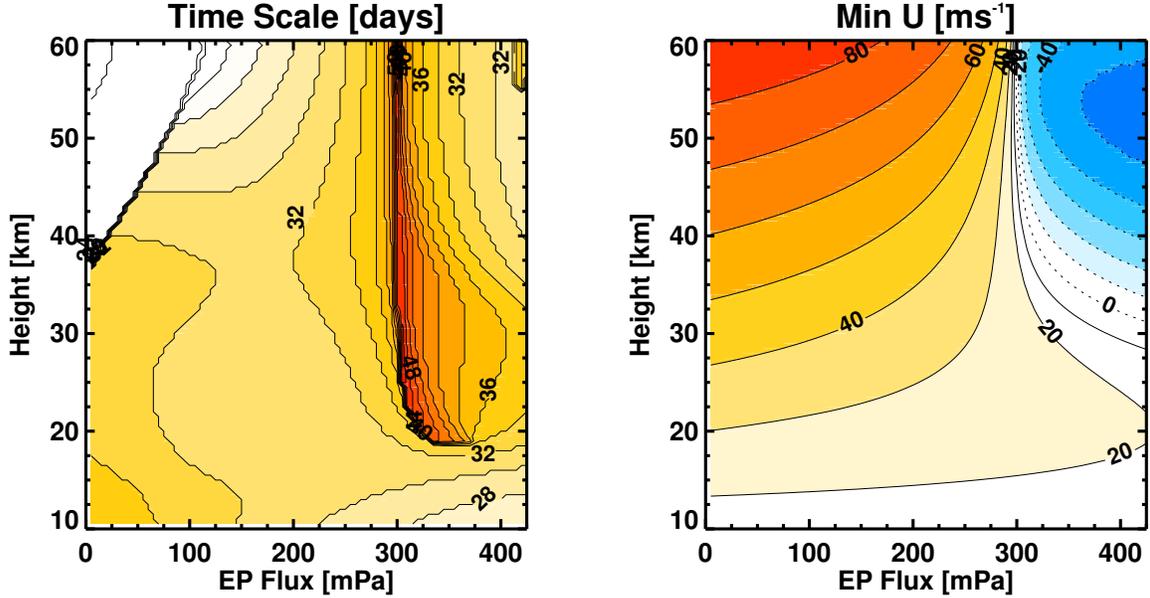


FIGURE 2.9. Pulse duration (left) and minimum zonal wind (right) as a function of pulse amplitude and of height. The model was forced in individual integrations with a 20 day pulse of incoming upward EP flux with amplitudes equal to that along the abscissa. The pulse duration contour spacing is 2 days and the zonal wind contour spacing is 10 ms^{-1} .

An additional feature of the positive feedback falls out from this analysis. As shown by Figure 2.8, when the model is forced by incoming upward EP flux pulses with sufficiently large amplitude to cause a SSW-like transition, the duration of the initial, bottom boundary pulse increases with height. Contrarily, when the pulses of incoming wave activity flux are of small amplitude, the durations decrease with height. To further elucidate this, Figure 2.9 shows the pulse durations and minimum zonal winds produced in integrations of the EP flux model forced by 20 day pulses of incoming upward EP flux. The pulse amplitudes in the different integrations span a range of values from 0 mPa to 425 mPa (the amplitude at which a SSW-like transition is forced for a 20 day pulse).

For amplitudes at which there is little deceleration of the zonal wind ($< \sim 200 \text{ mPa}$), the durations decrease with height throughout the vertical column. At larger amplitudes, the zonal wind is decelerated to the point of reversal with the model domain. Along with

this wind reversal is a clear increase in the duration of the pulse with height. This variation in height of the durations of the initial pulse are then a way to diagnose whether strong positive feedback occurs. Note that the increase in height is not monotonic with height, but instead peaks below the level of wind reversal. The lower level peak shows that the lower stratospheric wave activity fluxes are large and sustained when the feedback is strong. In the upper model levels, zonal wind reversal inhibits propagation, leading to relative decreases in the pulse durations. This suggests that an appropriate upper level for this diagnostic is within the lower stratosphere, well removed from the levels at which the zonal wind reverses.

2.3.2. NONZERO BASELINE.

The above modeling results exclusively consider the EP flux model; accordingly, only the internal wave-mean flow feedback is active. To gain insight into how the external wave-mean flow positive feedback interacts with the transient wave forcing utilized here, we perform the above sine-squared pulse experiment in both the geopotential and EP flux models with nonzero baseline. For a zero baseline in the geopotential model the external feedback can only operate during the pulse time scale (i.e. when the bottom boundary geopotential is greater than zero). However, a nonzero baseline bottom boundary geopotential allows the external feedback to operate even once the pulse has ceased.

We find that the external feedback is strengthened as the baseline is increased. Furthermore, the external feedback allows small amplitude pulses – which are insufficient to force a SSW-like transition on their own – to condition the model state in such a way that sufficient wave activity is fluxed into the model to complete the transition. As the baseline is increased, the geopotential model is more able to utilize the limitless wave source below the model boundary. Thus the external feedback more strongly impacts the evolution of the wave activity flux within the geopotential model as the baseline is increased.

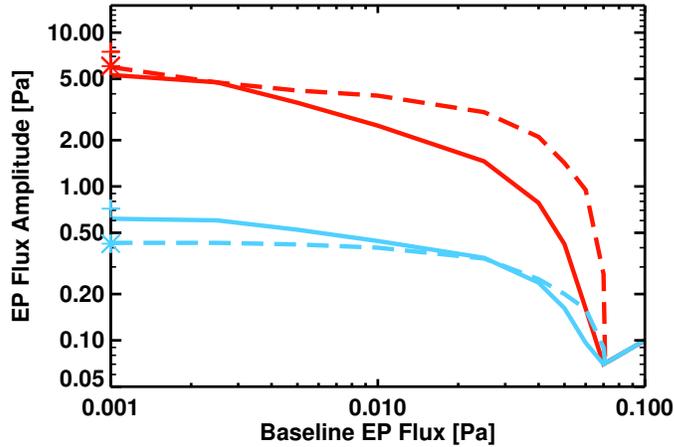


FIGURE 2.10. EP flux amplitude of bottom boundary sine-squared pulses necessary to force SSW-like transitions for the geopotential model (solid) and for the EP flux model (dashed) as a function of baseline EP flux. Results for pulse time scales of 20 days are shown in blue while results for pulse time scales of 2 days are shown in red. Symbols along the left ordinate represent zero baseline results, where ‘+’ is for the geopotential model and ‘*’ is for the EP flux model.

Figure 2.10 shows the bottom boundary wave activity flux amplitude necessary to drive a SSW-like transition as a function of the baseline for both models. Results for only two pulse time scales are shown – 20 day time scale in blue and 2 day time scale in red. For direct comparison, we present results from the geopotential model in equivalent wave activity flux values rather than in geopotential values.

When the baseline is zero (symbols), the geopotential model requires at least 25% higher bottom boundary wave activity flux amplitudes than does the EP flux model. That is, the EP flux model more easily forces SSW-like transitions when the external feedback is severely reduced in the geopotential model. We suggest this is so because even simple forms of prescribed bottom boundary geopotential perturbations in the geopotential model result in complex forms of incoming wave activity flux. The magnitudes of wave activity flux in either model rely on the magnitudes of geopotential, on the phase shifting with height of the

geopotential, and on the zonal wind (which affects both the magnitudes and phases of the geopotential). Thus specified bottom boundary geopotential perturbations poorly specify incoming upward wave activity fluxes.

With nonzero baseline, the geopotential model overcomes the poor specification discussed above by drawing wave activity from below the bottom boundary. Incoming wave activity flux amplitudes necessary for SSW-like transition tend to be lower in the geopotential model than in the EP flux model for nonzero baseline. As the baseline approaches the bifurcation points for each model, the incoming wave activity flux amplitudes necessary to force a SSW-like transition collapse to the baseline value. This is evident for baseline larger than ~ 0.070 Pa in Figure 2.10, where the amplitudes of both models at both pulse time scales are equal to the baseline value.

In the geopotential model, the initial bottom boundary pulse of wave activity flux conditions the model mean state to subsequently flux sufficient bottom boundary wave activity into the model in order to complete the transition. In this way, the external feedback allows a SSW-like transition to be forced with smaller pulse amplitudes than would be needed without the feedback. In line with how the external feedback works, these reductions in necessary amplitudes are however accompanied by increases of at least 25% in the total incoming wave activity flux. For the transient forcing results presented here, the total incoming wave activity flux is always higher in the geopotential model than in the EP flux model.

In both models, the time scale at which the minimum in total incoming wave activity flux is located becomes shorter as the baseline is increased. The external feedback allows pulses to force transitions with smaller amplitudes, which allows the associated total incoming wave activity flux to be smaller as well. Because this reduction is of largest magnitude at the shortest pulse time scales, the minimum in total incoming wave activity flux moves towards

shorter time scales. This is perhaps not surprising, given the strong dependence of necessary amplitudes on pulse time scale (cf. Figure 2.7) and that there is a minimum amplitude to drive a transition (cf. Figure 2.3). Amplitudes for very long pulses can only be reduced to a certain degree before reaching the steady state transition amplitude, while amplitudes for very short pulses may be reduced substantially before reaching this point. Through this, we find that the external feedback acts to shorten the wave forcing time scales as the feedback is strengthened (i.e. as the baseline is increased).

2.4. TROPOPAUSE EFFECTS

In order to study the impact of the tropopause on the forcing of SSW-like events in our versions of the Holton-Mass model, we consider three tropopause setups. The first has no vertically-varying stratification (as measured by the buoyancy frequency squared, N^2); the entire model depth is held at the stratospheric stratification of $4 \cdot 10^{-4} \text{ s}^{-2}$. We refer to this as the “stratosphere setup.” Note that this is equivalent to the control setup used up to this point. The second is a hyperbolic tangent structure in height starting at 11 km and increasing over a depth of 4 km from tropospheric to stratospheric stratification. We refer to this stratification profile as the “tropopause setup.” The third profile includes a tropopause inversion layer (TIL), for which N^2 near the top of the tropopause jumps to a much higher stratification and then quickly relaxes to stratospheric values (cf. Birner 2006). For this profile, referred to as the “TIL setup,” we utilize the tropopause setup, but add a 1 km, squared sinusoid of maximum amplitude $2 \cdot 10^{-4} \text{ s}^{-2}$ to the profile, centered a half kilometer above the center of the hyperbolic tangent. This relates to a maximum TIL stratification of $6 \cdot 10^{-4} \text{ s}^{-2}$, a magnitude similar to that found in climate model output (Birner et al. 2006).

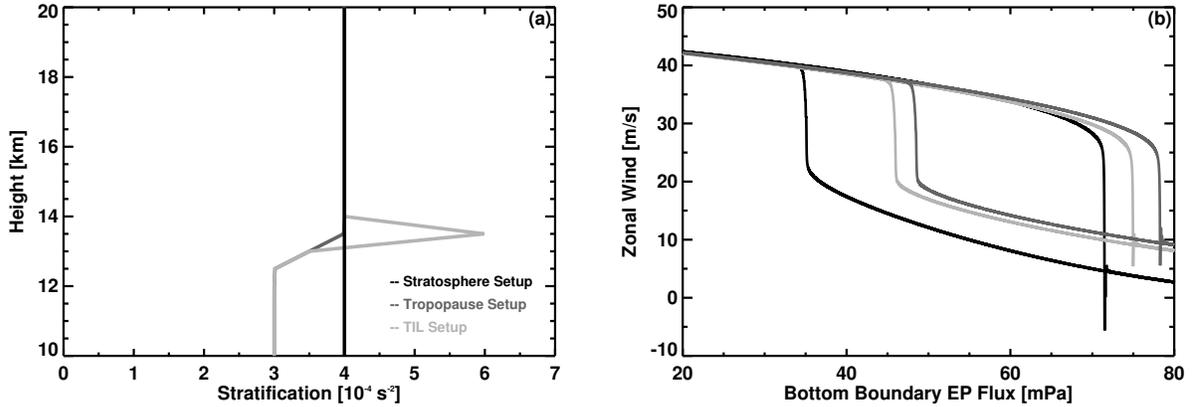


FIGURE 2.11. (a) Vertical stratification profiles for the three setups: stratosphere setup (black), tropopause setup (dark gray), and TIL setup (light gray). (b) As in Figure 2.3 (b), but for each stratification profile and without the stability states analysis.

We firstly analyze the unforced model states under different magnitudes of tropospheric stratification. We find that as the stratification of the model troposphere is lowered (i.e. the jump in stratification at the tropopause is increased), the relevant index of refraction at the tropopause levels becomes increasingly negative, asymptoting to $-\infty$. The tropopause then will reflect increasing amounts of upward wave activity flux as the tropopause stratification jump is increased. In the real atmosphere, this reflection would be at least partially overcome by horizontal wave propagation along the tropopause. Meridional wave propagation is prevented in the Holton-Mass model however, leading instead to highly unphysical results as the tropopause stratification jump is increased. We therefore opt for a tropospheric stratification of $3 \cdot 10^{-4} \text{ s}^{-2}$ (greater than the observed $\sim 1 \cdot 10^{-4} \text{ s}^{-2}$). No such highly unphysical results come from the jump in stratification for our TIL setup and thus we feel its magnitude is appropriate for this type of study. Furthermore, we find the results discussed below hold qualitatively both for different maximum stratification of the TIL setup and for different depths of the TIL.

The above idealized tropopause setups are shown schematically in Figure 2.11 (a). Note that the stratification profiles shown are at the model resolution of 500 m. To determine the steady states for each stratification profile, we perform the same slowly ramping, quasi-steady EP flux experiment as before (see Section 2). Figure 2.11 (b) is similar to Figure 2.3 (b) except that the results of all three of our idealized tropopause setups are shown.

Comparing the tropopause setup (dark gray) and TIL setup (light gray) to the stratosphere setup (black), one observes that including a tropopause of either form shifts both the upper bound and the lower bound of the bistable regime to higher magnitudes of bottom boundary upward EP flux. The higher magnitudes of the upper bounds imply that the tropopause and TIL stratification profiles reduce the stratospheric wave activity flux, requiring that more be injected into the model in order to drive a transition between steady states. Similarly, the higher magnitudes of the lower bounds of the bistable regime for both the tropopause and TIL setups indicates that the necessary forcing to keep the model state within the strongly perturbed regime has increased. These results are in line with Chen and Robinson (1992), who showed that the tropopause reduces upward wave propagation into the stratosphere. Furthermore, with both the tropopause and TIL setups, the zonal wind at 30 km in the strongly perturbed stable state is approximately 5 m s^{-1} larger than in the stratosphere case, indicating that the level of strongest wave forcing has shifted upwards.

If the tropopause acts to decrease the upward EP flux entering the stratosphere and thus the model requires more incoming EP flux for transition, why is the TIL setup upper bound closer to that of the stratosphere case than is the tropopause case? Its higher maximum stratification would seem *prima facie* to result in a higher upper bound than even the tropopause setup.

By analyzing the vertical structure of the wave activity flux in the presence of our TIL, we find EP flux divergence (wave generation) along its upper bound. This divergence comes about from a localized reversal of the meridional gradient of mean potential vorticity, satisfying the Charney-Stern necessary condition for baroclinic instability (Charney and Stern 1962). In the Holton-Mass model, this term takes the form:

$$(2.20) \quad \partial_y \bar{q} = \beta + \epsilon l^2 U - \frac{f_0^2}{N^2} \epsilon \left(\partial_{zz} U - \frac{1}{H} \partial_z U - 2 \frac{\partial_z N}{N} \partial_z U \right),$$

where f_0 is the Coriolis parameter, β is the meridional derivative of the Coriolis parameter, l is the meridional wavenumber of the model channel, ϵ is a factor resulting from the Fourier sine series expansion of $\sin^2 ly$, and H is the scale height. Due to the vertical structure of the TIL stratification, the final term in (2.20) is large and negative enough such that $\partial_y \bar{q} < 0$.

This reversed meridional gradient of mean potential vorticity provides a small, but important source of EP flux within the model domain, as it helps abate the reflection and dissipation of wave activity by the tropopause. We demonstrate the effect of this source by making use of the imposed zonal wind deceleration experiment of Figure 2.5 (a), but with the tropopause and TIL setups. Figure 2.12 shows the zonal wind and upward EP flux evolution for (a) the tropopause setup and (b) the TIL setup. The bottom boundary EP flux magnitude and form of the imposed zonal wind deceleration are identical in the two runs. The amplitude of the imposed deceleration is increased to $39 \text{ m s}^{-1} \text{ day}^{-1}$ to account for the higher threshold needed to overcome for transition. In this experiment, the TIL wave source adds sufficiently to the internal wave-mean flow feedback so that a transition is forced whereas no such transition occurs in the tropopause setup.

It is interesting to note that such mean potential vorticity gradient reversals may be forced in our TIL setup even when the model parameters do not produce a perpetual gradient

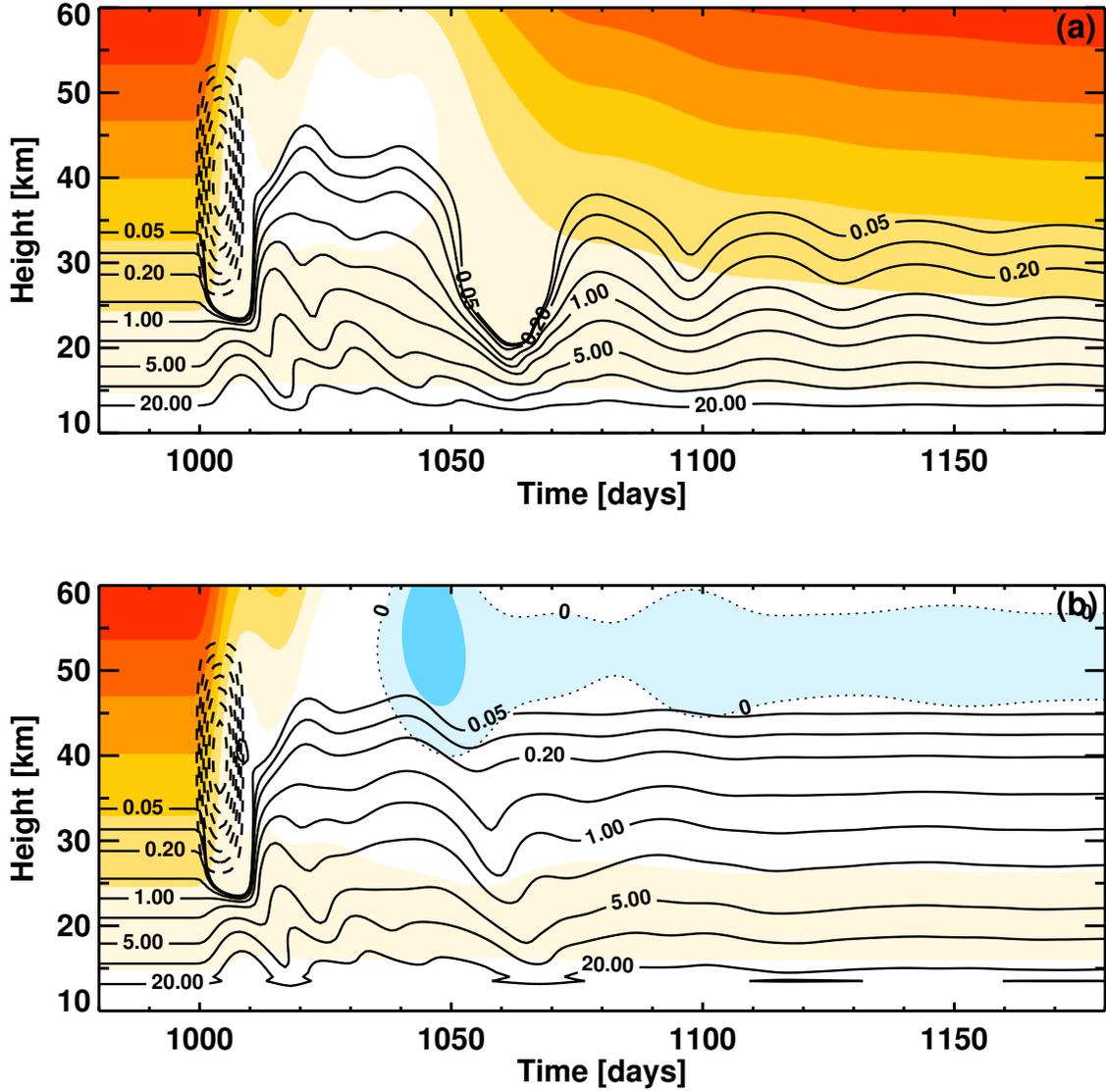


FIGURE 2.12. As in Figure 2.5, but for the EP flux model with (a) our tropopause setup and (b) our TIL setup. The maximum imposed deceleration is now $39 \text{ m s}^{-1} \text{ day}^{-1}$.

reversal. For instance, increasing the radiative equilibrium (initial) zonal wind at all levels increases the positive contribution of the second term in (2.20) while keeping the other terms the same, allowing $\partial_y \bar{q}$ to be greater than zero. By perturbing this model state with a relatively small bottom boundary upward EP flux, we find that a meridional potential vorticity gradient reversal may be forced. The dominant term in such a reversal becomes the

zonal wind curvature term of (2.20), which is zero here when the model is unperturbed but becomes negative and large as tropopause-level wave activity flux convergence decelerates the wind. In this way, only a reduced meridional potential vorticity gradient magnitude along the top of the TIL is necessary to provide an in situ wave source within the lower stratosphere.

2.5. CONCLUSIONS

Wave-mean flow interaction theory shows that upward EP flux (sometimes approximated by its heat flux component) represents the appropriate measure to describe the planetary wave forcing of SSW-like events. This theory also describes a wave-mean flow positive feedback where wave activity flux convergence within the stratosphere decelerates the zonal wind there, allowing more wave activity to be drawn upwards. A quasi-geostrophic, beta plane channel model of such wave-mean flow interaction – the Holton-Mass model – contains this feedback mechanism, yet is forced by bottom boundary geopotential perturbations. Here, we recast the Holton-Mass model so that upward EP flux is specified at its bottom boundary.

In the standard Holton-Mass model with bottom boundary geopotential forcing, the positive feedback exists both internally and externally. The external feedback allows the model to draw wave activity upwards from a potentially limitless source below the model domain, which makes this feedback unrealistically strong and may create unphysical results. We therefore isolate the internal portion of this positive feedback by specifying the incoming wave activity flux. Furthermore, our two model bottom boundary setups allow us to distinguish characteristics of both the external and internal feedback components.

Through analysis of the stable steady states, we find that the external component of the positive feedback, active in the geopotential model, is the strongest driver of the oscillating zonal wind state associated with the Holton-Mass model. It is through this external feedback that the model state may “turn on” and “turn off” incoming waves. Model states which are conducive to wave propagation allow incoming wave activity fluxes to decelerate the wind, while model states which are not conducive to propagation (i.e. which result in wave reflections) result in acceleration of the wind. In the EP flux model, the effect of wave reflection is damped at lower levels, which acts in part to prevent zonal wind oscillations except for at unrealistically large amplitudes of incoming wave activity flux. This is not unexpected as steady, conservative wave forcing should result in a steady zonal wind response.

The external feedback is not, however, the cause of the parameter-dependent bistable regime previously identified in the Holton-Mass model. We find that this bistability is a more fundamental result of this simple model: even for a given incoming upward wave activity flux, the zonal wind may exist at one of two substantially different steady states. This may apply hypothetically to the real world where a particular seasonal progression into winter (i.e. the state of the zonal wind at the beginning of winter) determines how the stratosphere will be affected by the DJF mean upward EP flux. For very weak or very strong mean upward EP flux, the DJF mean zonal wind state may be determined by the waves, but for moderate upward EP flux (within the bistable regime), the zonal wind may determine its own mean state by affecting how and where the waves propagate and subsequently force the zonal wind itself.

We demonstrate this mean flow determination of its own final state by imposing a perturbation on the zonal wind and allowing the wave-mean flow feedback to drive a stable

state transition. In doing so, we find that the EP flux model is able to generate a SSW-like transition by reducing internal EP flux convergences and increasing wave propagation, thereby increasing the amplitudes of wave activity flux at middle to upper model levels. The geopotential model instead drives the same transition with an order of magnitude increase in incoming EP flux. Because the geopotential model can tap into a limitless reservoir of wave activity below the model domain, much smaller levels of wind perturbation are necessary to set off this transition.

Through an idealized time-dependent incoming EP flux forcing experiment, we find, in agreement with previous results, that the stratosphere has a “preferred” wave forcing time scale on the order of 10 days. This preferred time scale manifests as a minimum in the total incoming upward EP flux necessary to drive a transition. At time scales longer than 20 days, the total incoming EP flux increases linearly with increasing time scale. At time scales shorter than 20 days, the necessary amplitudes are excessively large. These features present considerable requirements on the atmosphere to force SSW-like transitions. Neither excessively large amplitude pulses with very short (~ 1 day) duration nor pulses with very long (~ 100 day) durations are not found in the reanalysis dataset, consistent with these results.

We suggest that the preferred time scale is related to the interplay between internal wave propagation (e.g. group velocity) and the height of the reflecting layer. The positive feedback appears to be strongest when the model is forced with pulses near this preferred time scale. Pulses of shorter duration require larger integrated incoming EP flux to drive the same zonal wind variability. Pulses longer than this preferred time scale are subjected to this requirement as well, but also to large integrated radiative damping.

In analyzing the mechanisms responsible for this feedback time scale, we identified a feature which may be a useful diagnostic. When SSW-like transitions are forced with our time-dependent pulses, the durations of the pulses of wave activity flux in the interior of the model domain become longer than that of the transient bottom boundary pulse. In fact, these durations increase with height. When only weak zonal wind decelerations are forced by our incoming pulses, the durations instead decrease with height. In this way, pulse duration changes with height are a way of diagnosing whether the positive feedback is strong or weak.

By performing these time-dependent forcing experiments with the geopotential model, we find that the external feedback significantly affects the results of this transient forcing experiment. As the source of available wave energy at the model bottom boundary increases, the external feedback becomes stronger and the bottom boundary EP flux pulse amplitude necessary to drive a transition becomes smaller. The magnitude of this decrease in necessary pulse amplitude is largest at short pulse time scales, thereby shifting the minimum in total incoming EP flux to these short pulse time scales. The external feedback, in essence, allows the model to draw up the necessary wave activity to force a transition following these short duration pulses (which concurrently results in much higher total incoming EP flux).

Two idealized forms of the tropopause were included in the EP flux model: a simple tropopause and a TIL. We find, as expected, that the tropopause setup and TIL setup both reduce the EP flux entering the model stratosphere, relative to the model integrated with no vertically-varying stratification (the stratosphere setup). It follows that forcing steady state transitions requires larger EP fluxes when a tropopause is present.

However, the TIL setup requires a smaller increase in incoming EP flux because of in situ wave generation caused by a reversal of the meridional gradient of mean potential vorticity. We find that these gradient reversals may exist in the model either from the selected

parameters, or may be forced by zonal wind decelerations caused by wave activity flux convergences. This wave generation provides an important addition of wave activity to the internal feedback, increasing EP flux amplitudes in the stratosphere relative to amplitudes in the tropopause setup. Even if the external feedback component is weak or removed, in situ wave generation from the TIL will help decelerate the winds aloft, thus strengthening the wave-mean flow internal feedback process. However, it is not clear at this time how applicable this result is to the real world. Further study of this potential impact is required.

In summary, the reformulated model presented here shows that bistability of the zonal wind is a robust, internal feature of this type of simple model; that there exists internal wave time scales upon which zonal wind is most easily forced (in an integrated wave activity flux sense); and that the TIL may be an important source of large-scale wave activity within the stratosphere. The simplified model used here allows us to perform the idealized experiments needed to eke out these features of the system.

Given the simplicity of the model utilized here, it is imperative that evidence of the non-linear behaviors discussed above be sought through other means. One such way is through similar experiments within models of higher complexity. However, because the increased complexity does not easily allow for – and may in fact prohibit – the very specified experimentation performed here, we are left to perform suitable idealized experiments and subsequently apply appropriate diagnostics to eke out the dynamical behaviors we expect.

A different, and perhaps more ready-to-use, way of finding evidence of the above results is through analyzing observational data as we do below in Chapter 3.

CHAPTER 3

WAVE-MEAN FLOW FEEDBACKS IN REANALYSES

3.1. WAVE FORCING TIME SCALES OF SSWs

Material from Section 3.1.1 is published in Sjöberg and Birner (2012). Material from Section 3.1.2 and Section 3.3 is published in Sjöberg and Birner (In Press).

3.1.1. LOWER STRATOSPHERIC HEAT FLUX.

An observational analysis meant to eke out the relationship between the time scales of wave forcing and SSWs is presented here. Stratospheric wave forcing is generally represented by the vertical component of the EP flux vector, which may itself be approximated by the eddy meridional heat flux (Andrews et al. 1987, Newman and Nash 2000). Since these measures mainly differ by a static stability term which is nearly constant above the lowermost stratosphere, the present observational analysis is performed using the eddy meridional heat flux as a proxy for wave forcing. These heat flux data are derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Re-Analysis (ERAi) dataset, are averaged about 45-75°N as in previous studies (e.g. Newman and Nash 2000, Newman et al. 2001, Polvani and Waugh 2004), and cover a period of January 1979 to January 2012 at a time resolution of 6 hours. The observational analysis was also performed using ECMWF 40 year Re-Analysis (ERA-40) and was found to only show minor differences.

While 100 hPa is a typical choice for analyzing the relationship between eddy heat flux and SSWs (e.g., Polvani and Waugh 2004, Charlton and Polvani 2007), this level is also located within the bottom of the polar vortex. That is, at 100 hPa the wave forcing cannot be considered to be completely independent of the evolution of a SSW itself. A lower level, typically 300 hPa (e.g. Polvani and Waugh 2004, Charlton and Polvani 2007), is therefore

often analyzed to more properly account for the tropospheric wave forcing entering the stratosphere. However, since the 300 hPa level is sometimes above the tropopause and sometimes below the tropopause within the range of 45-75°N, the eddy heat flux data are no longer a proper proxy for wave activity (upward EP flux) due to differences in the static stability term between upper troposphere and lower stratosphere. This study therefore utilizes the 200 hPa level – located below the polar vortex but above the tropopause.

The World Meteorological Organization (WMO) defines a major SSW event (hereafter simply SSW event, unless stated otherwise) as a wind reversal from westerlies to easterlies at 60°N and at 10 hPa¹. For the considered period of 34 winters (December through February, DJF hereafter),² there are 18 winters with a total of 20 SSW events and 16 winters with no SSW events. This analysis will make use of three data periods based around the occurrence (or non-occurrence) of a SSW event. The first period contains all winters in which no SSW occurred – hereafter termed non-event winters. The second period contains all winters in which at least one SSW occurred – termed event winters. The third period contains the date of each SSW event and the preceding 45 days – termed the pre-event periods. The dates of the events contained in this dataset were identified by using the algorithm discussed in Charlton and Polvani (2007). For reference we note the climatological winter (DJF) mean heat flux at 200 hPa of approximately 18 Kms⁻¹ (with standard deviation of 11 Kms⁻¹). For the following discussion, heat flux values larger than this background value are considered to be “large.”

¹We note that this definition does not currently have a proper citation, and that the field is engaged in discussion over whether to maintain or modify this definition (Butler et al. Submitted).

²One of these winters is the period January-February 1979, which is included because a major warming SSW occurred in late winter. Another winter is the period December 2011-January 2012, which is included because a minor warming occurred in January 2012. This minor warming is relevant later.

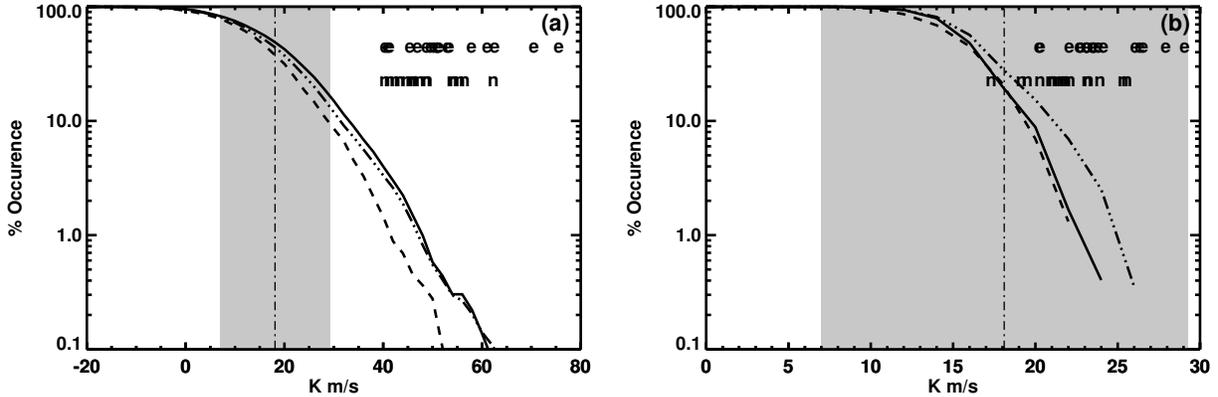


FIGURE 3.1. Cumulative frequency distributions for meridional heat flux (left) and for 40 day averaged heat flux (right). In each of the plots here, the solid curve denotes the pre-event periods, the dash-dot-dot-dot curve denotes the event winters, and the dashed line denotes the non-event winters (see text for definitions). Individual winter maximum heat flux values are plotted for event winters (e's) and for non-event winters (n's) both for instantaneous heat flux (a) and for 40 day averaged heat flux (b). The climatological winter (DJF) mean heat flux and one standard deviation are shown on each plot (dash-dots with gray shading). Note that the y-axes are plotted on a logarithmic scale.

Previous studies of SSWs show strong correlations between high values of meridional heat flux and major warming events (e.g., Baldwin and Dunkerton 1989, Polvani and Waugh 2004). We therefore first compare distributions of heat flux values for our three named periods (non-event winters, event winters, pre-event periods). Cumulative frequency distributions of the (6-hourly) heat flux (at 200 hPa, averaged over 45-75°N) are shown in Figure 3.1 (a). The pre-event periods (solid) and event winters (dash-dot-dot-dot) display little difference throughout the range of heat flux values. In contrast, the frequency of occurrences of large heat flux values during non-event winters (dashed) are systematically less than the frequency of occurrence during both the pre-event periods and event winters. For example, the probability to meet or exceed 20 Kms^{-1} is about 31% in non-event winters compared to 43% in pre-event periods. The probability to meet or exceed 40 Kms^{-1} is about 1.3% in non-event winters compared to 4.0% in pre-event periods.

However, while large heat flux values are less frequent in non-event winters they still occur. For example, absolute counts (not shown) are similar in the range 30-40 Kms^{-1} between the pre-event periods and non-event winters and they drop below ~ 10 in each case for heat flux values greater than 50 Kms^{-1} . Furthermore, the maximum yearly heat flux values do not show a clear separation between event winters (indicated by e's in the figure) and non-event winters (indicated by n's in the figure), but do show that these maxima are typically somewhat larger in event winters.

Polvani and Waugh (2004) argue that weak vortex events are highly correlated with large magnitudes of time-averaged heat fluxes, with highest correlations for time averaging of 30 days or longer. While not all weak vortex events are SSWs, all SSWs are weak vortex events. This suggests that large magnitudes of time-averaged heat flux are crucial to force SSW events. They propose that the 40 day averaged heat flux is perhaps best to capture this correlation, consistent with the earlier Newman et al. (2001) analysis. We may therefore expect a separation between occurrences of large 40 day averaged heat flux in non-event winters and occurrences of large 40 day averaged heat flux in event winters and pre-event periods.

Figure 3.1 (b) shows cumulative frequency distributions for the 40 day averaged heat flux. The 40 day averaged heat flux values are confined to within one standard deviation of the climatological mean instantaneous heat flux (indicated by gray shading). The frequency distributions show a similar difference between non-event and event winters as in (a). For example, the probability to exceed an averaged heat flux value of 20 Kms^{-1} is about 13% in non-event winters, while it is about 21% in event winters. Further, the maximum yearly 40 day averaged heat flux values show a clear tendency toward larger values during event winters compared to non-event winters. However, the pre-event periods distribution follows

much more closely that for non-event winters and appears systematically different from that for event winters. A closer look into this at first surprising behavior reveals that for most SSW events the period of enhanced heat fluxes continues a few days past the event date (typically by 3-4 days), leading to largest averaged heat fluxes when these days past the event date are included. Post-event days are by definition only included in the event winters sample, but not in the pre-event periods. Modifying the pre-event periods to include 5 days past the SSW event puts the distributions for pre-event periods and event winters in much closer agreement. Nevertheless, the additional enhancement of the 40 day averaged heat flux past the SSW event date can hardly be said to precede and therefore cause the event.

The results from the 40 day averaged heat flux values reinforce the concept that SSW events are typically associated with large time-averaged heat fluxes near the tropopause. However, the 40 day averaged heat flux *preceding* SSWs does not appear to be strongly enhanced compared to non-event winters. Furthermore, it is in principle not clear whether large time-averaged heat fluxes arise from an extremely large pulse of short duration within the averaging period, or from moderately large values extending over most of the averaging period. Results in Harnik (2009) suggest that short pulses preferentially lead to wave reflection with insufficient zonal wind deceleration to cause SSWs, while SSWs are associated with longer lasting pulses. We thus more closely investigate the evolution of the heat flux preceding SSWs.

As a first step, a composite heat flux preceding SSWs is formed (similar to Fig. 3d in Harnik (2009)), averaging the heat flux relative to the SSW date and is plotted as the solid curve in Figure 3.2 (a). Also shown in (a) are a composite minor warming heat flux (dashed, see below for minor warming definition) and the climatological winter mean heat flux and corresponding standard deviation (dash-dotted with gray shading). The time-averaged heat

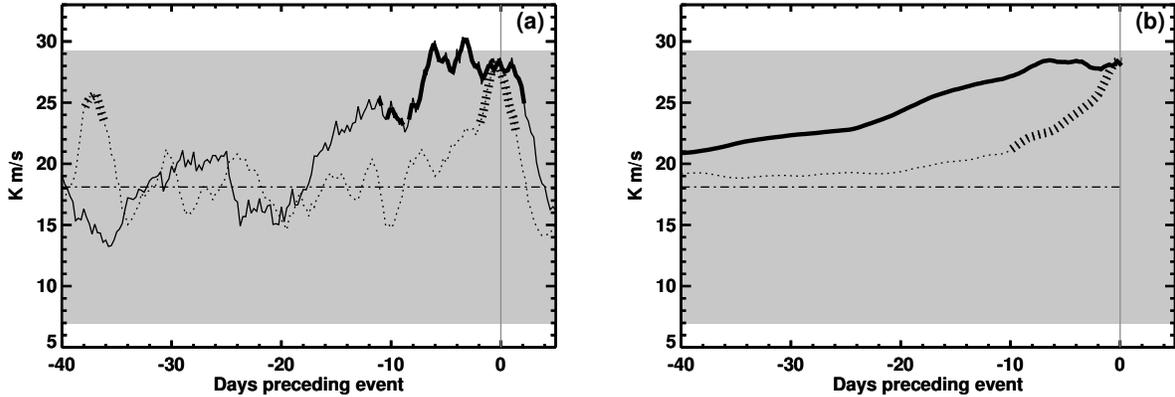


FIGURE 3.2. (a) Composites of heat flux for the pre-event periods (solid), for the identified minor warming events (dashed), and for the climatological winter mean (dash-dots) as a function of day preceding the event. (b) Composites of time-averaged heat flux for the same periods as in (a). The averaging is done over the number of days preceding the warming. Thick curves mark statistically significant differences – at the 95% level – relative to the full winter period. The gray shading denotes the one standard deviation range around the climatological winter mean heat flux.

flux for each of these is plotted in (b) as a function of the number of days preceding the SSW over which it is averaged.

A minor warming is identified here if the zonal mean zonal wind at 10 hPa, 60°N drops below 20 ms^{-1} but stays above 0 ms^{-1} . That is, minor warming events are identified as weak vortex events that do not lead to a major SSW. We analyze these minor warming events so as to compare our (major) SSW heat flux composites with composites from cases of moderately strong wave forcing. The upper bound of this range is chosen because it determines periods of large departures from the climatological winter zonal mean zonal wind speed (of approximately 30 ms^{-1} at 10 hPa, 60°N). We also impose the requirement that once the zonal mean zonal wind becomes larger than 20 ms^{-1} , it must remain larger than 20 ms^{-1} for at least 10 days. Also, no minor warming event may fall within 20 days prior to or following a major SSW. Doing so forces these minor warming events to be well-separated from other minor and major SSW events, helping separate analysis of the wave forcing

affecting major SSW events and minor warming events. This minor warming identification finds 21 cases meeting these criteria in the ERAi dataset.

The composite heat flux profile preceding SSWs only briefly exceeds 30 K ms^{-1} (Figure 3.2 (a)). Heat flux values of 30 Kms^{-1} roughly correspond to the climatological winter mean plus one standard deviation and the probability to exceed 30 Kms^{-1} is approximately 17% for the pre-event periods (cf. Figure 3.1). While these constitute significantly elevated heat flux values, they could certainly not be classified as “extreme” values. Rather it is more notable that the SSW composite profile exceeds the climatological mean heat flux value for almost 20 days prior to the date of the event, with the statistically significant differences extending to about 10 days prior to the event date. In contrast the composite minor warming profile shows a fairly short-lived pulse of heat flux that begins approximately six days prior to the minor warming event (statistical significance only extends to 2 days prior to the event). No such pulse exists in the SSW pre-event periods composite. The minor warming composite also shows a statistically significant similar pulse between 36-38 days preceding the event, suggesting a degree of periodicity in these events (reminiscent of vacillation cycles, e.g. Kuroda (2002)) on a time scale of ~ 1 month.

Similar results may be seen in Figure 3.2 (b), wherein the average heat flux for short averaging periods (which fall near to the event date) is dominated in the case of the minor warmings by the final pulse of heat flux. Statistically significant enhancements of the averaged heat flux are only found for averaging of up to 10 days preceding the minor warming event. The SSW pre-event periods composite, in contrast, shows statistically significant enhancements out to about 50 days (note: the figure only shows averaging periods up to 40 days). The averaged heat flux in this case remains relatively constant out to about 7 days of averaging and steadily decreases towards the climatological mean after that point.

Comparing the SSW pre-event composites with those from the winter mean and from our identified minor warming events gives a clearer picture of how meridional heat flux – and thus the wave forcing – evolves prior to a SSW event. In agreement with the results in Harnik (2009), the large time-averaged heat flux values that are well correlated with SSWs do not appear to arise from brief periods of extremely large heat fluxes but instead from long periods of moderately large heat fluxes. It is important to note that while the latter appears to be more common, the former in principle still remains a possibility (unless prevented through reflection events, see Harnik (2009)).

To test this result further, we compare the average durations over which the heat flux exceeds a given magnitude within each of the three periods considered in this study. From the analysis above, we expect that the average duration for large forcing values is significantly enhanced for the pre-event periods compared to the overall winter climatology. This comparison is accomplished by firstly averaging the durations for which the heat flux exceeds a specified value over the full winter period (including both event and non-event years). To properly compare these with the pre-event periods, we first select the occurrence closest to the SSW date of an exceedance of the specified value. We constrain this final date of the heat flux exceedances to precede the SSW by 10 days or less. While we do not wish to analyze wave forcing well-separated from the SSW event itself, there is typically a brief lead time for the maximum amplitude of wave forcing at 200 hPa relative to the date of the SSW. For the ERAi dataset and our set of identified SSW events, the maximum heat flux leads the SSW date by about 6 days on average (with a standard deviation of about 4 days). This implies that most maximum heat flux values will fall within 10 days prior to the SSW event. Our method is not sensitive between 5 and 15 days for this chosen constraint. An additional constraint on the analyzed exceedances is that they must last longer than the resolution of

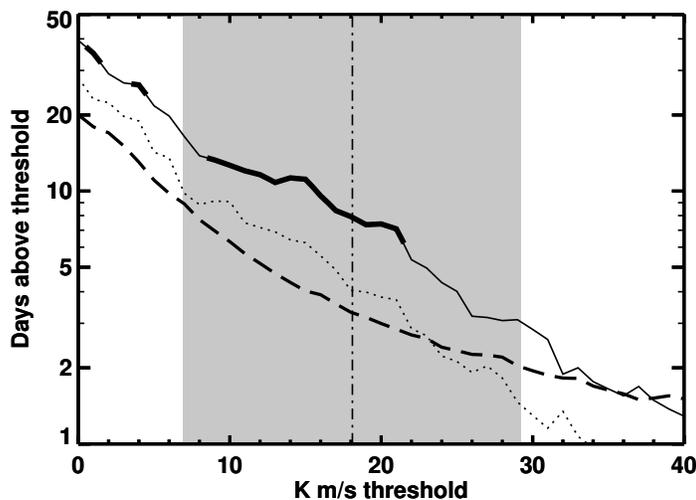


FIGURE 3.3. Average number of days (ordinate) the heat flux stays above a certain threshold (abscissa) for the pre-event periods (solid), for our identified minor warming events (dotted), and for the full winter period (long dashes). Thick curves mark statistically significant differences – at the 95% level – relative to the full winter period. The climatological winter mean heat flux and one standard deviation are shown in dash-dots with gray shading. Note that the y-axis is plotted on a logarithmic scale.

the data (here, 6 hours) so as to disallow inclusion of possibly erroneous data points. The primary results of this comparison hold even if these point exceedances are included.

Figure 3.3 compares the average duration that the heat flux exceeds a given value for the pre-event periods (solid) and for all winter data (dashed)³. Statistical significance at the 95% confidence level of the difference between these curves is assessed based on a Student’s t-test – thick lines mark regions of statistical significance. The degrees of freedom for the pre-event periods is simply the number of SSWs (20) found in the ERAi data set. The degrees of freedom per winter for the full winter period are taken to be the number of duration events if that number is below 40, and 40 otherwise (this number corresponds to the ratio of the approximate length of a winter – 100 days – to the decorrelation time scale

³The average duration for non-event winters is only slightly reduced compared to all winters, by less than 1 day over the range of heat flux thresholds considered.

of the heat flux – about 2.5 days). The average duration is significantly larger (by a factor of 2-3) for the pre-event periods than for the winter climatology for heat flux thresholds between $\sim 10\text{-}20 \text{ Kms}^{-1}$ (recall the climatological winter value of 18 Kms^{-1}). At larger heat flux thresholds, in particular outside one standard deviation of the climatological mean, the average durations are similar between the two periods with statistically insignificant differences. Isolated regions of statistically significant differences between the pre-event and the climatological curve are also obtained for strongly reduced heat flux thresholds (near $\sim 1 \text{ Kms}^{-1}$ as well as $\sim 4 \text{ Kms}^{-1}$), indicating that extended periods of strongly reduced heat fluxes are less likely to occur in the pre-event period.

For heat flux exceedances around the climatological mean value the average duration is on the order of 1 week in pre-event periods. This suggests that on average in order to force a SSW, the wave forcing entering the stratosphere needs to exceed climatological mean values for a period of at least 1 week. The strength of this wave forcing appears to be secondary – moderately large heat flux values, around or just above the climatological mean, appear to be sufficient. The duration for larger heat flux exceedances decays exponentially and is not much different between the overall climatology and the pre-event periods, indicating that these more extreme heat flux pulses are short-lived and less likely to force a SSW (according to Harnik (2009) short-lived wave forcings more likely lead to wave reflection events). This analysis shows that the enhanced SSW composite heat flux values (above about 25 Kms^{-1}) during the 10 days preceding the event in Figure 3.2 (a) are a result of short lived (2-3 days according to Figure 3.3) wave pulses with slightly different timing relative to the individual SSWs. That is, forcing of the individual SSWs does not appear to be so much due to enhanced wave forcing, but more due to the wave forcing not dropping below the

climatological value for an extended period. Note that our analysis of durations for the pre-event periods by definition excludes the frequent heat flux exceedances a few days past the SSW date (according to Figure 3.2 (a), the heat flux typically stays above the climatological value for 3-4 days past the event). Including these days past the event date would make the average duration for the pre-event periods even larger. However, heat flux exceedances past the SSW date can hardly be said to precede and therefore cause the event.

The duration analysis was also performed for the identified minor warming events (dotted line in Figure 3.3). Even though these minor warmings also show enhanced durations compared to the climatology for heat flux exceedances in the range 10-20 Kms^{-1} , these are not statistically significant. Furthermore, the average durations for the minor warmings are about a factor of two smaller than for the (major) SSWs⁴. For large heat flux thresholds the minor warming curve drops below the climatology, largely because many minor warmings do not exhibit any exceedance (i.e. zero duration) at these large thresholds.

We have also tested the predictive capability of periods of long lasting moderately enhanced wave forcing by identifying corresponding events in the 200 hPa heat flux time series and analyzing the zonal wind response at 10 hPa. We hypothesize that long duration wave forcing events (~ 10 days or longer) of moderate strength lead to SSWs. For this analysis we find all points in the data which exceed a given heat flux threshold for a given number of days. At each of these points, the minimum zonal-mean zonal wind (at 10 hPa, 60°N) within 10 days of the end of this exceedance is determined. A similar procedure is undertaken to analyze time-averaged heat flux: we determine the minimum zonal-mean zonal wind within 10 days of a point that shows a time-averaged heat flux value above a given threshold for a

⁴Due to the relatively small number of degrees of freedom for both the minor and the major warming ensembles, the difference between the minor and major warming curves is only statistically significant at the 90% confidence level for a range of heat flux thresholds around the climatological mean value.

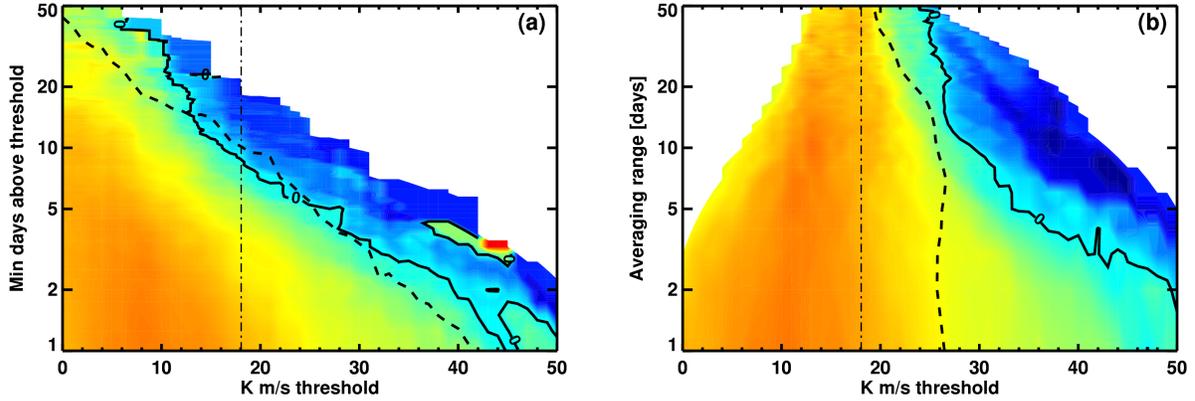


FIGURE 3.4. Mean minimum zonal wind within 10 days of (a) exceedance of a given heat flux threshold lasting a given number of days and (b) exceedance of a given time-averaged heat flux value for a given averaging range, shown in color contouring. Green, yellow, and red colors denote positive values, blue denotes negative values. Overplotted in (a) are the 0 ms^{-1} contour (solid) and the mean durations during pre-event periods (dashed, same as full line in Figure 3.3) both shown as functions of heat flux threshold. Overplotted in (b) are the 0 ms^{-1} contour (solid) and the composite time-averaged heat flux during the pre-event periods (dashed, same as full line in Figure 3.2 b) both shown as functions of the averaging range. The climatological winter mean heat flux is plotted in dash-dots. The y-axes are plotted on a logarithmic scale.

given number of days in the averaging. If this minimum zonal-mean zonal wind is less than zero the heat flux event is identified to have caused a (major) SSW.

The results of this analysis are plotted in Figure 3.4 where (a) is for the duration procedure and (b) is for the time-averaged heat flux procedure. Overplotted on (a) is the pre-event periods mean duration from Figure 3.3, while overplotted on (b) is the pre-event periods composite time-averaged heat flux from Figure 3.2 (b). One observes in (a) that the typical minimum zonal-mean zonal wind is greater than 0 ms^{-1} at any given threshold for durations shorter than the pre-event periods mean duration at that threshold. Similarly, (b) shows that westerlies are predominant for time-averaged heat flux thresholds less than the pre-event periods composite at any given averaging range. Conversely as the durations or

average heat flux thresholds increase, the determined zonal-mean zonal wind minimum falls below zero and indicates a major SSW event.

We find, not surprisingly, that in general combinations of small heat flux thresholds (below the climatological value) with shorter minimum durations (below ~ 1 week) lead to a large number of such heat flux events, most of which are not followed by SSWs. Likewise, combinations of large heat flux thresholds (above the climatological value) with longer minimum durations (above ~ 1 week) lead to only a few heat flux events, almost all of which are followed by SSWs. However, only about half of the 20 identified SSWs in the ERAi data set could be “predicted” this way. A similar but slightly reduced predictive capability was achieved using time-averaged heat fluxes for time averaging of about 10 days.

From our observational analysis we conclude that SSWs are typically forced by planetary wave forcing (meridional heat flux) of moderate strength that lasts 1 week or longer. The duration of the wave forcing appears to be a stronger controlling factor in producing SSWs than the strength of the wave forcing, confirming and extending the results in Harnik (2009). The high correlation between time-averaged wave forcing and the occurrence of SSWs as suggested by Newman et al. (2001), Polvani and Waugh (2004) appears to be a result of the minimum required forcing duration.

3.1.2. UPPER TROPOSPHERE TO UPPER STRATOSPHERE WAVE ACTIVITY FLUXES.

We extend the above pulse analysis to different pressure levels by considering upward Eliassen-Palm fluxes, given by

$$(3.1) \quad F^p = \cos \phi \left[-\overline{u'\omega'} - \frac{\overline{v'\theta'}}{\partial_p \theta} \left(\frac{\partial_\phi (u \cos \phi)}{a \cos \phi} - f \right) \right]$$

In the above analysis, heat fluxes were a good proxy for F^p both because it is the dominant term and because $\partial_p\theta \approx \text{constant}$ between 45°N and 75°N at 200 hPa. While the former remains true at other pressure levels, the latter does not – levels below 200 hPa begin to include the tropopause and/or the troposphere, where the assumption of $\partial_p\theta \approx \text{constant}$ is no longer valid. This makes heat flux a poor proxy for wave activity fluxes, and thus the full EP flux field is utilized here.

This analysis focuses on December through February (DJF) pulses of upward EP flux in the ERAi dataset. Similar to above, we define a pulse here as a continuous anomaly from the seasonally evolving mean upward EP flux averaged over $45\text{-}75^\circ\text{N}$. Figure 3.5 (a) displays the durations and associated maximum amplitudes (scaled by the standard deviation, plotted as asterisks) of all the 200 hPa pulses during the DJF period. Overplotted here as the solid curve is the mean of the scaled maximum amplitudes as a function of duration. There is an apparent relation between the longest pulses and the largest amplitudes, evident in the linear correlation coefficient of ~ 0.67 . This level of correlation holds across the planetary-scale wavenumbers and through the lower stratosphere.

Further, duration of a pulse appears to be a better indicator of an anomalous wave event than does maximum amplitude, confirming results in Chapter 2. Figure 3.5 (a) shows that pulses with maximum amplitude larger than two standard deviations above the mean (marked by the horizontal line) span the range of time scales. Contrarily, pulses with duration longer than two standard deviations above the mean (marked by the vertical line) predominantly have large maximum amplitude. It is evident that the largest amplitudes of upward EP flux in the real world tend to coincide with the longest lived periods of upward EP flux. The large amplitudes necessary to drive a transition for very short duration pulses are not observed.

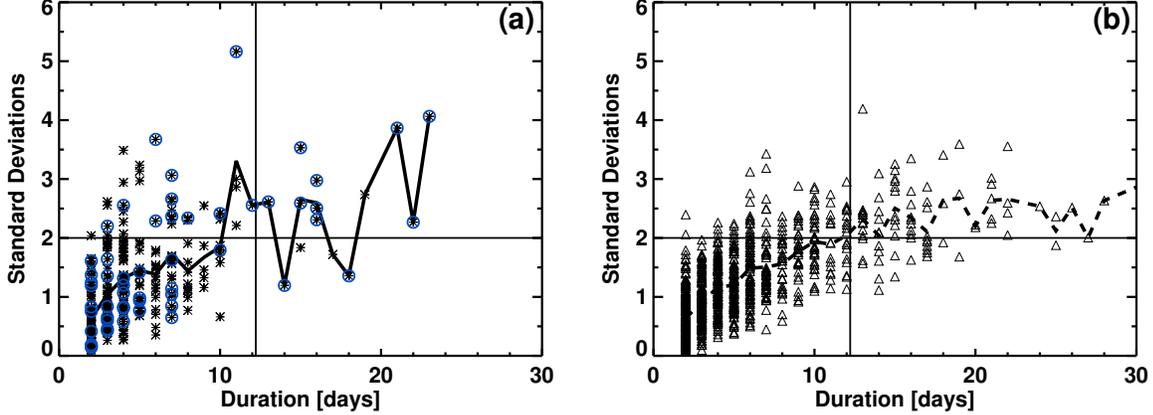


FIGURE 3.5. (a) Duration and maximum amplitude (scaled by standard deviations) of 200 hPa EP flux pulses greater than the DJF mean plotted as asterisks. The thin solid horizontal (vertical) line corresponds to where the maximum amplitudes (pulse durations) are two standard deviations above the mean. The thick solid curve plots the mean of the scaled maximum amplitudes as a function of pulse duration. Asterisks encircled with blue are pulses within ± 20 days of a major SSW. (b) As in (a), but for pulses from an autoregressive model of order 1. Triangles now denote the individual pulse durations and maximum amplitudes, and the dashed curve plots the mean of the scaled maximum amplitudes as a function of pulse duration.

This is validated by considering those asterisks which are contained by blue circles in Figure 3.5 (a), indicating which pulses are within ± 20 days of any SSW (identified using the technique of Charlton and Polvani 2007). Those pulses with duration shorter than 2 standard deviations and with maximum amplitude larger than 2 standard deviations tend to not occur near to a SSW event. Conversely, those pulses with duration longer than 2 standard deviations and with maximum amplitude larger than 2 standard deviations do tend to occur near to a SSW. The two fields for pulses near to a SSW are correlated at ~ 0.90 .

We test whether the presented data are indeed manifestations of a nonlinear, positive feedback by modeling the upward EP fluxes with an autoregressive model of order 1. This

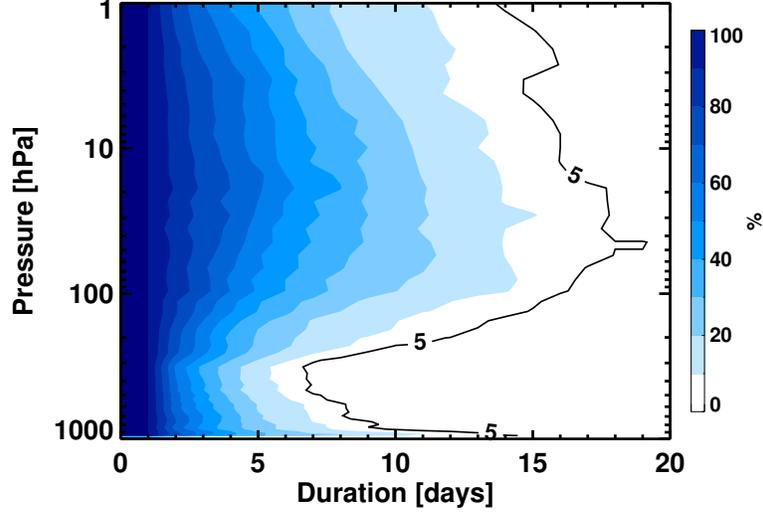


FIGURE 3.6. Percentage of wavenumber 2 upward EP flux events larger than the DJF mean which persist at least as long as that given along the abscissa, plotted as a function of pressure. The solid black contour is for 5%.

time series is generated following

$$(3.2) \quad F^p(t+1) = \rho F^p(t) + \epsilon + (1 - \rho) \langle F^p \rangle$$

where ρ is the autocorrelation of ERAi upward EP flux at a given pressure level, ϵ is a scaled random value, and $\langle F^p \rangle$ is the time mean of the upward EP flux at a given pressure level.

This modeled time series is then subjected to the same pulse identification as used on the reanalysis data. The triangles plotted in Figure 3.5 (b) give the durations and associated maximum amplitudes (scaled by the standard deviation) of the pulses from the modeled time series for 200 hPa upward EP flux. Overplotted here as the dashed curve is the mean of the scaled maximum amplitudes as a function of duration. The durations and maximum amplitudes for the modeled data are correlated at ~ 0.75 , higher than the overall set of pulses from the reanalysis dataset yet lower than the correlation for SSW pulses. This autoregressive modeling suggests that the relationship between pulse duration and amplitude in ERAi data is consistent with a red noise spectrum.

Figure 3.6 shows, in shading and as a function of pressure, the percent of upward EP flux pulses in ERAi which have duration at least as long as that given on the abscissa. Very few pulses in the real atmosphere last longer than 20 days. Only above 150 hPa, where radiative time scales are large (Hitchcock et al. 2010), does the 5% contour in the stratosphere exceed the 14 day duration mark. For pulses in the lower stratosphere – between 150-300 hPa – observing a pulse of longer than 14 days is uncommon. Taken in sum, the reanalysis results show that neither anomalously large amplitude yet short (~ 1 day) duration pulses nor very long (~ 100 day) pulses are readily observed. The observational analysis is thus consistent with our simple modeling suggestion of an order 10 day preferential wave forcing time scale associated with SSWs.

3.2. PULSE DURATION DIAGNOSTIC FOR STRONG WAVE-MEAN FLOW POSITIVE FEEDBACKS

The above observational analysis confirms a key result from the simple modeling: that the real world wave-mean flow positive feedbacks appear to have a preferential time scale on the order of 10 days. An additional finding from the simple modeling was that, when the wave-mean flow positive feedback is strongly active, the model pulse durations would increase with height. We search for such instances by further considering the pulses at all pressure levels which we identified above.

To do so, we consider a pulse at some given starting level – say, 700 hPa. If there is a pulse at the next higher level at an overlapping time as the anomalous pulse occurs at this starting level, then we take this to be a vertically-coherent pulse. This process is continued upward in height until the next higher level does not have a pulse existing at an overlapping time as the pulse on the considered level. We can then analyze how the duration

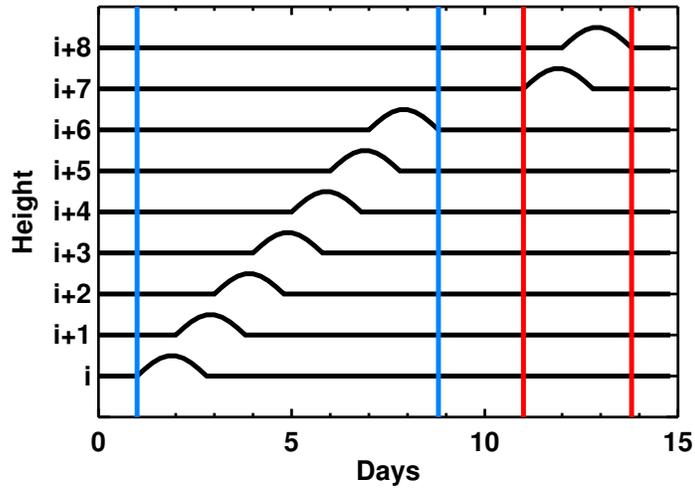


FIGURE 3.7. Schematic showing how the vertically-coherent pulse algorithm works. A pulse, first identified at level ‘i,’ is found to be vertically-coherent to level ‘i+6’ because there are temporally-coincident pulses connected at the next higher level. The pulse at level ‘i+7’ is considered a separate pulse because it is not temporally coherent. The two pulses, starting at level ‘i’ and at level ‘i+7,’ are delimited by the blue and red vertical lines, respectively.

of these vertically-coherent pulses varies with height. Note that, for a given starting level, this algorithm does not require that there is no pulse below the starting level. The starting levels considered here and for the remainder of this analysis are 700 hPa, 500 hPa, 300 hPa, and 200 hPa.

Figure 3.7 gives a schematic of how this algorithm works. Here, there are two independent pulses delimited by vertical lines of different colors. The blue lines delimit a vertically-coherent pulse beginning at level ‘i’ while the red lines delimit a vertically-coherent pulse beginning at level ‘i+7.’ In this schematic, the pulse durations are equal at all heights.

If a pulse is vertically-coherent as above, then we deem it to be vertically propagating. This may not be explicitly true in all cases – for instance, when wave generation along the path of a vertically-coherent pulse occurs. But as shown below, the characteristics of

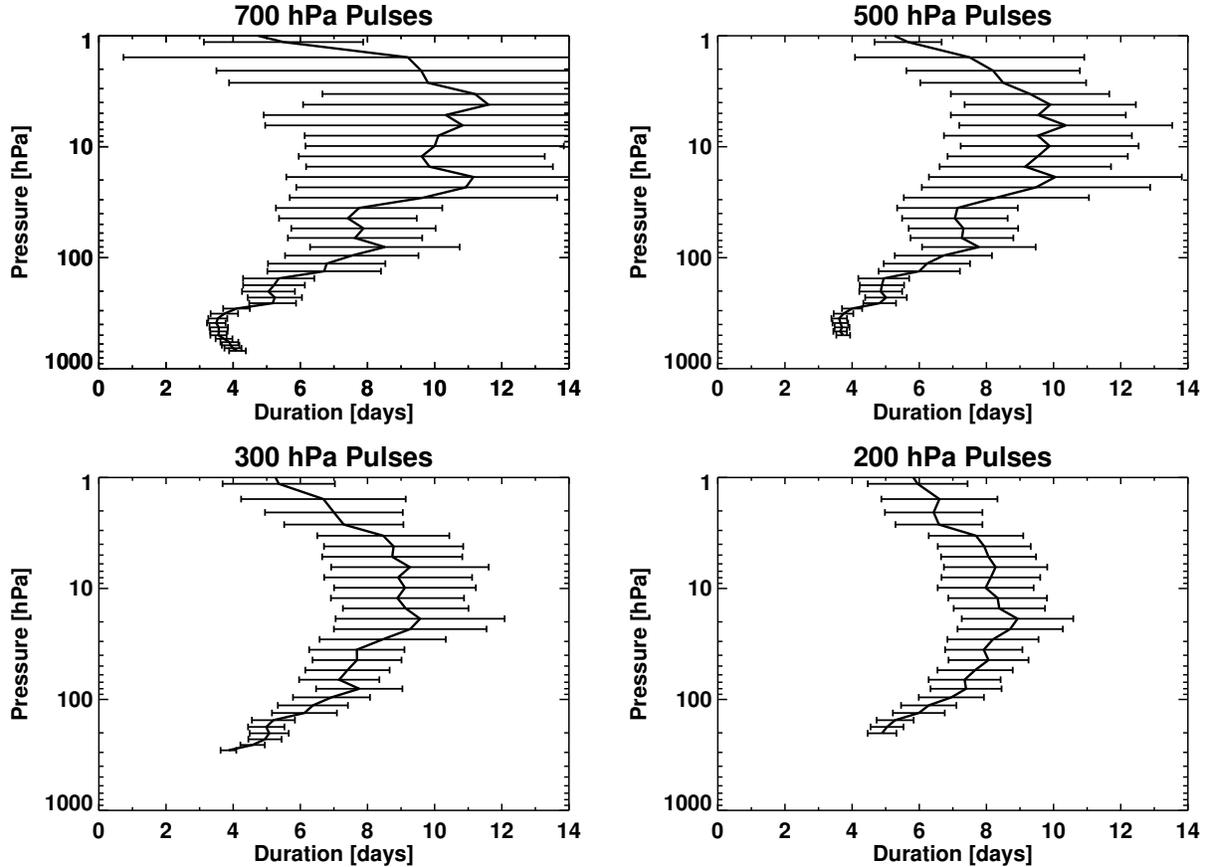


FIGURE 3.8. Mean durations, as a function of pressure, of the vertically-coherent pulses we identify at the four labeled starting levels. The bars at each data point indicate the 95% bounds following a Student's t-test.

these vertically-coherent waves is quite consistent with known characteristics of vertically-propagating waves. We are therefore confident that this algorithm is an effective proxy for identifying the vertically propagating pulses in which we are interested.

We first show relevant characteristics of these vertically-coherent waves, namely the typical (mean) durations and typical maximum amplitudes. Figure 3.8 shows these mean durations as a function of pressure for vertically-coherent pulses at each of the four starting levels. Overplotted on each panel is the 95% bounds following a Student's t-test. In line with Figure 3.6 above, this figure demonstrates that pulses tend to increase in duration when they propagate into the stratosphere. This is seen by the significant difference in mean durations

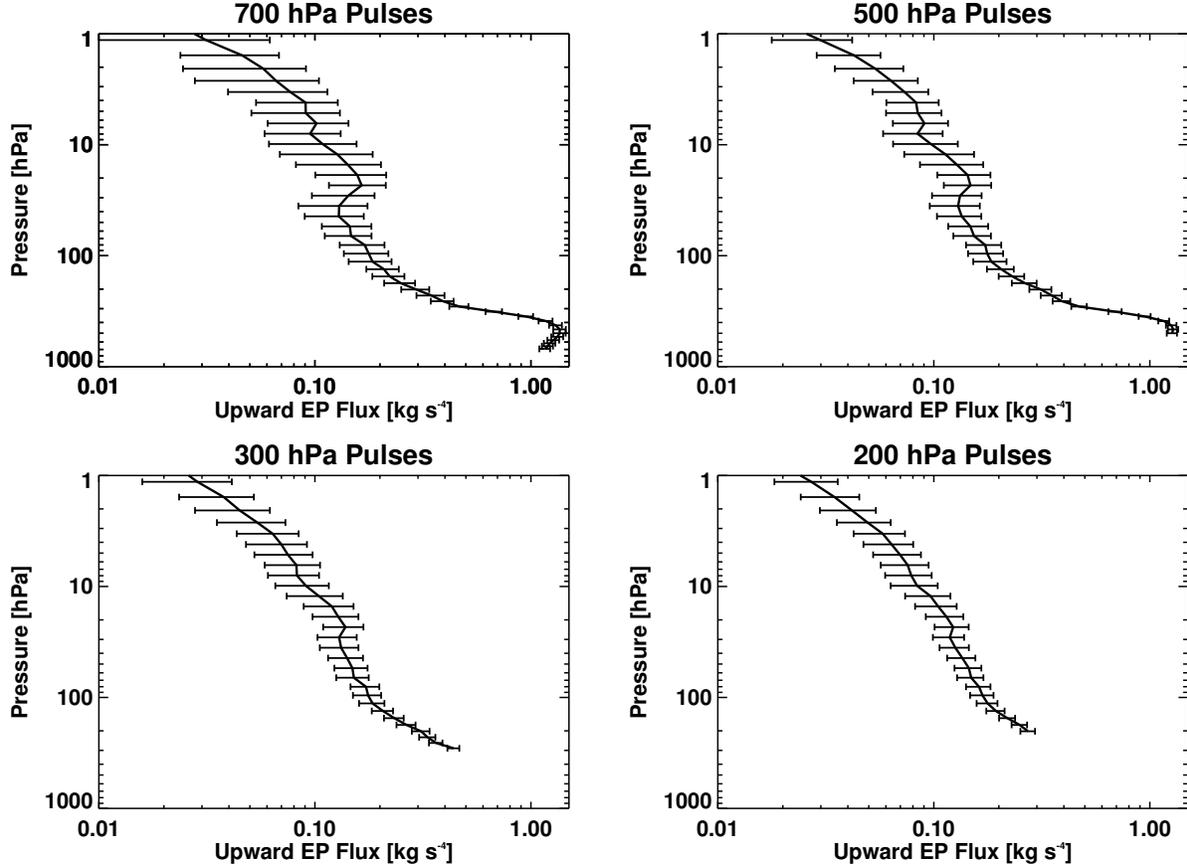


FIGURE 3.9. Mean of the maximum upward EP flux amplitudes, as a function of pressure, of the vertically-coherent pulses we identify at the four leveled starting levels. The bars at each data point indicate the 95% bounds following a Student's t-test.

between pulses above 100 hPa and pulses at each of the starting levels. We note that where these vertical profiles of pulse durations across different starting levels overlap, they do not significantly differ. Once these vertically-coherent pulses are above 100 hPa, the mean durations do not significantly differ with height, suggesting that the main influence on pulse duration occurs within the lowermost stratosphere (i.e. the region near the tropopause).

Figure 3.9 shows the mean of the upward EP flux amplitudes as a function of pressure for the vertically-coherent pulses at each of the four starting levels. As we expect, the amplitudes of the upward EP fluxes decrease with height, with the strongest of these decreases just below the tropopause. Once these vertically-coherent pulses are within the stratosphere, the

amplitudes still decrease but at a rate much less than in the upper troposphere. As with the pulse durations, the differences in maxima between pulses with different starting levels are small, and thus not significant.

As suggested by Figure 3.6 and Figure 3.8, the upper troposphere to the lower stratosphere is the region in which the pulses become significantly longer in duration than middle-tropospheric wave pulses. Though the full reasons for this are not well understood, it must in part be related to filtering of wavenumbers 3 and higher by strong zonal winds (Charney and Drazin 1961). Indeed we find that that the amplitudes of these wavenumbers 3 and higher vertically-coherent waves, while comparable in the middle troposphere, become an order of magnitude smaller than wavenumbers 1 and 2 near and above the tropopause⁵. Given that these high wavenumber pulses have much shorter duration than low wavenumber pulses, filtering of these components from the total EP flux field results in the durations of the total EP flux field pulses becoming longer. In addition, decreased wave-wave interactions here will decrease the dissipation of the planetary-scale waves (waves 1 and 2), keeping pulses larger in amplitude and thus longer in duration.

This is not a complete explanation, however, as vertically-coherent wavenumbers 1 and 2 pulses do not simply retain their long durations. As shown in Figure 3.10, wavenumber 1 pulses increase in duration between 300 hPa and 100 hPa while wavenumber 2 pulses increase in duration between \sim 400 hPa and 60 hPa. As the amplitudes of both wavenumbers either increase or decrease much less than at other pressure levels within their respective regions, we postulate that this is likely due to wave-mean flow interactions. Positive feedbacks would allow for this kind of wave growth and duration lengthening. The exact process through

⁵See Figure A.1 and Figure A.2 in the appendix for plots related to this wavenumber decomposition analysis.

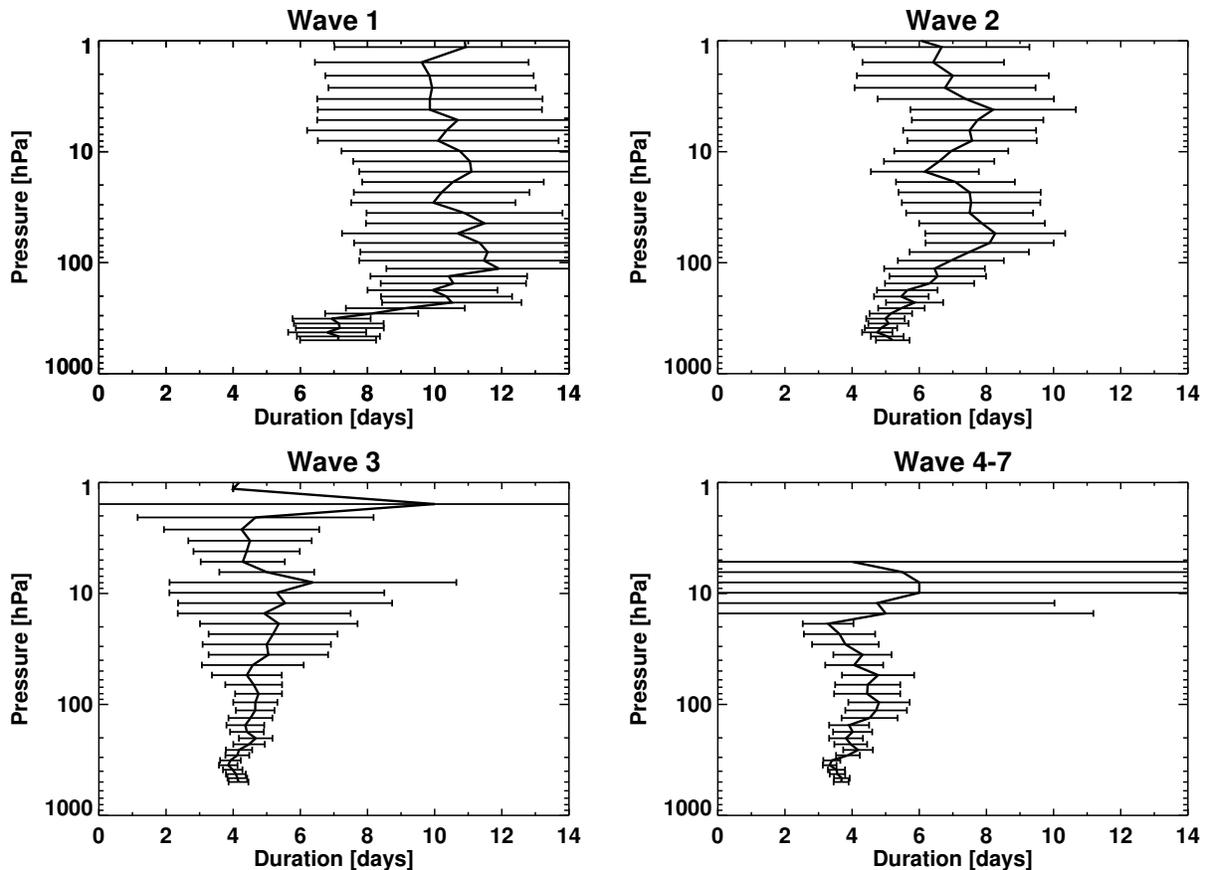


FIGURE 3.10. Mean durations of upward EP flux amplitudes, as a function of pressure, of the vertically-coherent pulses identified at 500 hPa. The panels from top to bottom and left to right show the durations for upward EP flux of wavenumbers 1, 2, 3, and 4-7, respectively. The bars at each data point indicate the 95% bounds following a Student's t-test.

which this occurs is not clear, however, leaving this aspect of the vertical changes in pulse durations for later study.

Longwave radiative damping may also factor into these changes of duration with height in the lower stratosphere. The time scales of this damping during winter are relatively longer in the lowermost stratosphere than in the troposphere (e.g. Hitchcock et al. 2010). Then amplitudes of waves which reach the lower stratosphere will be less damped than they were in the troposphere, perhaps leading to longer duration wave pulses. However, Figure 3.8 shows that the durations of pulses increase slightly or at least remain constant above 100

TABLE 3.1. Total number of vertically-coherent pulses for the given starting level and the percentage of these pulses which are coherent to at least 100 hPa.

Starting Level	# of Pulses at Starting Level	% of Pulses Reaching 100 hPa
700	335	9.6
500	367	13
300	347	20
200	284	40

hPa, in spite of radiative time scales decreasing at these levels. Wave-mean flow interaction must explain this maintenance or increase in the pulse durations, suggesting that radiative damping plays only a weak role here.

This is supported by our modeling of the positive feedback. In the case of the simple modeling of Chapter 2, the largest increases in duration occur below 25 km (cf. Figure 2.9), where the radiative damping is mostly invariant. The same holds for general circulation model results which will be presented in Chapter 4: the radiative damping time scales do not vary with height over the region where the durations increase the most. Thus while weak lower stratospheric radiative damping may act to amplify the increases in the durations, our model results suggest that the dynamics – primarily those associated with the positive feedback – are responsible for most of the increases in pulse duration within the upper troposphere to lower stratosphere.

We further study these vertically-coherent pulses by analyzing the depths to which the pulses are vertically-coherent. Figure 3.11 shows the cumulative frequency distribution of the pulses which are coherent to at least as high as the associated pressure level for the chosen four starting levels: 700 (dark blue), 500 (light blue), 300 (orange), and 200 hPa (red). For example, roughly 10% of pulses with starting level 700 hPa (dark blue) reach 100 hPa or above. Listed in Table 3.1 are the total number of pulses for each starting level (second column) and those pulses which reach 100 hPa or above for each starting level (third column).

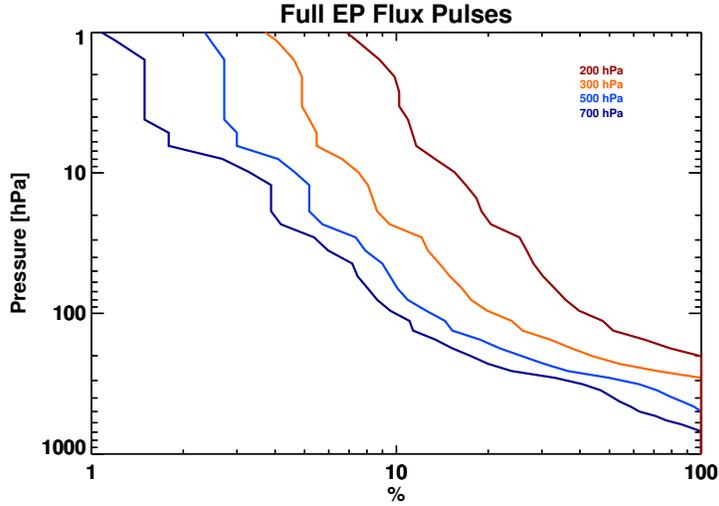


FIGURE 3.11. Cumulative frequency distributions of the pulses which reach at least as high as the associated pressure level for the chosen four starting levels: 700 (dark blue), 500 (light blue), 300 (orange), and 200 hPa (red).

The above results show that a nontrivial number of pulses which exist in the upper troposphere or lower stratosphere reach 100 hPa or higher. In this light, we classify pulses which reach at least 100 hPa (from our chosen pulse starting levels) as ‘deep’ pulses and all those which do not as ‘shallow.’ 100 hPa is a natural choice, as it is the level at which the durations and maxima of these vertically-coherent pulses becomes roughly constant. This shows that pulses below 100 hPa are strongly impacted by the positive feedback, while pulses above 100 hPa, while also feeling the positive feedback, are strongly impacted by the vortex. Thus requiring deep pulses to reach above 100 hPa may select a set of pulses which are not appropriate for studying the positive feedback. Furthermore, results from Section 2.3.1 show that a level below where the zonal wind reverses is appropriate for the duration diagnostic proposed there. 100 hPa often meets this criterion, except in a few instances of major SSWs. The following results largely hold if the level for identifying deep pulses is chosen to be either 50 hPa or 10 hPa.

By considering only these deep pulses, we have identified a subset of vertically-coherent pulses with which we may study a finding from the simple modeling results on wave-mean flow positive feedbacks. We found in the simple model (see Section 2.3.1) that pulse durations increase with height when the positive feedback is strong, but decrease with height when the positive feedback is weak. This was identified by forcing the model with transient pulses of bottom boundary upward EP flux of different maximum amplitude. When the pulse produces a SSW-like transition, the pulse duration increases with height; when the pulse only weakly decelerates the zonal wind, the pulse duration decreases with height.

The validity of these results may be tested by considering differences in wave characteristics and zonal wind evolution between wave pulses which have decreasing durations in height and wave pulses which have increasing durations in height. Our hypothesis is that the latter pulses will have much larger maximum amplitudes and thus drive much larger zonal wind decelerations.

To perform this analysis, we consider the deep wave pulses described above. At each starting level, we partition into subsets those deep pulses which have increasing durations between the starting level and 100 hPa and those deep pulses which have decreasing durations between the starting level and 100 hPa. We hereafter refer to the former as ‘increasing pulses’ and the latter as ‘decreasing pulses.’ Note that increasing and decreasing are relative to the mean duration changes in height – thus a pulse which initially decreases duration from the starting level but subsequently increases in duration to be longer than at the starting level will be labeled ‘increasing,’ etc.

Figure 3.12 shows the number of deep pulses identified as a function of starting levels and the percent of these deep pulses which are classified as increasing or decreasing. The total number of deep pulses at each starting level increases roughly as an inverse power

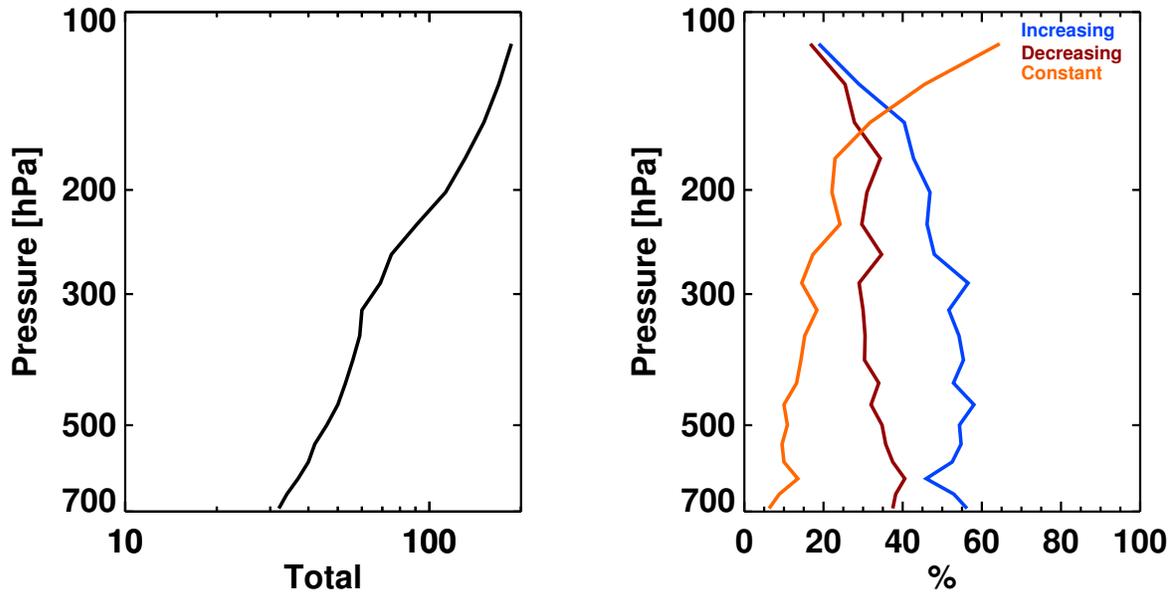


FIGURE 3.12. Total number of deep pulses as a function of starting pressure (left) and the percent of these deep pulses (right) which have increasing durations with height (blue), decreasing durations with height (red), and constant duration with height (orange).

law for higher starting level. Though the tropopause forms a strong boundary to vertical wave propagation (Chen and Robinson 1992), this effect only manifests weakly here as a slight decrease in the slope of the relationship. This is in part because we consider pulses as exceedances relative to each level’s mean; thus while the amplitudes decrease strongly across the tropopause, the pulses can still be vertically-coherent. Note that some of the pulse events overlap across different starting levels, and therefore the pulses at different starting levels are not independent sets.

The percent of these deep pulses which are increasing pulses remains roughly constant near 50%. The decreasing pulse percentage as well remains roughly constant at $\sim 35\%$ of all deep pulses. There is a sudden decrease in these percentages at levels near to 100 hPa, primarily because the distance between the starting level and 100 hPa is small at these levels. With so small a distance separating the levels, the probability of constant durations

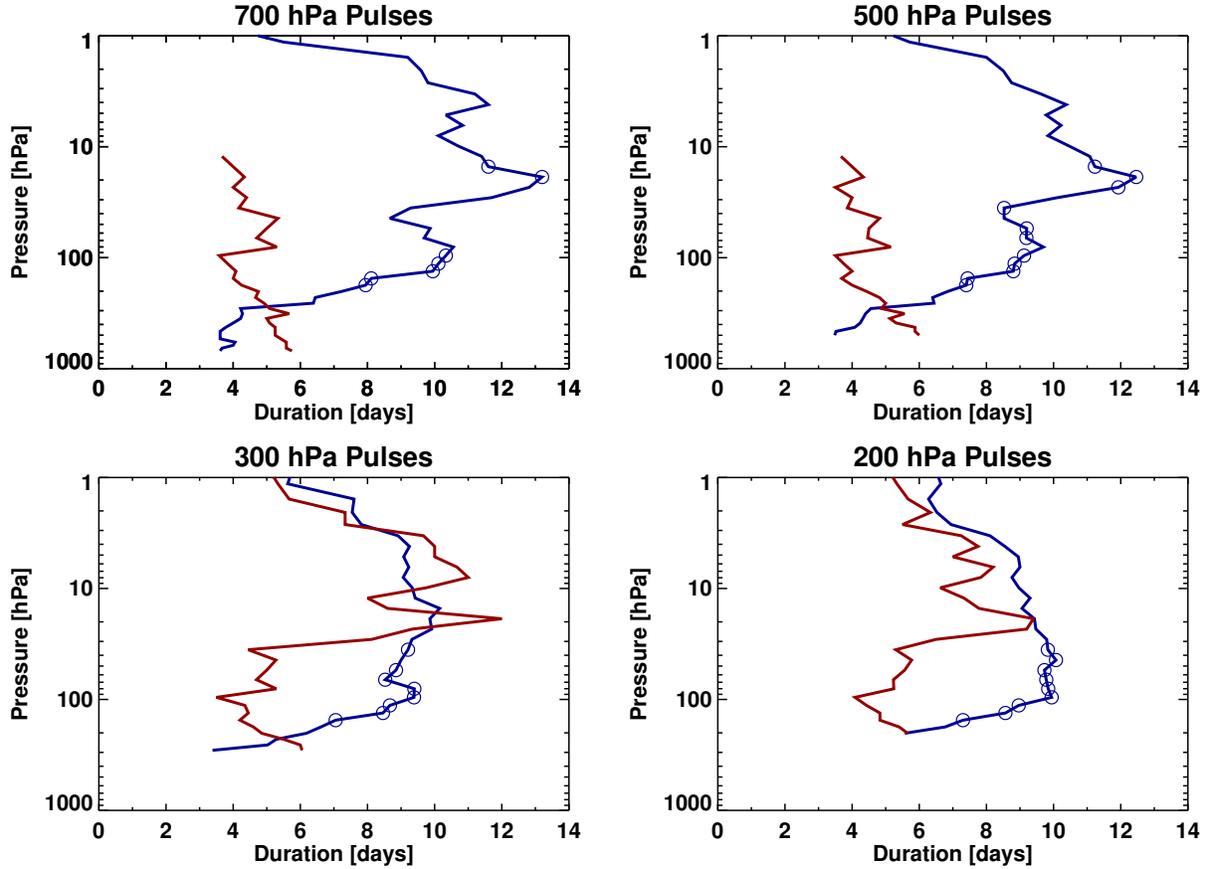


FIGURE 3.13. Mean durations of the increasing pulses (blue) and the decreasing pulses (red) for each starting level. Points which are significantly larger (at the 95% level) for the increasing pulses are circled.

in height increases, especially because the biggest changes in durations appear below these levels (between ~ 500 hPa and 200 hPa).

To show evidence for the prior hypothesis that increasing pulses are indicative of strong wave-mean flow positive feedback, we compare principal characteristics between the increasing and decreasing pulses. Figure 3.13 shows the mean durations for increasing pulses (blue) and decreasing pulses (red). Levels where the mean duration of increasing pulses is significantly longer than the mean duration of decreasing pulses are identified by blue circles. There is general overlap in the locations of significant difference amongst the four starting levels; increasing pulses at all starting levels are significantly longer just below and just

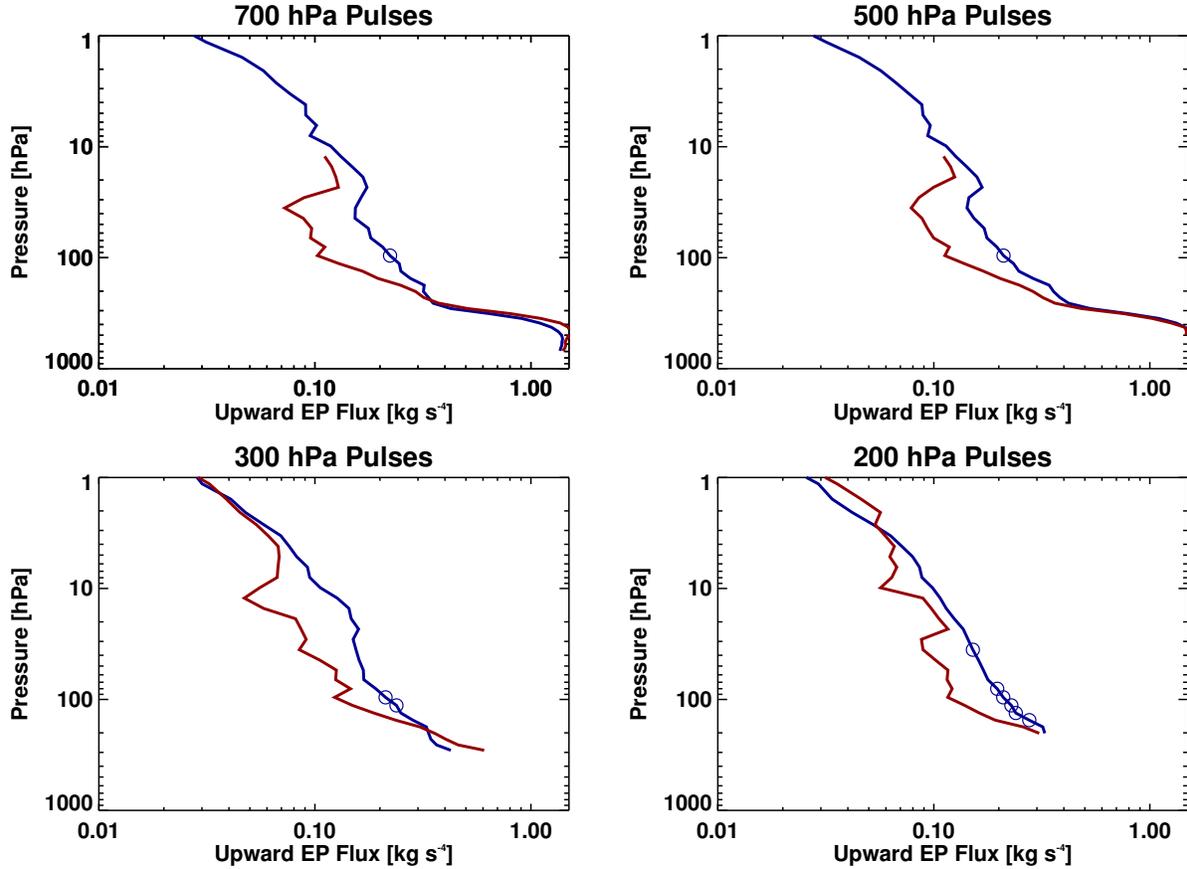


FIGURE 3.14. Means of the maximum upward EP flux amplitudes for increasing pulses (blue) and decreasing pulses (red) for each starting level. Points which are significantly larger (at the 95% level) for the increasing pulses are circled.

above 100 hPa. The result here is clear: increasing pulses have longer durations than do the decreasing pulses, especially, though not unexpectedly, at the level at which we define a pulse to be deep. Of interest to note is that the durations just below 10 hPa for the 700 hPa and 500 hPa starting levels are also significantly longer.

Next, Figure 3.14 shows the means of the upward EP flux amplitudes for increasing pulses (blue) and decreasing pulses (red). Increasing pulses at all starting levels are significantly larger in amplitude at 100 hPa. This significant difference extends downwards as the starting level is raised to higher levels. Though the amplitude differences are not significant above

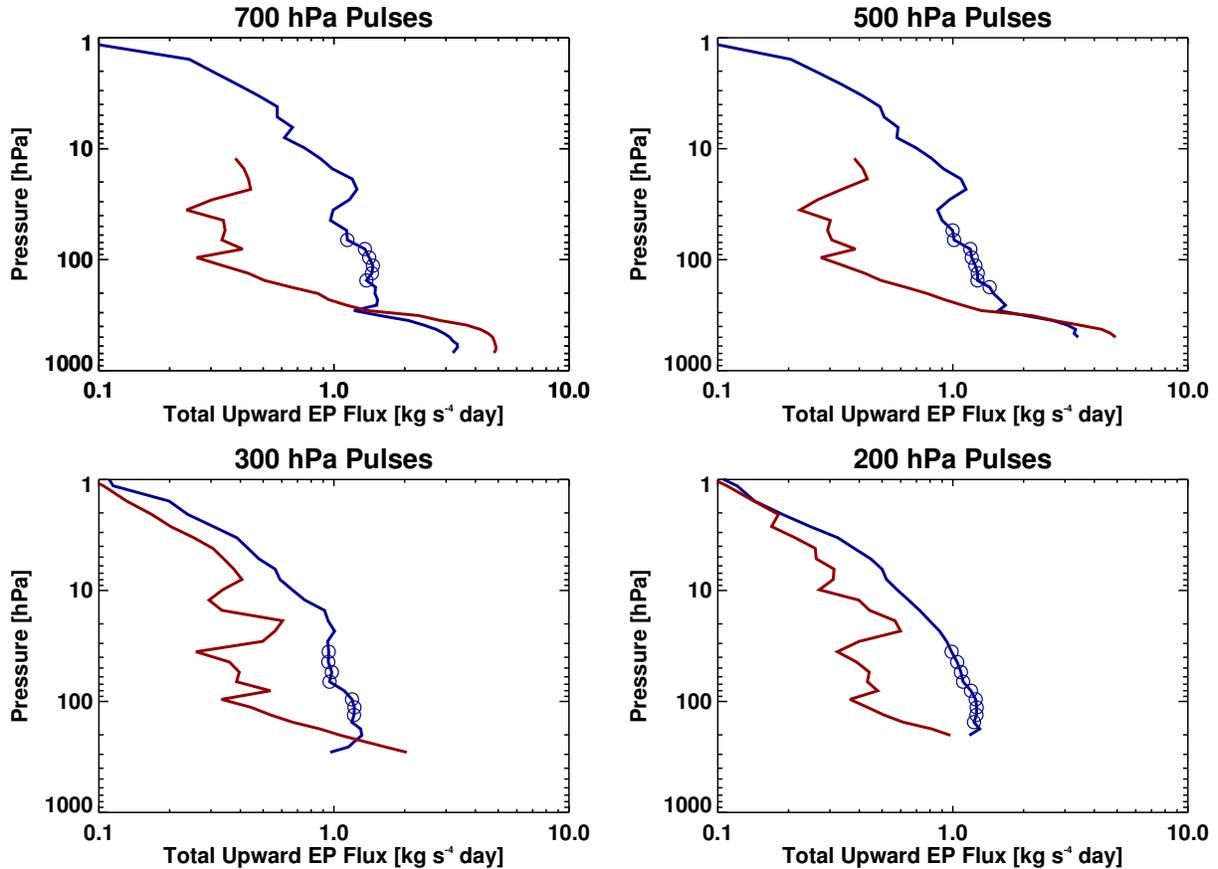


FIGURE 3.15. Means of the total upward EP flux amplitudes for increasing pulses (blue) and decreasing pulses (red) for each starting level. Points which are significantly larger (at the 95% level) for the increasing pulses are circled.

100 hPa, the results show the possibility that the amplitudes of increasing pulses are on average larger to at least 10 hPa.

Finally, Figure 3.15 shows the total upward EP flux for increasing and decreasing pulses. The different starting levels have a great deal of overlap in significance with almost all levels between 50 hPa and 200 hPa significantly larger for increasing pulses. As this total upward EP flux field represents, in ways, the compounding of the duration and amplitude fields, it quite well demonstrates support for our hypothesis that upward wave pulses which increase in duration with height are indicative of strong wave-mean flow positive feedback. Even where they are not statistically significant, the differences in total upward EP fluxes are

large, indicating that increasing pulses likely have a much larger capacity to force the zonal wind at upper stratospheric levels.

The associated changes in zonal wind are the last diagnostic to consider with regards to these pulses. Recall from the descriptions of the wave-mean flow positive feedback in prior sections that this feedback begins with fast winds which inhibit vertical wave fluxes. Wave fluxes will then dissipate in these fast winds, producing wind deceleration, thereby allowing enhanced wave propagation and deceleration.

Thus in analyzing our two subsets of deep pulses (increasing pulses and decreasing pulses), two hypotheses may be drawn on aspects of the zonal wind state and evolution. Following the preconditioning notion of McIntyre (1982), the first expectation is that the zonal wind prior to increasing pulse events is, on average, anomalously strong⁶. The second expectation is that the zonal wind in the period during and following the increasing pulses will be substantially more reduced than in the period during and following decreasing pulses.

To explore the above hypotheses, we consider ERAi zonal mean zonal wind at 60°N. The seasonal cycle is removed so that only anomalies are considered. For the shown composites of zonal wind, the central date (day 0) is taken to be the date at which the deep pulses – either increasing or decreasing with height – reach 100 hPa from their respective starting levels. Doing so accounts for differences in time taken for pulses to propagate, thereby giving a more consistent sense of the zonal wind response.

Figure 3.16 shows the anomalous zonal wind composited about this central date for increasing pulses at all four starting levels. In all cases, the zonal wind is anomalously strong prior to the wave pulses arriving at 100 hPa. Furthermore, as indicated by the overplotted

⁶A similar or opposite expectation is not placed on the zonal wind prior to decreasing pulse events as the theory does not describe such events.

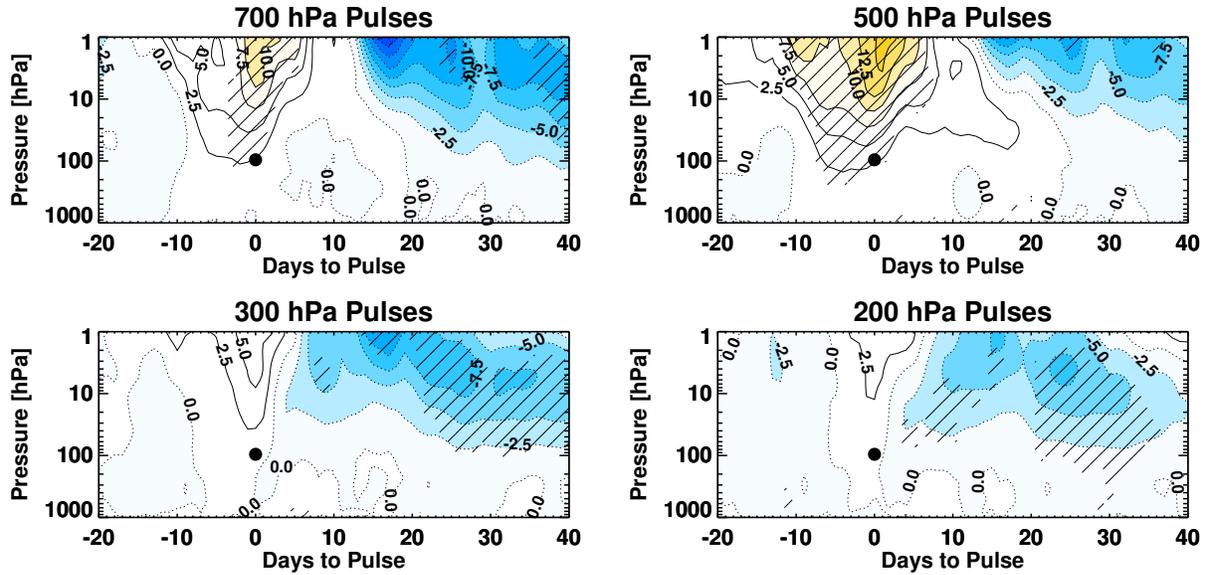


FIGURE 3.16. Anomalous zonal wind composited about increasing pulses with the four starting levels. Regions which are significantly different from 0 ms^{-1} (i.e. are significantly anomalous) at the 95% level are hatched. Day 0 corresponds to the date at which all the pulses in each composite arrive at 100 hPa, indicated by the filled circle.

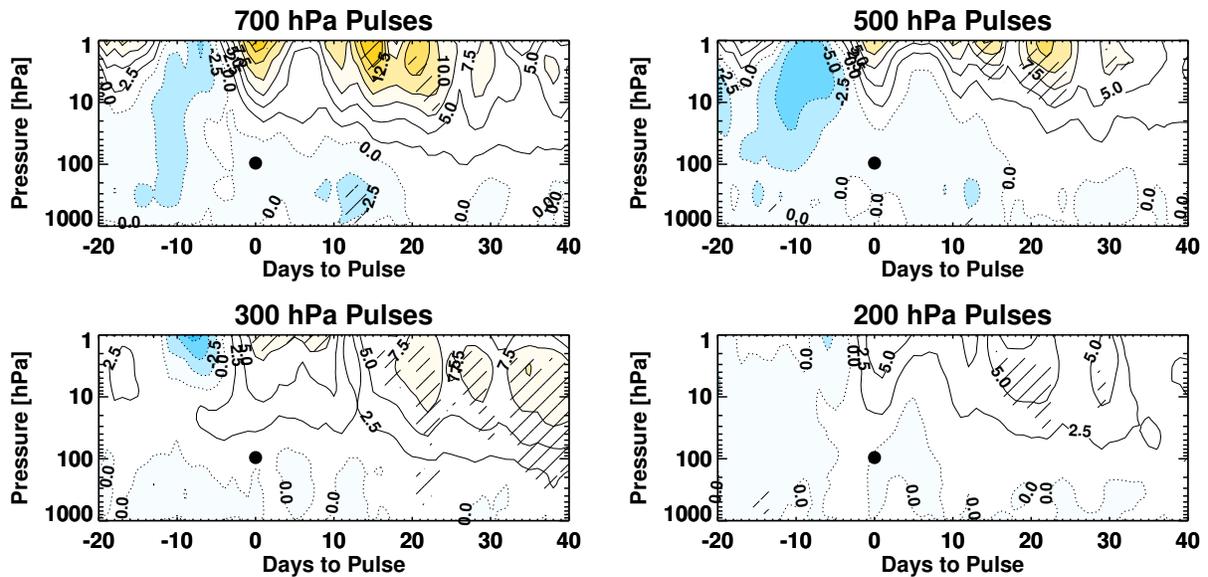


FIGURE 3.17. As in Figure 3.16, but for decreasing pulses.

hatching, the anomalies are significantly large for the 700 hPa and 500 hPa starting levels. This falls in line with the first expectation above: that zonal winds should be strong prior to strong wave-mean flow positive feedback events.

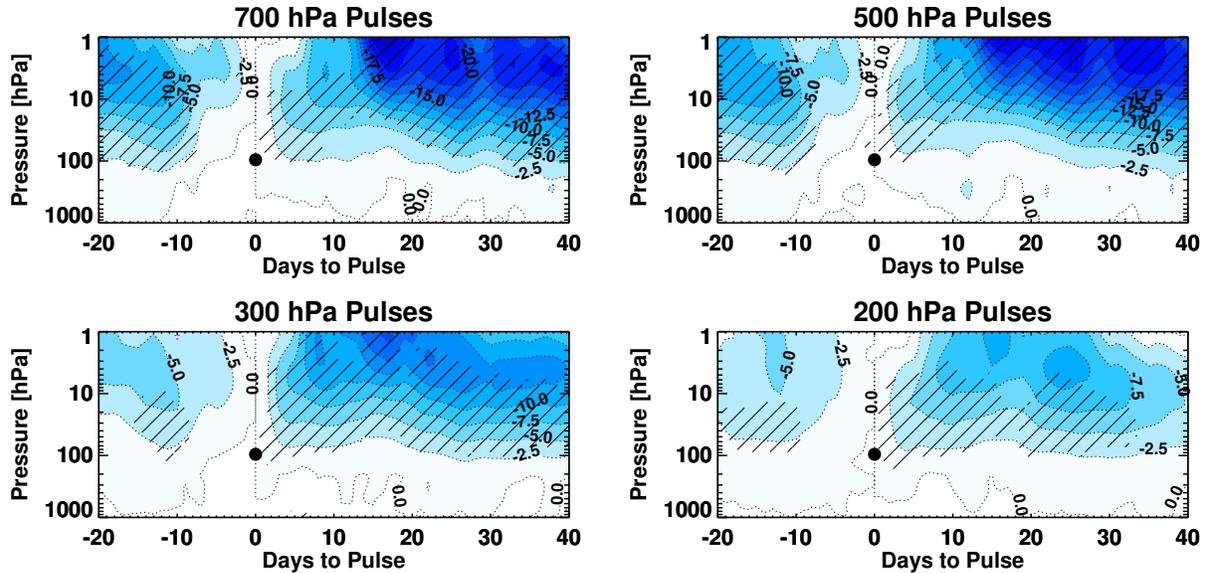


FIGURE 3.18. Changes in zonal wind from the central date for increasing pulses at all starting levels. Regions which are significantly less than the central date zonal wind are hatched. Note that the panels are not centered about the central date (day 0).

For contrast, Figure 3.17 shows the anomalous zonal wind composites for decreasing pulses. The anomalies prior to decreasing pulses tend to be positive at high levels but negative at 100 hPa.

One may also note from Figures 3.16 and 3.17 that the evolution of the zonal wind anomalies following the central date are notably different between decreasing and increasing pulses. To better demonstrate this, we analyze the changes in zonal wind following the central date for both pulse types. To get at these changes, we subtract the vertical profile of zonal wind anomalies at the central date from the zonal wind anomalies for each pulse prior to compositing. Thus, the change in anomalous zonal wind at the central date will be zero by definition and all other values will be changes in the anomalies from this central date zonal wind.

Figure 3.18 shows the composite evolution of changes in zonal wind anomalies from the central date for increasing pulses at the four starting levels. Relative to the central date,

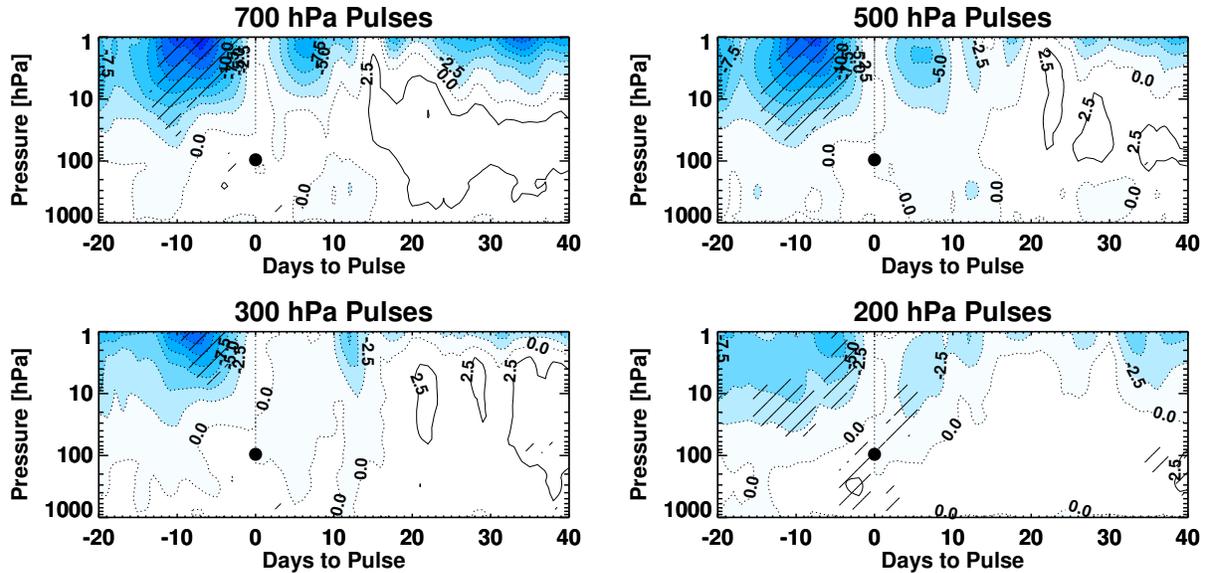


FIGURE 3.19. As in Figure 3.18, but for decreasing pulses.

the zonal wind is significantly decelerated in all four cases. For the 700, 500, and 300 hPa starting level pulses, this significant deceleration is maintained for at least 40 days beyond the central date. The 200 hPa pulses are significantly decelerated but only to ~ 35 days beyond the central date.

Figure 3.19 shows the same field, but for decreasing pulses. The decelerations following decreasing pulses are considerably weaker, with very few regions of significance. As well, for 15 days or longer beyond the central date, the zonal wind at levels including and below 10 hPa becomes anomalously positive.

These results shows that the second expectation is fulfilled: increasing pulses, being associated with strong wave-mean flow positive feedbacks, have much greater decelerations than do decreasing pulses. Thus for these starting levels – and presumably those in between – we have demonstrated that analyzing lower- to mid-stratospheric vertical variations in pulse duration is an effective diagnostic for instances of strong or weak wave-mean flow positive feedback.

3.3. TROPOPAUSE INVERSION LAYER IN REANALYSIS DATA

In the simple modeling framework, we found that an idealized form of stratification mimicking the tropopause inversion layer (TIL) would lead to enhanced vertical wave fluxes, relative to an idealized form of stratification mimicking a simple tropopause. The primary way in which this occurred was through perpetual or forced reversals of the meridional potential vorticity gradient. Such reversals are associated with instability and wave generation.

In seeking to verify the applicability of these results outside of the simplified modeling environment, we have searched for such potential vorticity gradient reversals in other sources. ERAi data show no clear signs of these reversals near the tropopause (T. Birner, personal communication). Thus the perpetual reversal we find above is not a robust condition of the real atmosphere. However, vertical smoothing of fields due to data assimilation is known to decrease the maximum stratification of the TIL (Birner 2006) and could thereby affect how these gradient reversals are resolved. Idealized general circulation modeling with sufficiently high resolution to capture the TIL therefore is a next step in determining the potential impact of such reversals.

3.4. CONCLUSIONS

Characteristics of transient wave forcing leading to major stratospheric warming events are investigated in ERAi reanalysis data. The largest values of wave activity entering the stratosphere are not observed to vary dramatically between years with and without SSW events. Previous studies (Newman et al. 2001, Polvani and Waugh 2004) have shown the 40-day averaged heat flux values to correlate well with SSW events. We find a large part of the difference in 40-day averaged heat flux between nonevent winters and event winters to come from heat flux enhancements past the event date. These enhancements past the event date

can hardly be said to cause the SSW event. Furthermore, the correlation between 40-day averaged heat flux and the occurrence of SSWs does not fully detail the evolution of the wave forcing prior to a SSW. This wave forcing could primarily take one of two forms: a large but rapid burst or a period of moderately large values. By compositing around SSW events, it is shown that the latter of the two forms – a sufficiently long period of anomalously large wave forcing – is most typical (in agreement with Harnik 2009). An additional composite for our identified minor warming events showed no such long-period anomalous wave forcing.

We verify these results by examining the duration of the heat flux exceeding a given threshold for the periods leading up to a SSW. For anomalously large heat flux thresholds (around or just above the climatological mean), this duration is found to be significantly larger during these pre-event periods (~ 1 week) compared with the all winter mean (~ 3 days). This difference in forcing duration becomes statistically insignificant for wave forcings stronger than about half a standard deviation above the climatological value, showing that strong and in particular extreme wave forcings are generally not very long lived preceding SSWs. This suggests that the strength of the forcing plays a secondary role in forcing SSWs; rather it appears crucial for the forcing to maintain a minimum strength (around or just above the climatological value) over a sufficiently long period. It is important to note that the above results are of statistical nature: large variations in forcing duration and strength preceding SSWs exist. Nevertheless, our results indicate certain characteristics of the wave forcing that appear favorable for producing SSWs. SSWs have often been linked to tropospheric blocking events (e.g. Martius et al. 2009). Blocking events are characterized by long time scales, consistent with our findings. However, the majority of blocking events do not produce an SSW, suggesting that internal stratospheric characteristics play an important

role in allowing tropospheric wave activity of long time scale to propagate sufficiently deep into the stratosphere (cf. Scott and Polvani 2004).

The heat flux results are bolstered by considering pulses of upward EP flux data from ERAi data. As with the heat flux analysis, we define pulses as continuous exceedance of a given value; here we take this value to be the seasonally-evolving mean. Linear statistics show that there is good correlation between the durations of these pulses and their maximum amplitudes, a result which is consistent with a red noise spectrum. Furthermore, by analyzing durations at all pressure levels, we find that pulses with duration exceeding 20 days are highly uncommon in ERAi. Together, these results cohere with those from the heat flux analysis: neither extremely large amplitude but short (~ 1 day) pulses nor extremely long (~ 100 day) pulses are readily observed. Thus, the reanalysis data are consistent with the notion of an order 10 day preferential wave forcing time scale.

An algorithm was developed and implemented to identify vertical-coherence in the above upward EP flux pulses so as to analyze vertical variations in pulse durations. By considering these vertically-coherent pulses of EP flux, we find that the upper troposphere to lower stratosphere region to be a crucial region for the wave-mean flow positive feedback. Whereas there are significant increases in the duration of pulses up to 100 hPa, above 100 hPa, the characteristics of the pulses do not largely vary with height. This highlights a possible sensitivity to analyzing wave pulses prior to strong wave-mean flow interaction events – such as SSWs – at 100 hPa or higher.

Sets of pulses with deep vertical coherence were selected and subsequently partitioned into subsets of pulses with increasing duration with height and pulses with decreasing duration with height. The hypothesis is that the pulses which increase duration with height indicate that strong wave-mean flow positive feedback is occurring. This hypothesis may be shown to

be true if, at least, 1) the precedent zonal winds are fast until deceleration by the increasing pulses, and 2) the deceleration following the pulses is notably strong relative to other pulses (e.g. decreasing pulses).

It is found, in general, that durations are significantly larger near and below 100 hPa for increasing pulses than for decreasing pulses. The amplitudes in this region are also significantly larger for increasing pulses. Composites of anomalous zonal winds show that the winds are indeed fast prior to increasing pulses and that the decelerations following these pulses are much larger than the decelerations following decreasing pulses. The expectations thus upheld, we may conclude that this diagnostic – increasing pulse duration with height – is an effective means of identifying strong wave-mean flow positive feedbacks.

We search for reanalysis support of our TIL results from the simple model, but find that it is not upheld. This may be due to a number of reasons, particularly that reanalysis data has too high of vertical smoothing of fields due to data assimilation to properly resolve the TIL stratification. Since this is an unresolvable problem of the observational data utilized here, we are left to investigate this topic in higher order models (general circulation models).

To summarize, we find reanalysis data support for our notion of preferential wave forcing time scales on the order of 10 day, we find that increasing pulse durations with height is a useful diagnostic for strong wave-mean flow positive feedbacks, and we find that reanalysis data do not contain strong evidence of wave generation due to the TIL.

The reanalysis data thus presents us with a mixed bag of results, lending support to some prior findings and unable to support others. While these reanalyses are extremely useful sources of global data at regular intervals, they are not without their problems. It is possible that the features we find – and those which we do not – are poorly represented and thus potentially misleading. This is however not likely for the large-scale features we

consider (e.g. zonal mean zonal wind and EP fluxes), but such representation errors are undetermined for small-scale features (e.g. TIL stratification). Given this uncertainty, we continue with our hierarchical approach to studying these wave-mean flow interaction and feedback problems through an intermediate step to simple models and observations: a general circulation model.

CHAPTER 4

GENERAL CIRCULATION MODEL WAVE-MEAN FLOW POSITIVE FEEDBACKS

4.1. MODEL INTRODUCTION

For the high order model experiments performed here, we utilize a general circulation model (GCM) based on the primitive equations in spherical coordinates. The specific model is the Geophysical Fluid Dynamics Lab (GFDL) “Memphis Patch 1”. We use the dry dynamical core of this model – i.e. the model does not resolve moisture, chemistry, and radiation, amongst other fields. The spectral-truncated version of this model is used. Thus we set the horizontal resolution through spherical harmonics, while the vertical and temporal resolutions are specified as straightforward inputs.

Because the model does not include radiation or chemistry, we must specify both a background radiative equilibrium state to which the model will damp and the radiative damping which the model will use to do so. Following Polvani and Kushner (2002), this background state is set by an analytical radiative equilibrium temperature setup appropriate for studying wintertime polar stratospheric dynamics. For this setup, the tropospheric equilibrium temperature is prescribed to a highly idealized state following Held and Suarez (1994). The stratosphere, unlike in Held and Suarez’s setup, is set to approximately mimic the radiative equilibrium temperature profile during winter solstice as shown by Fels (1985). The prominent feature in the winter stratosphere, as described in the Introduction, is the polar vortex. This feature is included in the specified radiative equilibrium temperature profile through a relatively cold pool bounded in height by a nominal tropopause and the model lid, and in latitude by the pole and $\sim 50^\circ$ latitude. The most notable parameter of this specified profile is the “strength” parameter that controls how cold the vortex becomes. This parameter is

denoted by γ and will be set to 4 K km^{-1} unless otherwise noted. Newtonian damping is specified as in Held and Suarez (1994).

To induce large-amplitude planetary waves, we include idealized topography following Gerber and Polvani (2009), who specify zonal topography of a given wavenumber. This is done by setting the surface geopotential height Φ_0 as

$$(4.1) \quad \Phi_0(\lambda, \phi) = g \sin^2 \left(\frac{\phi - \phi_0}{\phi_1 - \phi_0} \pi \right) (h_0 \cos(m\lambda) + h_{adj}) \quad \phi_0 < \phi < \phi_1,$$

and $\Phi_0 = 0$ for all other ϕ . Here, λ is longitude, ϕ is latitude, g is the gravitational constant, h_0 is the topographic height, h_{adj} is the mean height of the topography, and m is the zonal wavenumber of the topography. For the experiments shown here, the latitudinal bounds ϕ_0 and ϕ_1 are 25°N and 65°N , respectively, and m is set to 2.

For all experiments here, we set h_{adj} to 0 m. We found that the results and the model climatology are both quite sensitive to nonzero values for this parameter. Given that exploring such a sensitivity falls outside the scope of the current dissertation work, investigation of nonzero h_{adj} is left for future work.

The following settings are utilized to integrate the GCM. The time stepping is 600 s; output data resolution is one day. Integrations are 8000 days in length, with the first 2500 days discarded for spin-up. We use 40 vertical layers at specified sigma levels, the values of which are described in Polvani and Kushner (2002); interpolation from these sigma levels to pressure levels was performed in post-processing. We used triangular truncation of order 42 (T42). This corresponds to 128 longitudinal points and 64 latitudinal points in the horizontal. 8th order diffusion was applied. All other settings are as in Polvani and Kushner (2002).

4.2. TOPOGRAPHY EXPERIMENTS

4.2.1. TIME MEAN STATE.

We begin with the GCM by performing integrations at various topographic heights. As discussed in Gerber and Polvani (2009) and Jucker et al. (Submitted), the frequency of modeled major SSWs displays a large sensitivity to the topographic height h_0 . Namely, the frequency of SSWs goes up for larger h_0 . Both of these studies discuss this point, but only Gerber and Polvani find that there is a large topographic parameter h_0 at which the SSW frequency becomes smaller. Given that Jucker et al. utilize a different profile of radiative equilibrium temperature than that which we (and Gerber and Polvani) use, we expect our results to fall more in line with Gerber and Polvani (2009). We however utilize a much larger and refined parameter space for h_0 and thus more fully draw out the sensitivity than either of these two previous studies.

For this analysis, SSWs are parsed into both ‘major’ and ‘minor’ SSWs. Major SSWs are identified as zonal mean zonal wind reversals at 60°N, 10 hPa which are separated by at least 20 days. Minor SSWs follow the traditional definition of minor SSWs, in which the temperature gradient between 60°N and 90°N reverses. Note that the events are allowed to overlap, which differs from typical definitions. Figure 4.1 plots the SSW occurrences as a function of topographic parameter for both major SSWs (red) and minor SSWs (blue).

Considering this strong dependence of SSW occurrence on topographic parameter, we study the various states of the model with particular emphasis on pulses of waves. The pulse diagnostics we use have been introduced and discussed in Chapter 3; these include pulses at all levels, vertically-coherent pulses, and pulses with either increasing or decreasing duration with height. Our objective here is to identify what determines this particular dependence of SSW frequency on topographic parameter and how aspects of the wave-mean flow positive

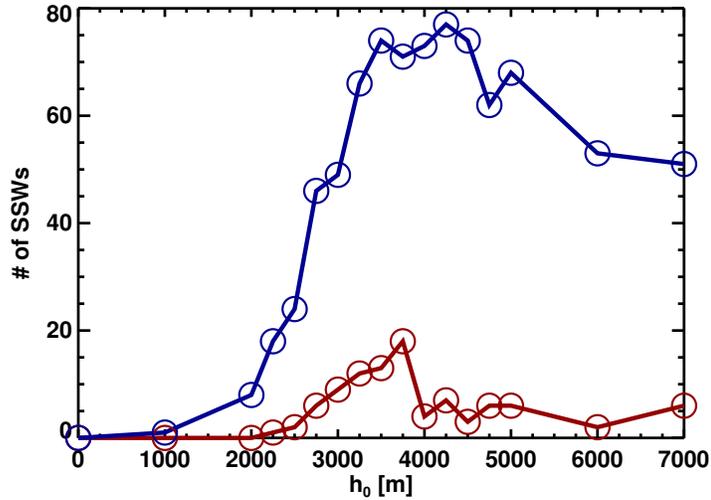


FIGURE 4.1. Occurrence of two classes of SSWs in the GFDL GCM integrated in perpetual winter for 5500 days with different levels of topographic height parameter h_0 . Major SSWs are plotted in red and minor SSWs are plotted in blue. Circles outline the actual data points.

feedback which we have heretofore studied – such as wave forcing time scales – manifest here.

We firstly analyze the time mean state of the model through the fields shown in Figure 4.2: zonal mean zonal wind, temperature, EP flux divergence, and upward EP fluxes, all averaged between $45\text{--}75^\circ\text{N}$. Note that the upward EP fluxes in this chapter follow the definition given by Eq. 3.1, but are multiplied by the radius of Earth (≈ 6371 km). All fields here are plotted as a function of topographic parameter and pressure.

As h_0 increases, the zonal wind (upper left) naturally decreases to a minimum at the topographic height associated with the maximum occurrence of SSWs. The upper level zonal winds become somewhat faster for still larger height parameter, as expected with the associated decrease in SSWs. Temperature (upper right) follows a similar evolution: warming in the middle stratosphere in association with the maximum occurrence of SSWs and cooling at still larger topographic heights. The same is true of EP flux divergences

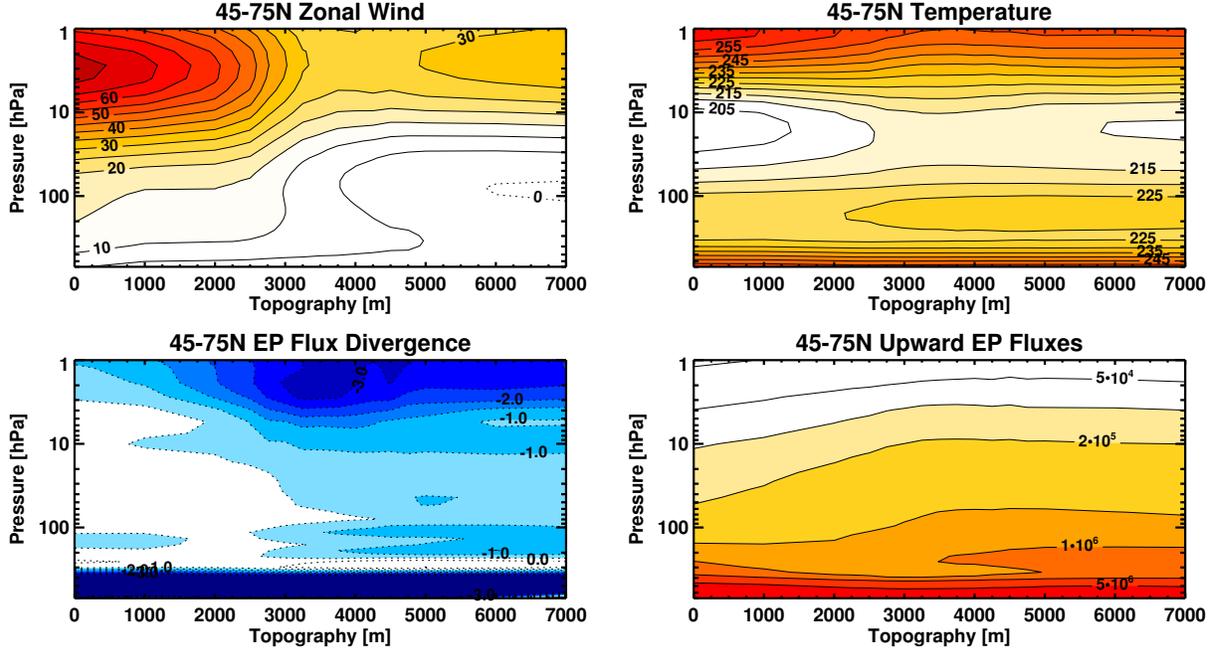


FIGURE 4.2. Time mean zonal wind (upper left), temperature (upper right), EP flux divergence (lower left), and upward EP fluxes (lower right), all averaged between $45\text{--}75^\circ\text{N}$ and plotted as functions of topographic parameter and of pressure. The zonal wind contour spacing is 5 ms^{-1} , the temperature contour spacing is 5 K , the EP flux divergence contour spacing is $0.5 \text{ ms}^{-1}\text{day}^{-1}$, and the contour spacing for the upward EP fluxes is on a logarithmic scale.

(bottom left). A maximum in upper level convergence coincides with the maximum in SSW occurrence, with smaller values of convergence at both shorter and higher topography.

The upward EP flux evolution is more distinct than these fields. Above 100 hPa, the upward EP fluxes become larger for higher topography up to the topography at which the maximum in SSWs occurs (3750 m). These upper level EP fluxes decrease slightly for still larger topography. In the middle troposphere to lowermost stratosphere, the upward EP fluxes decrease from the no topography integration until ~ 3000 m topography. Beyond this, the EP fluxes increase slightly.

We note that even in the no topography case, where baroclinic instabilities on synoptic scales dominate, planetary wave activity fluxes are generated and of comparable magnitude to those amplitudes with nonzero topography. The primary means for this generation is

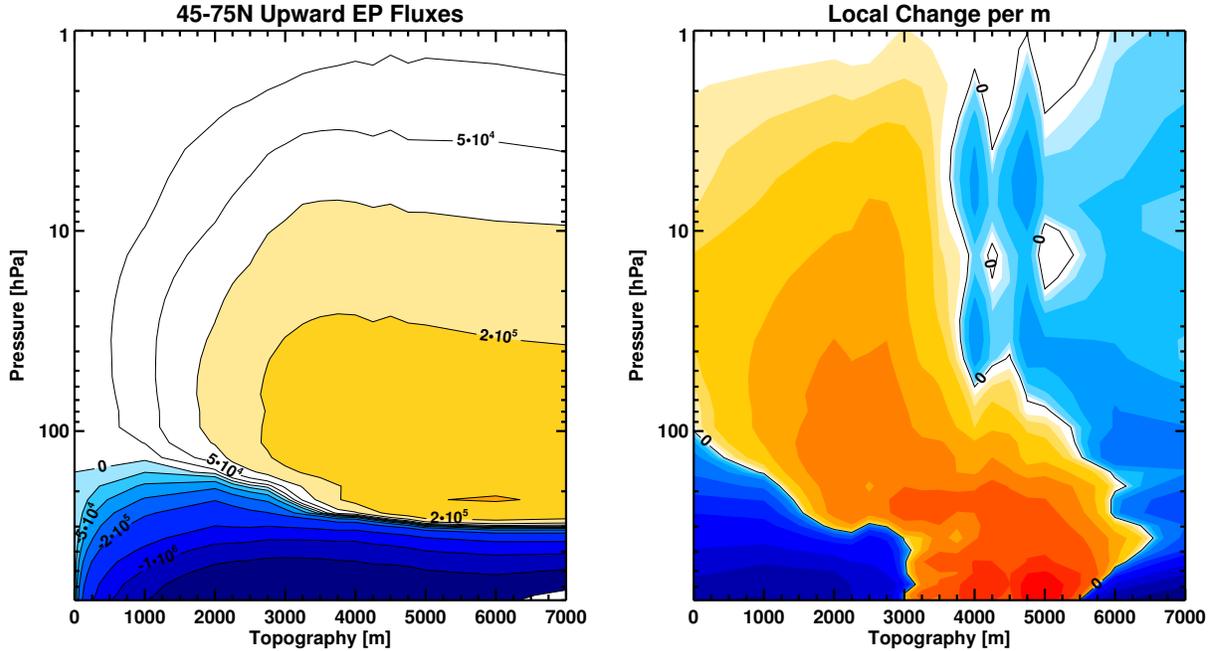


FIGURE 4.3. (Left) Differences in time mean values of upward EP fluxes from the 0 m integration. Positive EP flux values indicate larger values at that pressure level and topographic parameter than at that pressure level in the no topography integration. (Right) Local change in the time mean 45-75°N upward EP fluxes per meter change of topographic parameter as a function of pressure and topographic parameter. Positive values (warm colors) indicate where the upward EP flux increases as the topographic height increases at a given topographic parameter.

through wave-wave interactions leading to inverse energy cascade. Though most turbulence within the model will act to flux energy to smaller spatial scales, a nontrivial amount of this turbulence will behave as two-dimensional homogeneous turbulence as a result of small vertical scales relative to forcing length scales (Celani et al. 2010). This two-dimensional turbulence fluxes energy to larger scales, leading to nonzero planetary-scale wave amplitudes despite no apparent forcing mechanism.

The changes in wave activity fluxes with changing topography are elucidated by showing in Figure 4.3 the anomalies of upward EP flux relative to the no topography case. Shown as well are the local changes in upward EP flux per meter change in topography. This figure clearly shows how the upper level EP fluxes increase up to ~ 4000 m topography and

subsequently decrease, while also showing the lower level decrease to ~ 3000 m topography and subsequent increase. Only for a small parameter range between ~ 3000 m and 3750 m does the upward EP flux locally increase throughout the depth of the troposphere and stratosphere for increasing topography.

That this range of increasing upward EP flux at all levels corresponds to the range of maximum occurrences of SSWs is likely not due to coincidence. Rather, it is consistent with our understanding of the wave-mean flow positive feedbacks. The stratospheric wave fluxes have reached sufficiently large amplitude so as to strongly and frequently induce the feedback cycle, amplifying stratospheric and upper tropospheric wave amplitudes. For larger topography than 4000 m, the large stratospheric wave amplitudes instead seem to stifle SSW variability by maintaining consistently slow winds, especially in the lower stratosphere (cf. Figure 4.2). These low zonal winds in the lower stratosphere are not conducive to upward wave fluxes (as will be shown below) and thus inhibit the ability of the waves to force strong zonal wind variability aloft. This may be seen through the faster time mean zonal wind above 10 hPa found at topographic heights of 5000-7000 m.

An additional feature of interest from Figure 4.3 is that the tropospheric upward EP fluxes for nonzero topography are always smaller than the 0 m topography integration, in no small part because of the strong decrease between 0 m and 3000 m topographic height. Histograms are useful here in determining why the tropospheric upward EP fluxes would decrease as topographic height is increased. Histogram analysis can identify changes in the distribution about the mean value for a field, where increases in negative or decreases in large positive upward EP flux occurrence as topography changes may be easily quantified.

Here, histograms of 500 hPa upward EP fluxes are calculated. Figure 4.4 shows the total occurrences of a given upward EP flux for all integrations as contours. As topography is

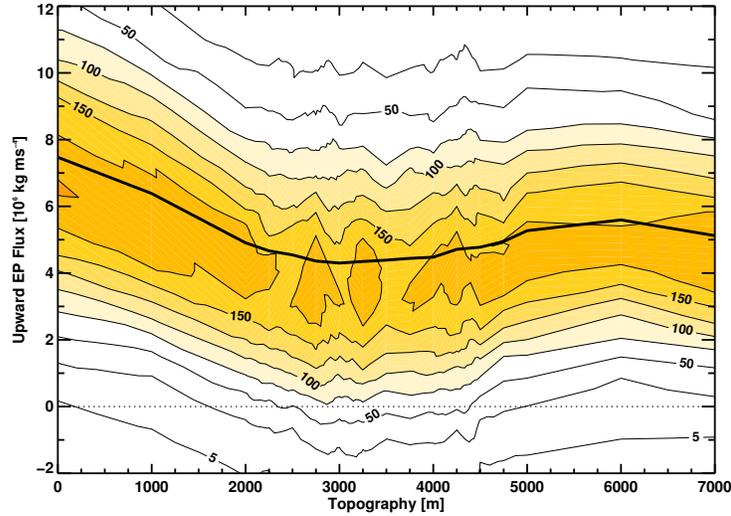


FIGURE 4.4. Contours of total occurrences of a given values of upward EP flux at 500 hPa as a function of topographic parameter. The contour spacing is 25 occurrences; the 5 occurrence contour is also plotted. The thick solid curve marks the mean value of the 500 hPa upward EP fluxes. The dotted line denotes zero upward EP flux.

increased to 3000 m, the largest amplitudes of 500 hPa upward EP fluxes occur less frequently and negative amplitudes occur much more frequently.

To find from where these changes in the distributions come, Figure 4.5 shows contours of the total occurrences of a given value of 500 hPa upward EP flux for all integrations, here decomposed into wavenumbers 1, 2, 3, and 4+ (top left, top right, bottom left, and bottom right, respectively). The contour spacing is doubled, but the scale of the ordinate is the same as that in Figure 4.4 so that the results are directly comparable.

Wavenumber 1 upward EP fluxes display no substantive change across these four integrations, though close inspection shows minor differences in the shape of the histograms with slight shifts towards lower values. Changes toward smaller and negative values with topography for wavenumber 3 are larger, but still not large enough to explain the full field changes.

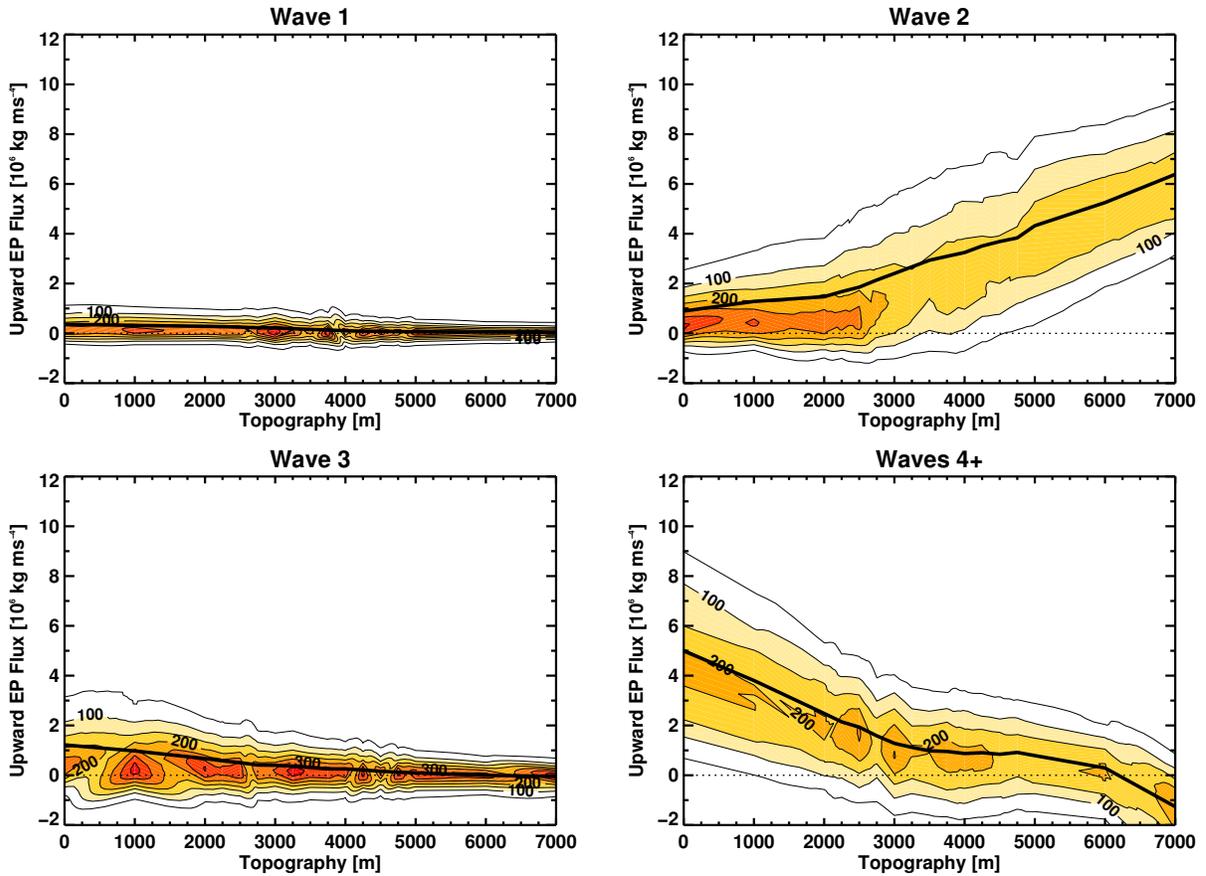


FIGURE 4.5. As in Figure 4.4, but for waves 1 (upper left), 2 (upper right), 3 (lower left), and 4+ (lower right). The contour spacing here is 50 occurrences.

Rather, changes in the synoptic and small-scale waves (wavenumbers 4+) affect the full wavenumber distribution most strongly. The bottom right panel of Figure 4.5 shows that the wavenumber 4+ EP fluxes become more frequently negative (downward wave propagation) and are more rarely large (relative to the 0 m integration) as topography increases to 3000 m. The distributions remain fairly similar for topographic heights between 3000 m and 5000 m. For higher topography, the mean of the wavenumber 4+ EP fluxes becomes negative.

In opposition to these decreases from wavenumbers 1, 3, and 4+, the wavenumber 2 upward EP fluxes generally increase as the topography increases. This is both desired and expected, and is reflected in the shift of the mean of the histogram to large positive values. For topographic heights of 4000 m or higher, the increases in wavenumber 2 upward EP fluxes

are sufficient to overcome the decreases in other wavenumbers, resulting in a net increase in upward EP flux at 500 hPa.

These wavenumber 2 increases are only large beyond ~ 2500 m topography, however. For small topographic heights, the peak of the distribution remains nearly steady while the mean only increases slightly. This steady response to increasing topographic forcing of the same wavenumber might suggest that tropospheric wave trapping may be affecting the amplitudes. In this argument, wave reflection above results in anomalous downward EP fluxes back to this layer having the net effect of reducing the wave amplitudes there (e.g. Boville 1984, Perlwitz and Harnik 2003).

To determine if this is so, we consider an index of refraction argument. Simply, for a wave with a particular zonal wavenumber and a particular phase speed, positive index of refraction indicates regions where vertical wave propagation is possible while negative index of refraction indicates vertical propagation is not possible. At the levels where the index of refraction becomes zero or less, the waves will partly (or perhaps totally) reflect with evanescence above. If the steady amplitudes of upward EP flux at 500 hPa across this range of topographic heights is due to wave reflections above, then the index of refraction at a relatively near upper level should be negative for a sufficiently frequent amount of time.

The index of refraction n_{ref}^2 is calculated here for wavenumber 2, stationary waves following Harnik and Lindzen (2001):

$$(4.2) \quad n_{ref}^2 = \frac{\overline{q_y}}{\overline{U}} - \left(\frac{k}{r \cos \phi} \right)^2 - \left(\frac{f}{2NH} \right)^2$$

where $\overline{q_y}$, the meridional gradient of potential vorticity is given by

$$(4.3) \quad \overline{q_y} = \beta - \frac{1}{r^2} \partial_\phi \left(\frac{\partial_\phi [\overline{U} \cos \phi]}{\cos \phi} \right) + \frac{f^2}{R_d} \left(\frac{p\overline{\theta}}{T} \frac{\partial_p \overline{U}}{\partial_p \overline{\theta}} \right)$$

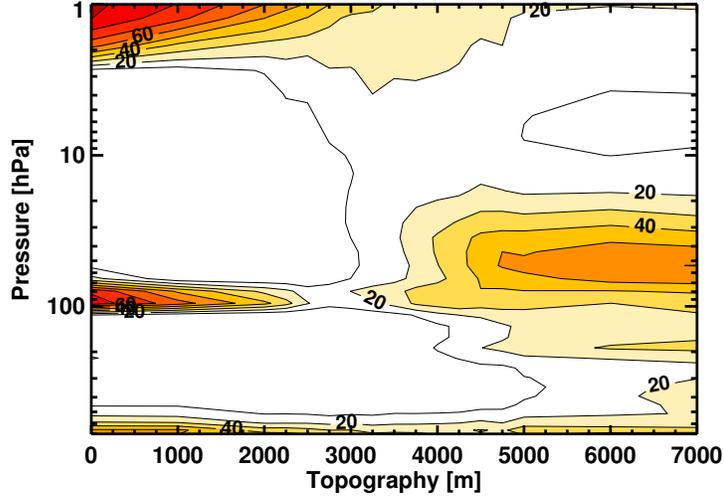


FIGURE 4.6. Percentage of time for which the 45-75°N index of refraction is less than or equal to zero at a given pressure as a function of topographic parameter. Zero or negative index of refraction implies inhibited wave propagation.

Here, k is zonal wavenumber, r is the radius of Earth (≈ 6371 km), $f = 2\Omega \sin \phi$ is the Coriolis parameter, N is the square root of the stratification, $H = 7000$ m is the scale height, β is the meridional derivative of the Coriolis parameter, R_d is the gas constant for dry air, T is the air temperature, and θ is the potential temperature. For this calculation, we follow Gerber (2012) and Simpson et al. (2009) and set N to a constant equal to $\sqrt{3.5 \cdot 10^{-4}} \text{ s}^{-1}$.

Figure 4.6 shows the percent of time over the integration during which the calculated wavenumber 2 index of refraction was less than zero as a function of topographic parameter and of pressure. This figure shows that, for small topography, the condition for wave reflection was met near 100 hPa for more than half of the integration. Between 0 m and 1000 m topography, the 100 hPa time mean meridional gradient of potential vorticity $\overline{q_y} \approx \beta$ and thus is relatively large (see Eq. 4.3). Though the 100 hPa time mean zonal winds for these topographic heights are only $\sim 15 \text{ ms}^{-1}$, this is sufficiently large to make n_{ref}^2 reverse in this region. For larger topography, the zonal wind in this region decelerates while $\overline{q_y}$ is nearly constant, resulting in a decrease of the frequency of negative index of refraction. That this

waning frequency coincides with the increases in upward EP flux at 500 hPa lends support for the hypothesis of reflection-induced amplitude reduction.

Also shown in Figure 4.6 is that, for topographic heights above 4000 m, a broad area below the middle stratosphere has negative wavenumber 2 index of refraction for at least half the integration length. This feature coincides with a time mean layer of easterlies that arises between ~ 110 hPa and 70 hPa for topography higher than 6000 m. These easterlies correspond to a critical layer for wave propagation at which the upward propagating waves will be both reflected and absorbed.

To show the effect on the upward EP fluxes of this broad region of frequent index reversal and associated critical layer, Figure 4.7 shows the anomalies (from the 0 m topography integration) of upward EP fluxes and associated local changes per meter of increasing topography. For topographic heights above 4000 m, the stratospheric level upward EP fluxes are nearly constant or slightly decreasing. Despite large increases in the incoming upward EP fluxes from the troposphere, the critical layer in the broad region of frequent index of refraction reversal results in the lower stratosphere preventing these wave activity fluxes from propagating through the region. We note that this result is in line with Gerber and Polvani (2009), but is shown in greater detail here.

Contrarily, the stratospheric upward EP fluxes for topographic heights between 0 m and 2500 m monotonically increase for increasing topography. This increase occurs in spite of the region of frequent index of refraction reversal near 100 hPa. The difference here is that this region of index reversal is relatively quite shallow compared to the region of index reversal for topographic heights larger than 4000 m. With considerably more levels at which the waves may reflect in the broad region for large topography, the integrated effect is to hold

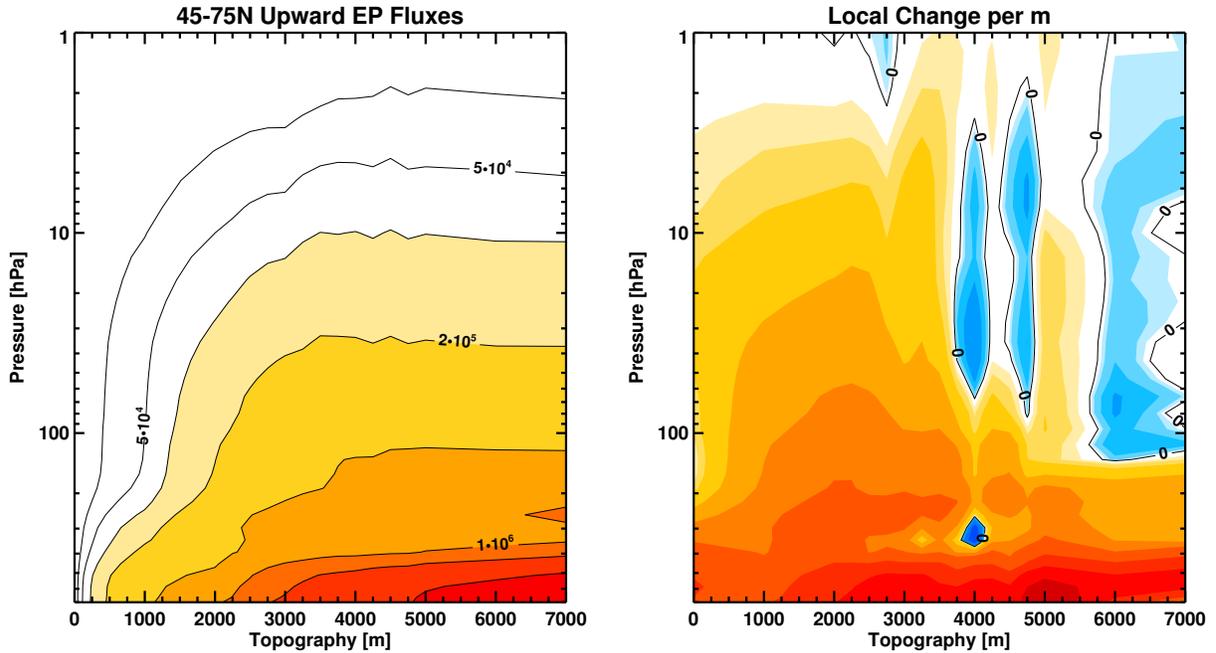


FIGURE 4.7. As in Figure 4.3, but for wavenumber 2 EP flux fields.

the upper level EP fluxes steady. The shallow region of reversal at small topographic heights thus may “leak” wave activity flux more easily.

While the above index of refraction arguments help explain the changes in wave 2 upward EP fluxes with increasing topography, it remains to be shown why the synoptic scale waves in the troposphere decrease in amplitude as the wavenumber 2 topography is raised. This may result from stratosphere-troposphere coupling in which changes of circulation and wave propagation within the stratosphere affect the tropospheric circulation and waves. Here, this may manifest as anomalous downward synoptic scale wave fluxes which occur ever more frequently and/or are of larger magnitude as the topography increases.

Another source of synoptic scale wave suppression may be from changes in the energy cascade of the system with changing topography. It is possible that for progressively higher wavenumber 2 topography, the model atmosphere produces stronger wavenumber 2 EP fluxes which break and dissipate less in the middle troposphere. This decrease in breaking and

dissipation will reduce the energy which cascades to larger wavenumbers there. These wave 2 EP fluxes may instead drive this energy cascade at higher levels. Through a similar histogram analysis at 300 hPa, we find some evidence of this (see Appendix), but this is highly inconclusive without further study (e.g. enstrophy budgets).

4.2.2. PULSE ANALYSIS.

We now consider characteristics of upward EP flux pulses and their relation to the topographic height parameter. We begin by identifying all instances of pulses on all levels and compositing characteristics of their durations and maxima. Figure 4.8 shows these composite characteristics as functions of topographic height and pressure. The durations of upward wave pulses (left) increase at all pressure levels up to ~ 3000 m topography. Below 10 hPa, these durations remain nearly steady for still higher topographic heights, while durations above 10 hPa are at a maximum between 3000 m and 5000 m. The pulse maxima of upward EP flux (right) are shown as anomalies from the 0 m topography integration. These pulse anomalies match the behavior of the full upward EP flux anomalies: in the troposphere, the amplitudes are everywhere less than the no topography integration, while in the stratosphere, the amplitudes maximize for the topographic height at which the maximum in SSWs occurs.

Vertically-coherent pulses are found following the algorithm described in Section 3.2. Figure 4.9 shows the mean durations and mean amplitude anomalies (relative to the 0 m integration) of upward EP flux for four starting levels: 700, 500, 300, and 200 hPa. The results are highly analogous to the results for all pulses shown above, but with slightly longer durations and higher amplitudes. The primary difference is that these vertically-coherent pulses show a clearer maxima in both fields between 3000 m and 5000 m.

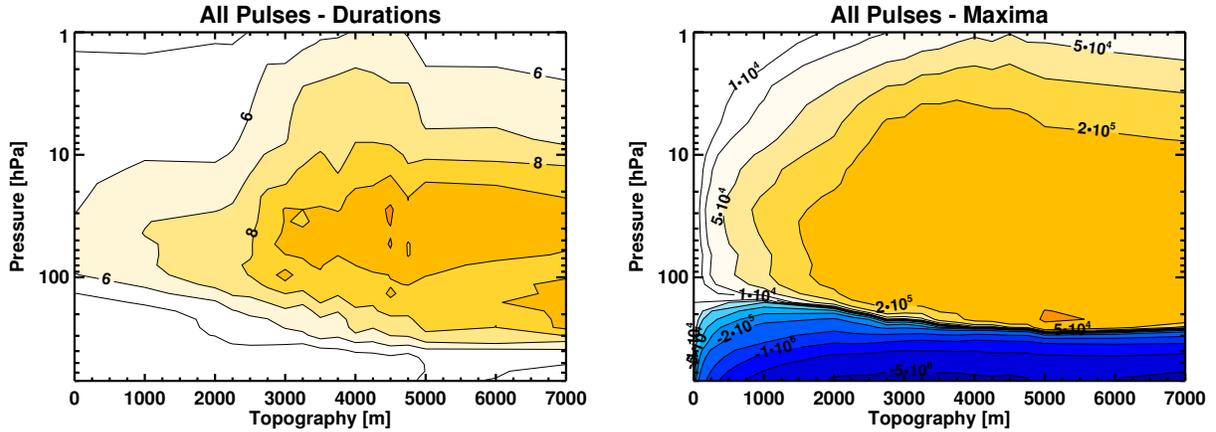


FIGURE 4.8. Durations (left) and 45-75°N upward EP flux maxima (right) for all pulses as a function of pressure and topography. The duration contour spacing is 1 day while the contours for upward EP fluxes are on a logarithmic scale.

The “deep pulses” are found from this set of vertically-coherent pulses, again defined as pulses which are vertically-coherent between a given starting level and at least 100 hPa. In doing so we obtain a set of pulses which is a good proxy for pulses with deep vertical propagation, though we note the caveat that these do not strictly represent vertically-propagating waves. This set of deep pulses is divided into those which have increasing durations with height up to 100 hPa and those pulses which have decreasing durations with height up to 100 hPa. As discussed in Section 3.2, the former are a suitable proxy for instances of strong wave-mean flow positive feedback.

Figure 4.10 shows the total number of deep pulses as a function of topography and of starting level in the left panel. The total number of deep pulses unsurprisingly increases as the starting level approaches 100 hPa. More interestingly, the maximum occurrence of deep pulses for any given starting level falls at the smallest topographic heights (< 2000 m). These occurrences show a nearly monotonic decrease with increasing topography, perhaps in part due to the aforementioned vertical propagation inhibition by the mean state which would prevent pulses from reaching 100 hPa.

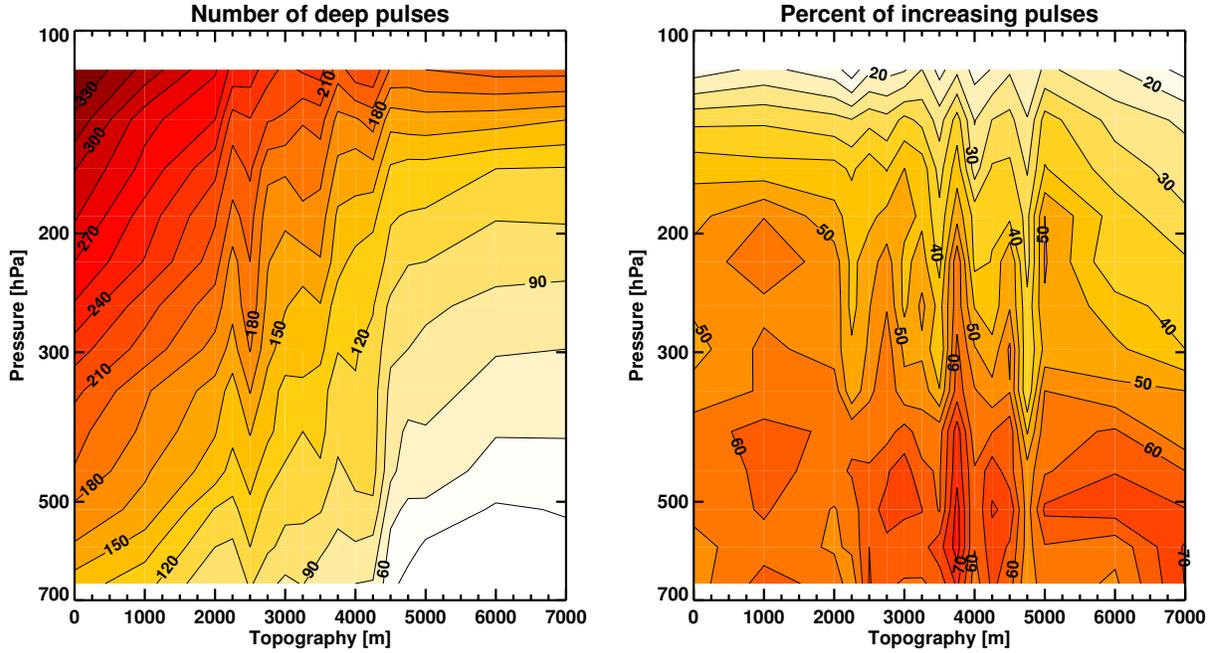


FIGURE 4.10. Total number of deep pulses as a function of topography and starting level (left) and the associated percentage of these deep pulses which are increasing with height (right). The contour spacing of the total deep pulses is 15, while the percentage contour spacing is 10%.

and above are small, resulting in small variability of the zonal winds (e.g. few or no SSWs). For larger topography, the time mean zonal winds decrease at these levels, showing that the frequency of fast zonal winds and, thus, strong positive feedbacks both decrease.

The right panel of Figure 4.10 shows the percent of these deep pulses which are increasing pulses at each topographic height and starting level. The vertical structure is consistent with that of Figure 3.12 in that most of the deep pulses with lower and middle tropospheric starting levels are increasing, with percentages only falling off near to 100 hPa. Though there is some variance with different topography, the changes in these percentage values are small.

Pulse characteristics (duration and amplitudes) of these increasing pulses are not remarkably different in structure and magnitude from those shown in Figure 4.9 and are thus not

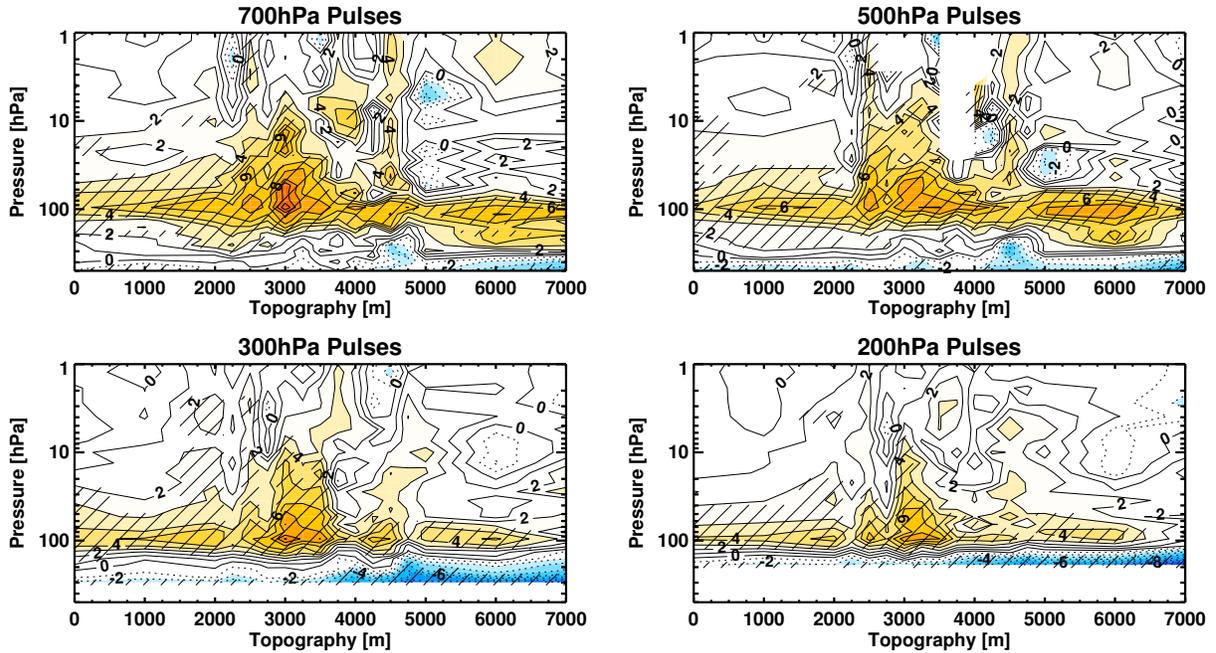


FIGURE 4.11. Differences in durations between increasing pulses and decreasing pulses, where positive values indicate that increasing pulses are longer. Results for 4 starting levels are shown: 700 hPa (upper left), 500 hPa (upper right), 300 hPa (lower left), and 200 hPa (lower right). Significant differences at the 95% level are hatched. Contour spacing is 1 day.

shown. Some figures showing these characteristics are included in the Appendix. We instead focus on differences between increasing and decreasing pulses here.

Figure 4.11 shows the differences in durations between increasing pulses and decreasing pulses for four starting levels. Positive values indicate where the increasing pulses are longer than the decreasing pulses. The increasing pulses are significantly longer at 100 hPa for nearly all starting levels and topographic heights. Further, the increasing pulses are significantly longer up to 10 hPa at least around 3000 m topography for all starting levels. The significance up to 10 hPa is over a larger range of topography for lower starting levels.

Figure 4.12 shows the analogous differences in amplitudes between the increasing pulses and decreasing pulses. The consistent features across starting levels are that the increasing

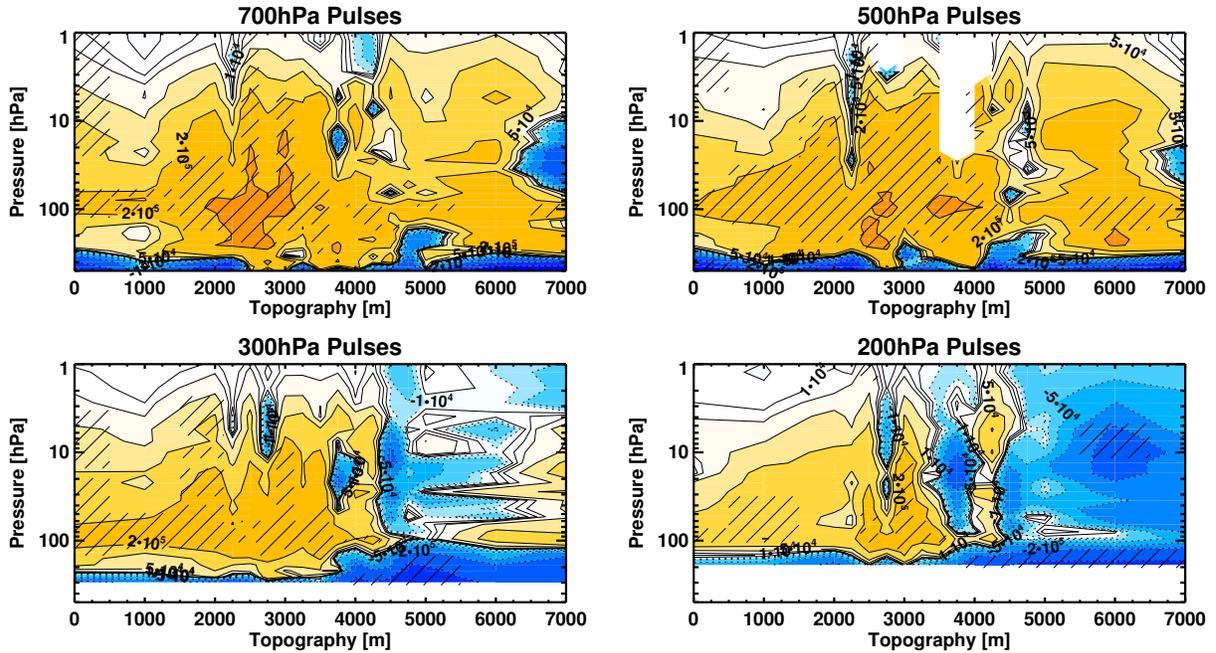


FIGURE 4.12. Differences in upward EP flux amplitudes between increasing pulses and decreasing pulses, where positive values indicate that increasing pulses are larger. Results for 4 starting levels are shown: 700 hPa (upper left), 500 hPa (upper right), 300 hPa (lower left), and 200 hPa (lower right). Significant differences at the 95% level are hatched. The contour spacing is on a logarithmic scale.

pulses have larger amplitudes for small to moderate topographic heights and that this difference is significant up to 10 hPa near 3000 m topography. Note that, as with the durations, the significance up to 10 hPa is over a larger range of topography for lower starting levels.

Finally, Figure 4.13 shows the differences in total upward EP flux between the increasing pulses and decreasing pulses. The total upward EP flux for increasing pulses is larger than that for decreasing pulses nearly everywhere above 100 hPa. Where this difference is significant is again centered around 3000 m topography, and broadens relative to topographic heights as the starting level is lowered.

From the above differences in pulse characteristics between increasing and decreasing pulses, the increasing pulses are again found to on average have longer durations and larger associated upward EP fluxes. Analysis of this behavior as a function of topography shows

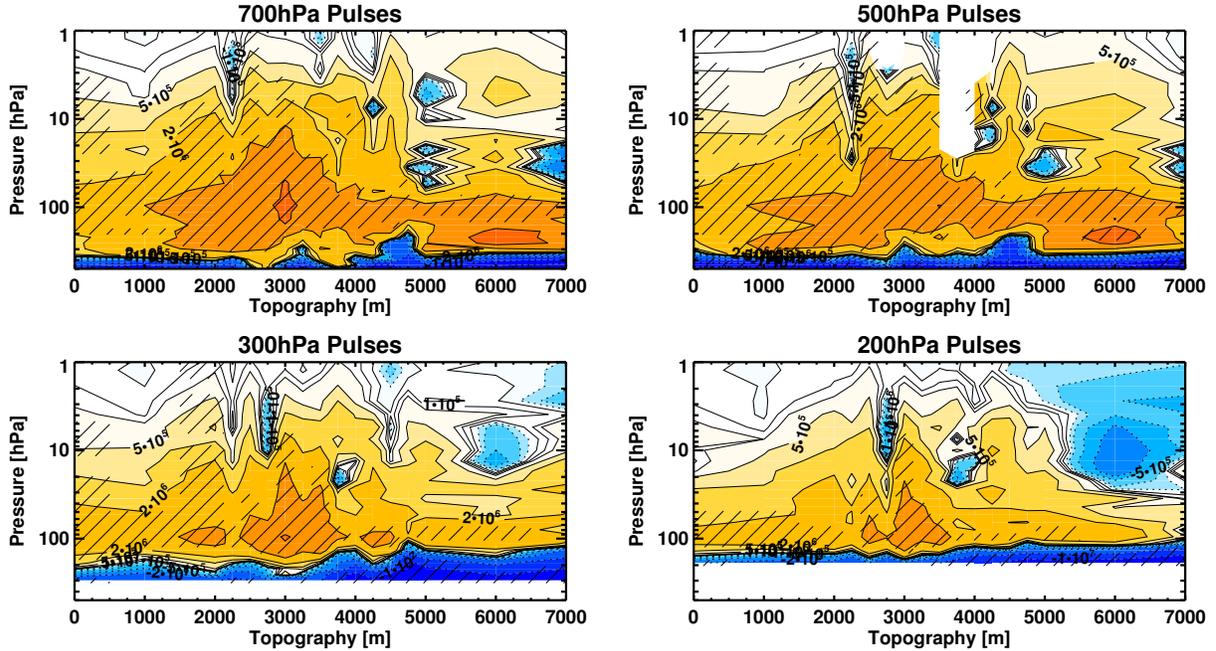


FIGURE 4.13. Differences in total upward EP flux between increasing pulses and decreasing pulses, where positive values indicate that increasing pulses are larger. Results for 4 starting levels are shown: 700 hPa (upper left), 500 hPa (upper right), 300 hPa (lower left), and 200 hPa (lower right). Significant differences at the 95% level are hatched. The contour spacing is on a logarithmic scale.

that the magnitudes of these differences vary as topography varies. Most notably, there is a range between approximately 2500 m and 4000 m topographic height in which the differences are usually larger and significant. Note that it is in this range of topographic heights that the most frequent occurrence of SSWs lies. This coincidence of strong zonal wind variability and strong increasing pulses suggests that the wave-mean flow positive feedback within this topographic parameter space is also quite strong.

To provide additional support for this, we analyze anomalous zonal wind states and evolution around these increasing pulses and decreasing pulses. Here the anomalies are relative to the time mean zonal wind at each pressure level. Figure 4.14 shows the means of the anomalous zonal winds for the 5 days prior to the central date for increasing pulses with the typical four starting levels. The choice of 5 days prior to the central date is made

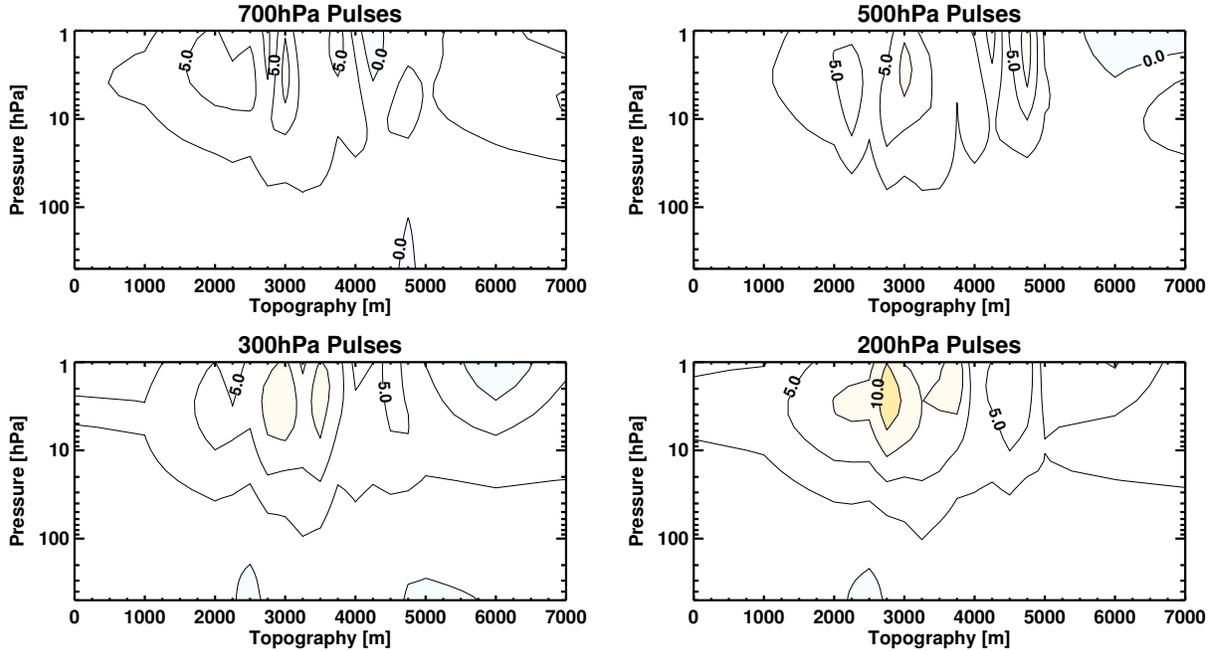


FIGURE 4.14. Means of the anomalous zonal winds for the 5 days prior to the central date for increasing pulses. The typical four starting levels are shown. Contour spacing is 2.5 ms^{-1} .

because it is within this period that the winds are strongest prior to increasing pulses (cf. Figure 3.16). Because these means do not show significance nor time evolution, we include some example plots showing this in the Appendix.

These results show that, on average, the zonal winds are anomalously fast prior to increasing pulses reaching 100 hPa, with the strongest winds between topographic heights of ~ 2500 m and 4000 m. However, as shown in Figure 4.15, the difference between zonal wind prior to increasing pulses and zonal wind prior to decreasing pulses are small in the GCM integrations. The increasing pulses have faster precedent zonal winds primarily for topographic heights between 3000 m and 5000 m, i.e. within the range of maximum SSW occurrence.

The final analysis of the zonal winds compares the changes in zonal winds from the central date for the following 20 days between increasing and decreasing pulses. Results of this are

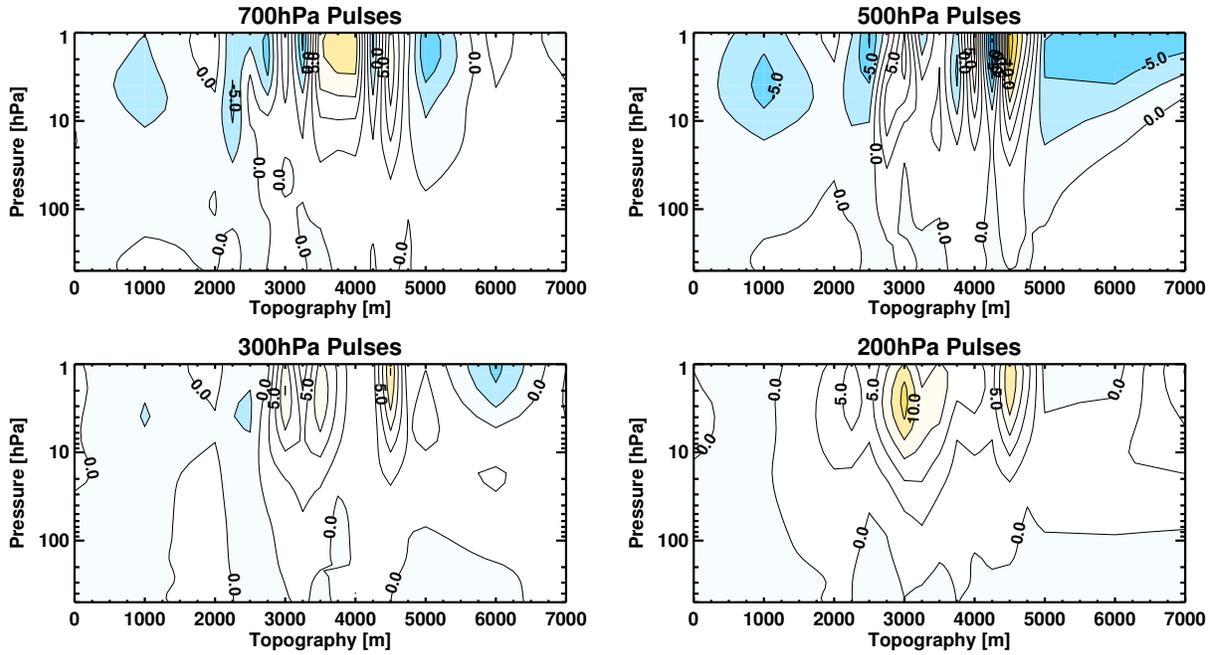


FIGURE 4.15. As in Figure 4.14, but showing the difference between increasing pulses and decreasing pulses.

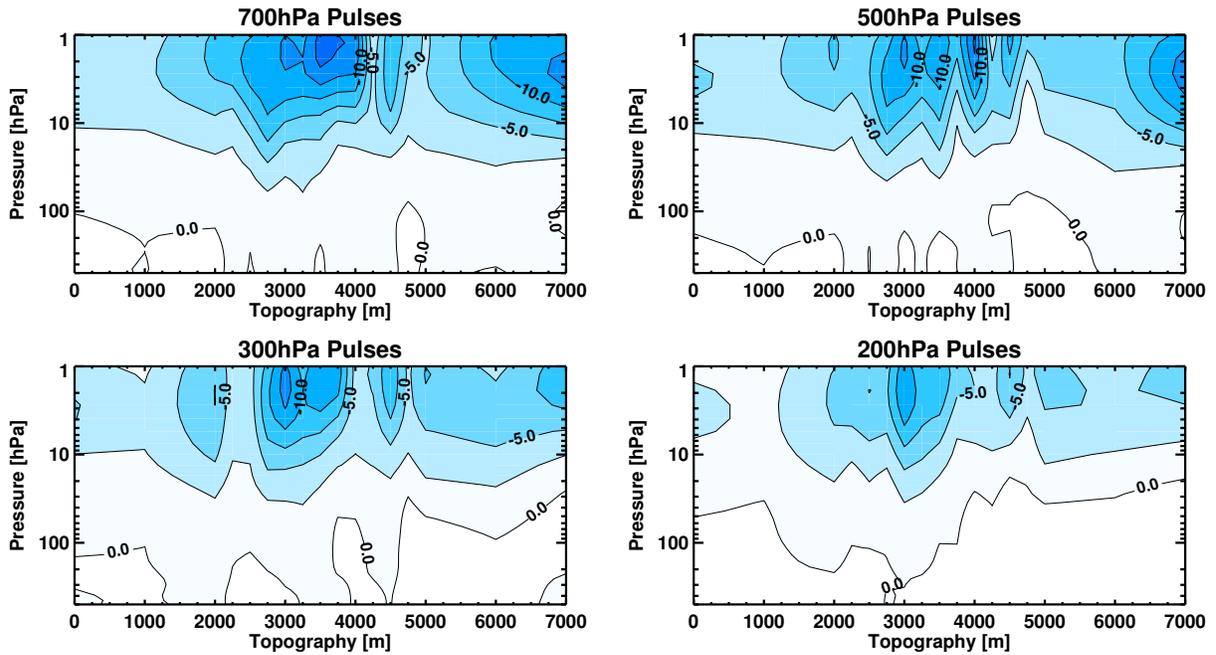


FIGURE 4.16. Differences between increasing pulses and decreasing pulses in the means of the anomalous zonal wind changes from the central date for the 20 days following the central date. The typical four starting levels are shown. Contour spacing is 2.5 ms^{-1} .

shown in Figure 4.16. Between 2500 m and 4000 m topography, the increasing pulses have considerably larger zonal wind decelerations at and above 10 hPa.

In all, the above pulse analysis demonstrates that strong wave-mean flow positive feedback events may differ in strength considerably dependent on the initial wave source (here, topography). The range of topographic heights corresponding to the range of most SSW occurrences coincides with a maximum in total wave fluxes for increasing pulses and a maximum in deceleration following these increasing pulses. This range of topographic heights – roughly 2500 m to 4000 m – appears to be the range in which the upward waves are neither too weakly nor too strongly forced, the former not being sufficient to sustain a strong feedback and the latter being too strong to allow the type of deep feedback cycle in which we are interested. As well, this analysis shows (see Figure 4.8) that the pulse durations in the GCM also align with the order 10 day preferential time scale, lending further supporting evidence for this result.

4.3. TROPOPAUSE INVERSION LAYER EXPERIMENTS

We have utilized the general circulation model with idealized topography to study how changes in tropopause inversion layer (TIL) stratification affects the upward EP fluxes in the stratosphere. From the simple modeling of Chapter 2, it was found that the time mean upward EP fluxes in the stratosphere are larger during integrations where the TIL stratification is larger. While these results were not upheld by the observational analysis of Chapter 3, the TIL in the ERAi dataset may not be properly represented. Thus the GCM presents a high-order environment in which to test the simple model results.

In order to do so, we adjust the radiative equilibrium temperature field (T_{eq}) in such a way as to increase or decrease the stratification at the model level at which the TIL lies.

Ideally, manipulation of the equilibrium temperature field would allow us to specify the stratification added at the TIL level with no other changes of stratification at any other level. This would provide the cleanest experiment by which to test the results from the simple model.

However, such an adjustment has been found to be harder to achieve than initially expected. While it is straightforward to produce an amplification of the TIL stratification, such amplifications also result in reduced stratification values above the TIL, the depth of which may be large relative to the depth of the TIL. In a way, these reductions act to greatly lengthen the vertical extent of the TIL. In experiments with the simple model which we do not show, it was found that stratospheric upward EP fluxes do not follow the relationship discussed in Section 2.4 when the TIL is deep (e.g. > 8 km). Instead, large TIL stratifications with deep TIL structures resulted in reduced time mean stratospheric upward EP fluxes.

The above is in line with the results from most TIL experiments we performed with the GCM: upward EP fluxes in the stratosphere are reduced as the TIL is strengthened. Rather than focusing on these results, however, we elect to present results of an experimental setup which does generally produce the expected enhancements of upward wave activity flux. The results discussed below are not wholly conclusive and are not intended to be broadly applicable to the real world. Rather, they give a proof of concept that the TIL stratification may indeed play a non-dissipative role in stratospheric wave amplitudes.

For this experiment, a pressure-dependent temperature perturbation (Teq_{add}) is added to the standard radiative equilibrium temperature (Teq_0). The form of this temperature

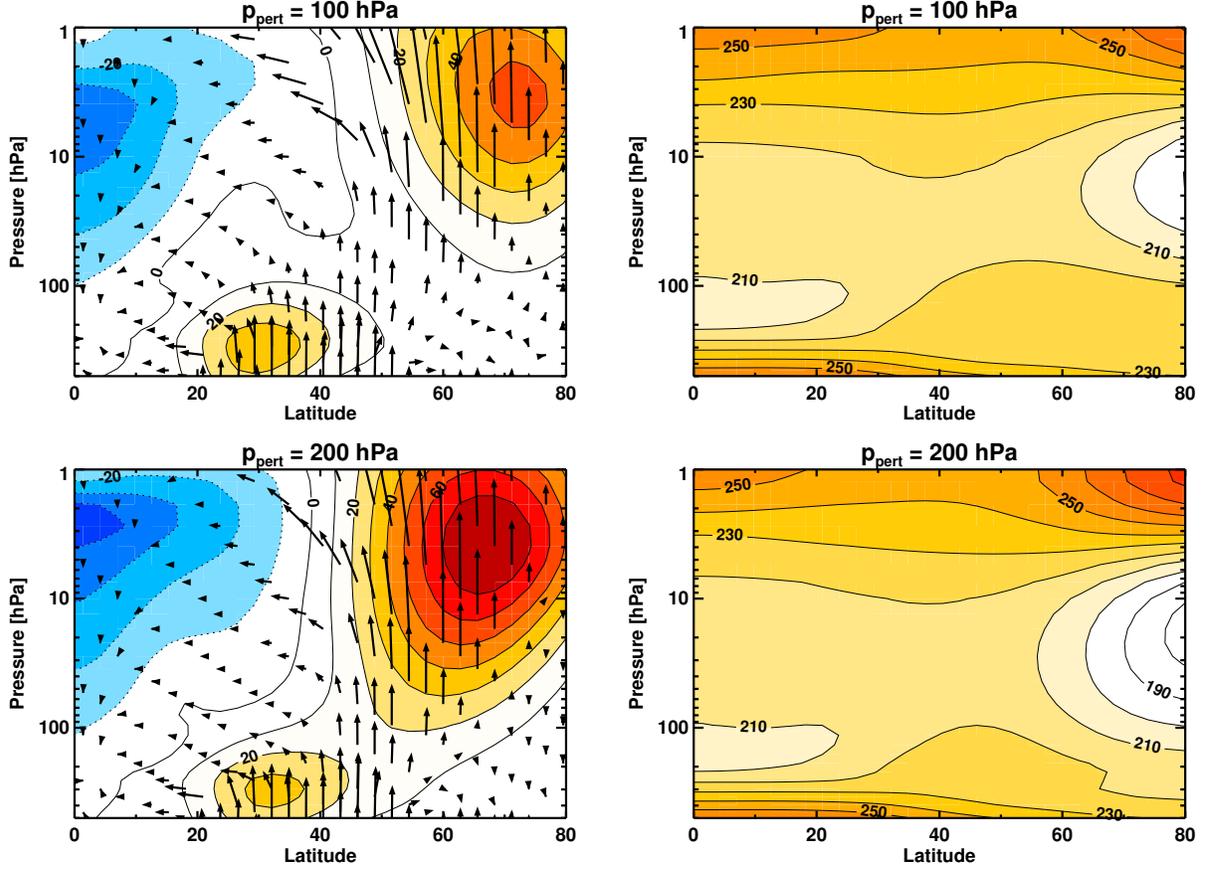


FIGURE 4.17. Zonal mean zonal winds (left) and temperatures (right) for integrations of the GCM with 3500 m topography. Overplotted on the zonal winds are EP flux vectors which are scaled by inverse density. The top row shows the time mean state for $p_{pert} = 100$ hPa while the bottom row shows the time mean state for $p_{pert} = 200$ hPa.

adjustment is

$$\begin{aligned}
 \text{Teq}_{add}(\phi, p) &= \lambda W(\phi) (p/p_{pert})^{\frac{13}{3}} && \text{for } p \leq p_{pert} \\
 \text{Teq}_{add}(\phi, p) &= 0 && \text{for } p > p_{pert},
 \end{aligned}
 \tag{4.4}$$

where λ is the magnitude of the temperature perturbation in K, W is the weighting function given by Polvani and Kushner (2002), and p_{pert} is the level at which the perturbation is first applied. For the experiments here, we set p_{pert} to 200 hPa and vary λ between -2 K and +20 K as follows: -2, 0, 2, 5, 10, 15, and 20 K.

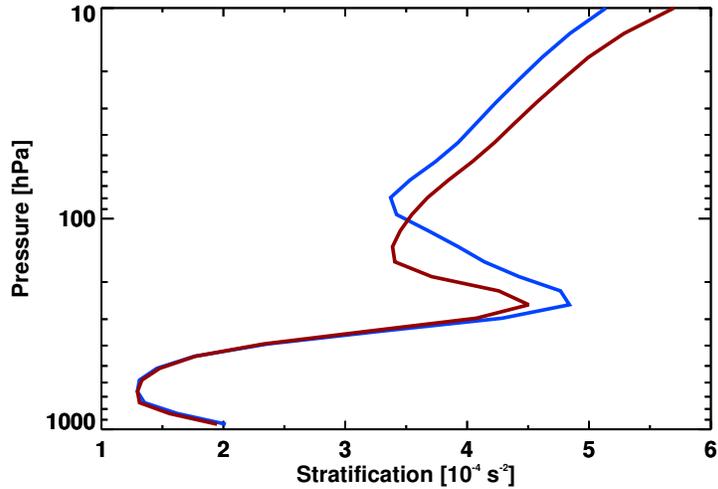


FIGURE 4.18. 45-75N mean stratification for the $p_{pert} = 100$ hPa integration (blue) and for the $p_{pert} = 200$ hPa integration (red).

Additionally, we adjust a parameter within the Polvani and Kushner (2002) Teq_0 : p_T , the height of a nominal tropopause. In their setup of Teq_0 , Polvani and Kushner essentially fit together a tropospheric Teq given by Held and Suarez (1994) and a stratospheric Teq which they derive to roughly match that of Fels (1985). The level at which these are fit together is set by p_T . In all experiments heretofore with this model setup, p_T is set to 100 hPa. For the experiments here, we instead set this parameter to be 200 hPa, equal to p_{pert} . Though this change produces a nontrivial change to the whole model background state, it also acts to greatly shorten the depth of the TIL, which is highly desirable here.

Figure 4.17 shows the time mean zonal winds and temperatures for two integrations with topography $h_0 = 3500$ m and temperature perturbation $\lambda = 0$ K. The top row shows the fields for $p_{pert} = 100$ hPa, while the bottom row is for $p_{pert} = 200$ hPa. Lowering the nominal tropopause cools the polar wintertime stratosphere. The colder stratosphere produces a stronger and deeper polar vortex, which in turn is an enhanced wave guide for upward propagating waves that allows these waves to propagate to large heights. These features all manifest in Figure 4.17.

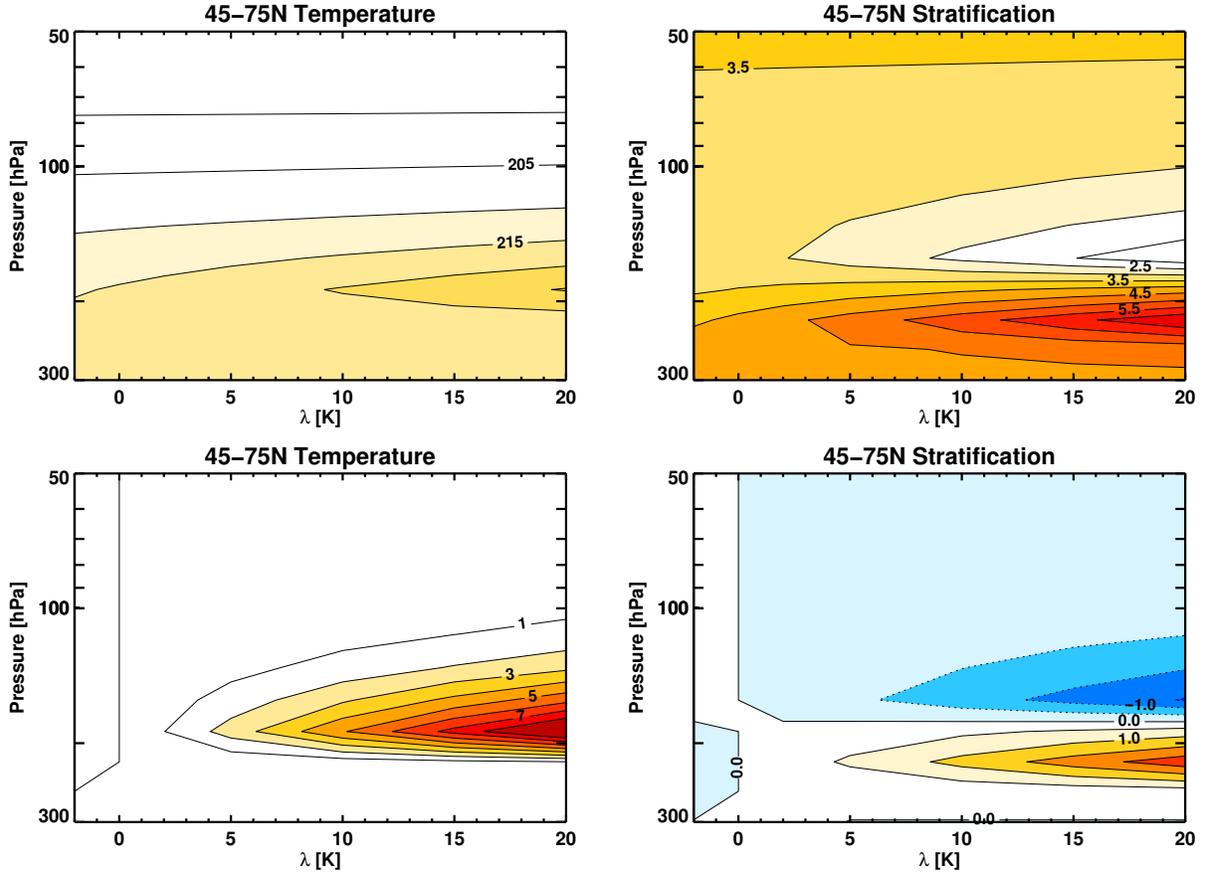


FIGURE 4.19. (Top) Radiative equilibrium temperature (left) and associated Brunt-Vaisala frequency squared (right) averaged between 45-75N as a function of λ . The values of stratification have been multiplied by 10^4 . The contour spacing for temperature is 5 K while the contour spacing for stratification is $0.5 \cdot 10^{-4} \text{ s}^{-2}$. (Bottom) The differences in radiative equilibrium temperature (left) and associated stratification (right) from the $\lambda = 0$ K integration. Here the temperature contour spacing is 1 K.

The main features of the zonal winds, temperatures, and EP flux vectors are still intact, showing that the time mean state of the altered setup (i.e. for $p_{pert} = 200$ hPa) is not so different as to be unjustifiably usable. Further, Figure 4.18 shows that the desired effect of this alteration is satisfied: the TIL depth decreases. While there is a concomitant decrease in the maximum stratification of the TIL, the experiments here are easily able to adjust this maximum.

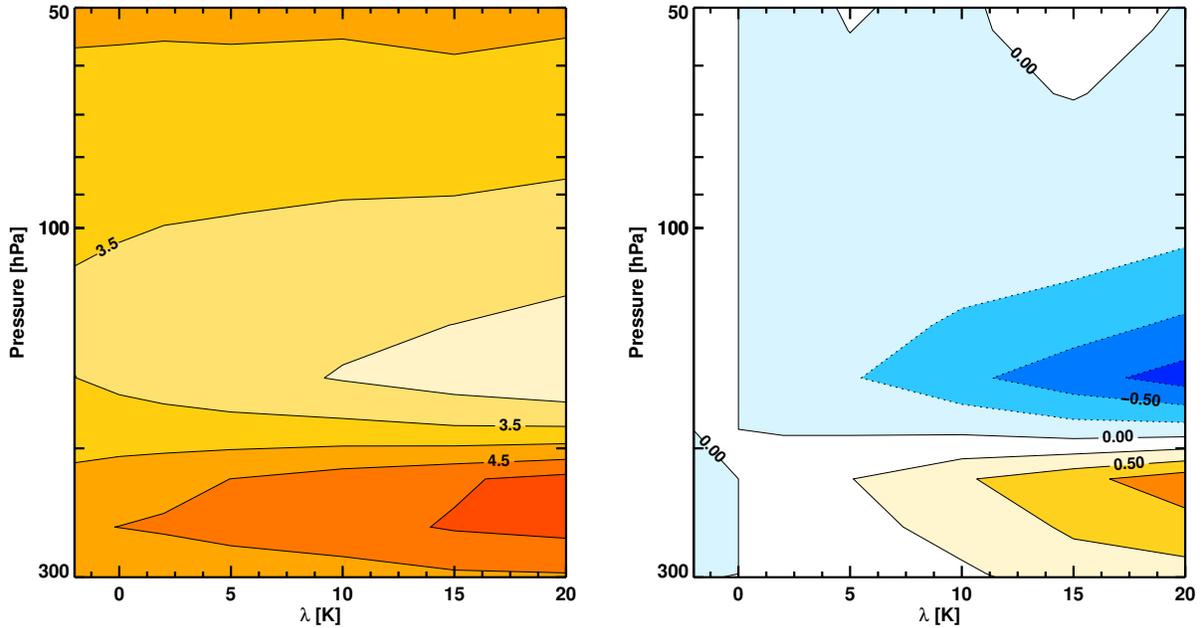


FIGURE 4.20. Time mean stratification averaged between 45-75N (left) and the differences in this stratification with the $\lambda = 0$ K integration (right), shown as a function of λ and of pressure. The contour spacing for the full stratification is $0.5 \cdot 10^{-4} \text{ s}^{-2}$ and is $0.25 \cdot 10^{-4} \text{ s}^{-1}$ for the difference in stratification.

The top row of Figure 4.19 shows the 500 hPa to 50 hPa radiative equilibrium temperatures and associated Brunt-Väisälä frequency squared averaged between 45-75°N. The bottom row of this figure shows the differences from the $\lambda = 0$ K integration for both fields. For the range of λ utilized here, T_{eq} only increases by 8 K at the most. Yet this increase and its vertical structure is sufficient to cause an increase in the radiative equilibrium stratification of up to $2 \cdot 10^{-4} \text{ s}^{-1}$. Above this increased stratification is a decrease in stratification of approximately half the magnitude. As noted above, this decreased stratification is not desirable for the experimentation here, but acceptable given the relatively shallow depth and small magnitude of the reduction.

In the model results, the correspondence between λ and TIL stratification is not as strong as in the radiative equilibrium calculations. Figure 4.20 shows these model stratification results and the differences from the $\lambda = 0$ K integration. The maximum change

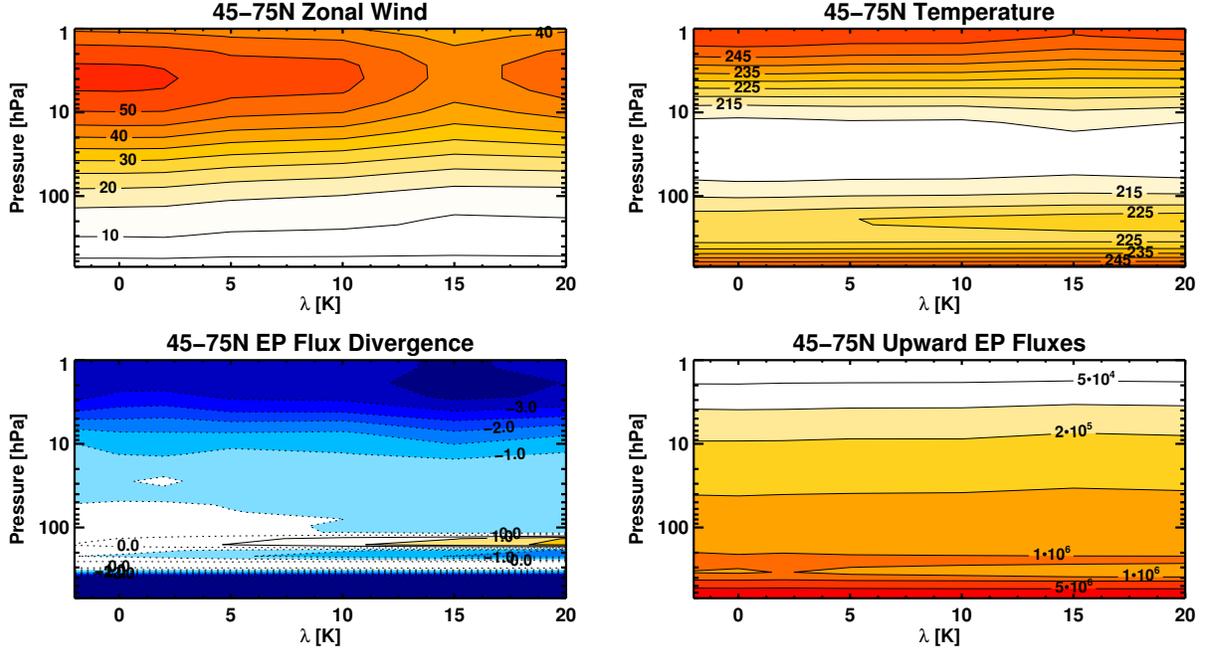


FIGURE 4.21. Time mean zonal winds (upper left), temperatures (upper right), EP flux divergences (lower left), and upward EP fluxes (lower right), all averaged between 45-75N and shown as function of λ and of pressure. The zonal wind contour spacing is 5 ms^{-1} , the temperature contour spacing is 5 K, the EP flux divergence contour spacing is $0.5 \text{ ms}^{-1} \text{ day}^{-1}$, and the contour spacing for the upward EP fluxes is on a logarithmic scale.

in TIL stratification is a factor of three smaller in the model results than in the radiative equilibrium calculations. The dynamics of the system (i.e. waves) may act to dampen the expected increase in stratification. Still, the applied equilibrium temperature adjustment Teq_{add} produces the desired changes in TIL stratification.

How the range of λ values – and thus the TIL stratification – affects the time mean state of the model is shown in Figure 4.21. The zonal winds (top left) decrease at all levels up to $\lambda = 15 \text{ K}$ and slightly increase above this, giving an initial indication that the wave amplitudes are increasing for increasing TIL stratification. This is bolstered by the relatively warm temperatures (top right) near 10 hPa for $\lambda = 15 \text{ K}$.

That the wave amplitudes increase with increasing TIL stratification is verified by the EP flux fields shown in the bottom row of Figure 4.21. There are nontrivial upward EP

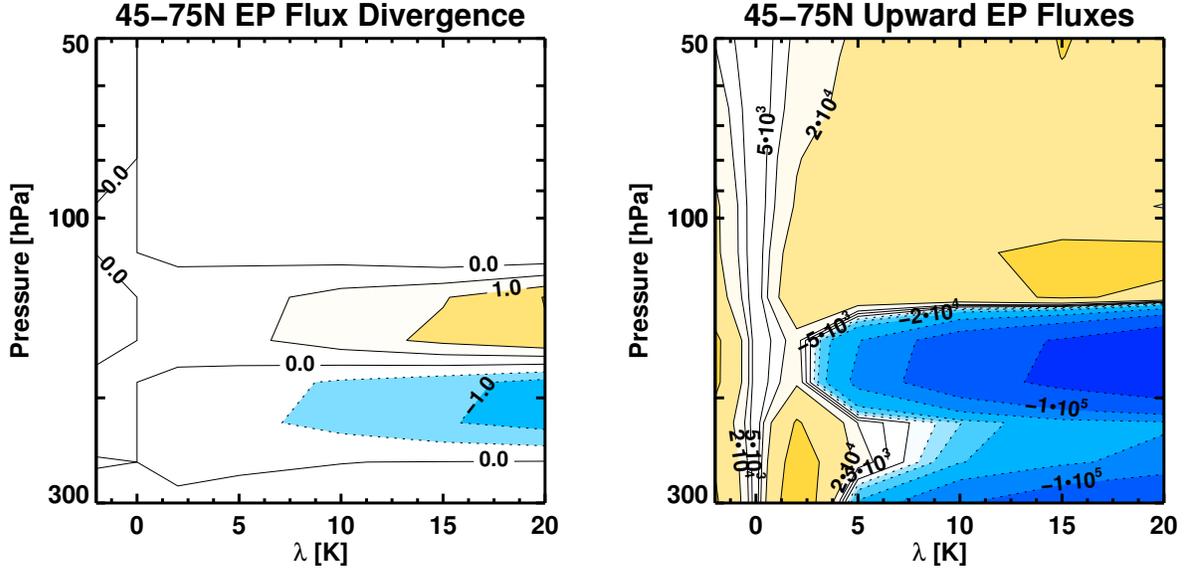


FIGURE 4.22. The difference from the $\lambda = 0$ K integration for the 45-75N EP flux divergences (left) and 45-75N upward EP fluxes (right).

flux divergences (wave generation) between 200 hPa and 100 hPa when the perturbation λ is larger than 5 K. As the divergence field shows that the primary sink of this wave activity flux is above 10 hPa, the wave amplitudes are expected to be larger in the lower and middle stratosphere for large λ . The upward EP flux field (bottom right) shows this to be true.

Figure 4.22 shows the anomalies of EP flux divergence and upward EP fluxes from the control case ($\lambda = 0$ K). This clearly shows the EP flux divergences above 200 hPa and that these divergences directly relate to large stratosphere-level upward EP fluxes. This wave generation is likely not the only source of the upward wave activity fluxes, however. Much of this could come from meridional wave propagation into the domain (45-75°N).

To assess this possible impact, Figure 4.23 shows the zonal winds and anomalous EP flux vectors for four values of λ . The anomalies are calculated as the differences between each integration and the $\lambda = 0$ K integration. In the upper troposphere and lower stratosphere, meridional EP fluxes across 45°N are anomalously positive (poleward) for nonzero λ and become increasingly so for larger λ . These increases in anomalous poleward EP fluxes

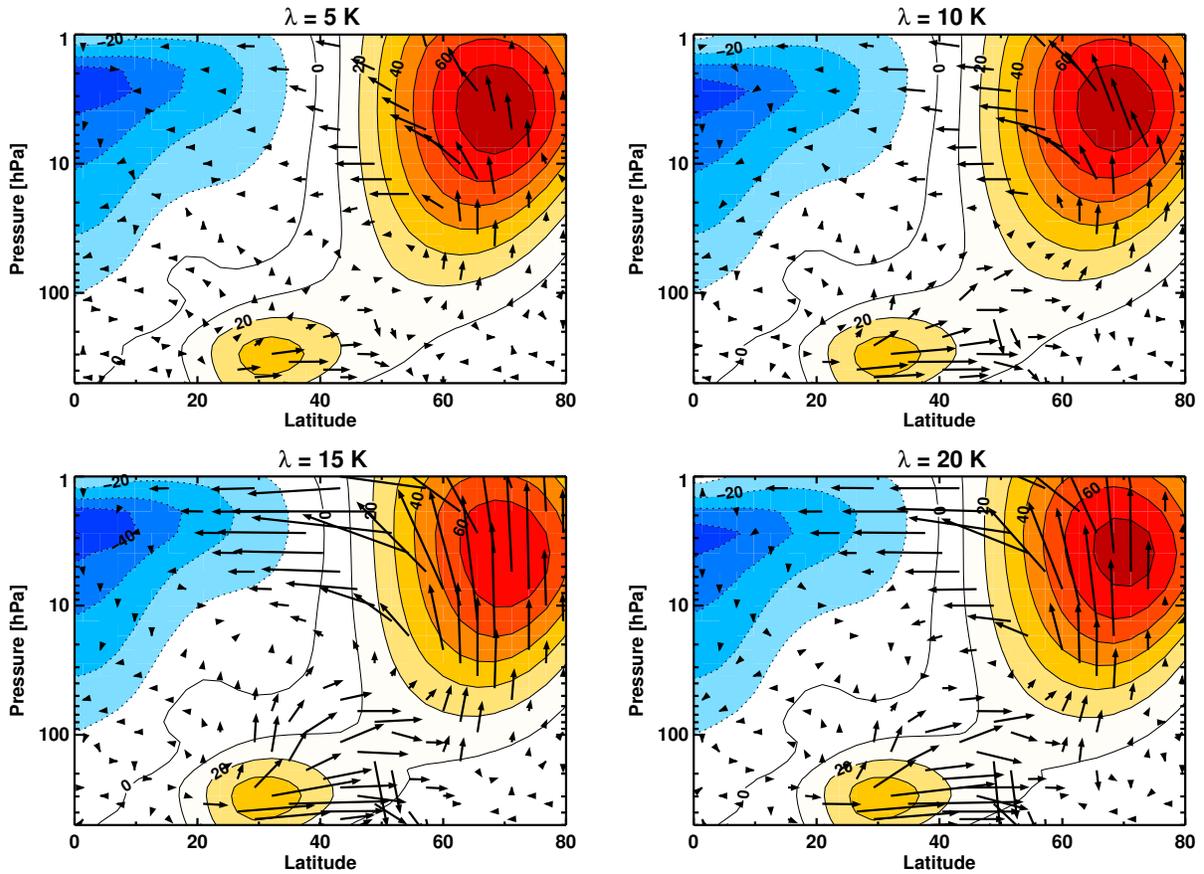


FIGURE 4.23. Zonal winds and overplotted anomalous EP flux vectors for four values of λ : 5 K (upper left), 10 K (upper right), 15 K (lower left), and 20 K (lower right). The EP flux anomalies are calculated as the difference between the given integration and the $\lambda = 0$ K integration. The vertical component of the EP flux is scaled by inverse pressure.

likely constitute some of the EP flux divergence and resultant enhanced middle and upper stratospheric upward EP fluxes.

An additional source of wave generation was shown by the simple model results. These results pointed to reversals of the meridional potential vorticity gradient as the likely culprit. Such reversals satisfy the Charney-Stern necessary condition for baroclinic instability and thus would indicate wave generation. Figure 4.24 shows the percentage of time for which the meridional potential vorticity gradient is reversed. At the largest values of λ (i.e. the

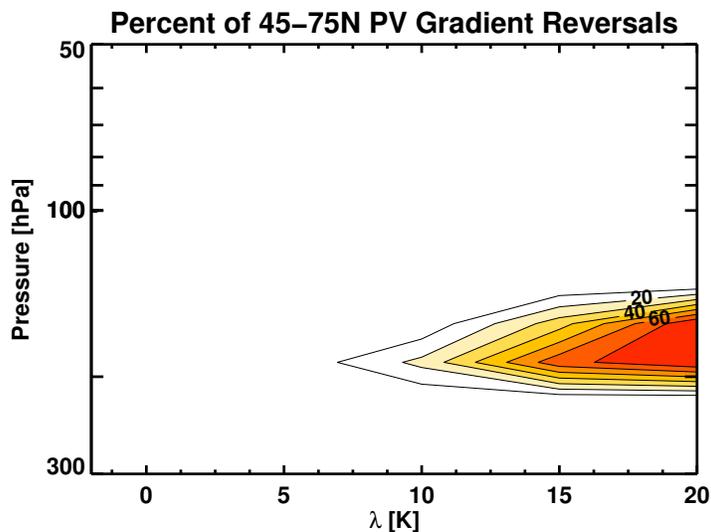


FIGURE 4.24. Percent of time for which the meridional potential vorticity gradient is reversed (< 0), shown as a function of λ and of pressure. The contour spacing is 10%.

largest TIL stratifications), the instability condition is satisfied for greater than 70% of the integration.

Given the frequency of reversed meridional potential vorticity gradient for large λ , this baroclinic instability must provide a nontrivial source of wave activity flux. Together, with the anomalous poleward EP fluxes in the upper troposphere and lower stratosphere, this results in enhanced wave activity fluxes in the middle and upper stratosphere. One effect of these altered wave activity fluxes is shown in Figure 4.25. Though the changes are not monotonic, the number of SSWs increases for increasing λ , suggesting that this added wave activity flux strongly affects the zonal wind variability.

To summarize, we adjust the radiative equilibrium temperature near the TIL level, forcing an increase in the maximum stratification there. The results above show that large maximum TIL stratification leads to relatively larger stratospheric wave activity fluxes. This added

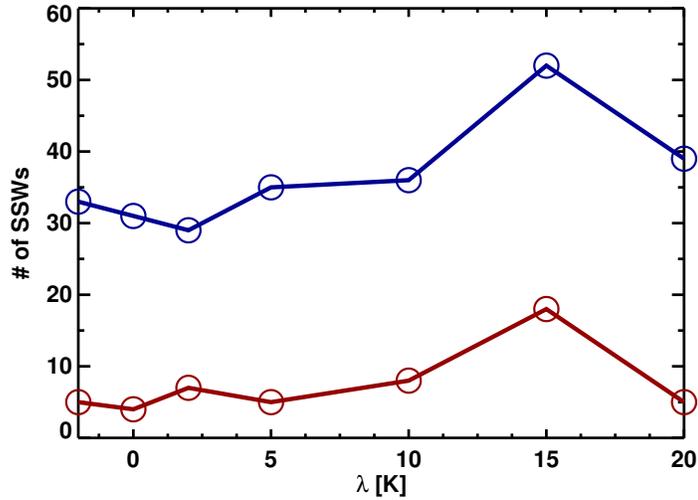


FIGURE 4.25. The total number of major SSWs (red) and of minor SSWs (blue) as a function of λ . Circles outline the actual data points.

wave activity flux into the lower stratosphere may be a nontrivial factor during strong wave-mean flow positive feedback events such as SSWs. It is still not clear whether this is a robust result that applies broadly to the atmosphere, however.

4.4. CONCLUSIONS

A general circulation model is utilized to investigate properties of the wave-mean flow positive feedback. The dry, dynamical core model is forced by idealized topography of zonal wavenumber 2 with specified height. This model setup has been utilized in previous studies, in part because its idealized nature allows for mechanistic studies on stratospheric dynamics and in part because it reproduces with fidelity known coupling between the stratosphere and troposphere.

Integrations are performed spanning a large range of topographic heights. In line with previous results, we find a nonlinear functional form of the occurrence of SSWs on topographic height. SSWs do not occur for small topographic heights and occur infrequently for large topographic heights; sudden warmings reach a maximum occurrence rate for moderate

topographic forcing. Analysis of the upward EP fluxes in these integrations reveals a similar nonlinear behavior. In the stratosphere, upward wave amplitudes reach a maximum at moderate topographic forcing and remain nearly constant for still larger topography. This latter feature is consistent with inhibited upward wave propagation shown by frequent negative values of the index of refraction in and below this region. Tropospheric wave amplitudes meanwhile strongly decrease for progressively larger topography up to moderate topographic forcing, where the amplitudes begin to slightly increase with topography. Decomposing the upward EP fluxes into zonal wavenumbers reveals that synoptic-scale waves are most responsible for the initial decrease at small topography, with wavenumber 2 providing the slight increase between moderate and large topographic heights. Though some of the reductions in synoptic scale wave amplitudes results from organization onto wavenumber 2, it is possible that weakened energy cascade at these levels may play a part too.

A pulse analysis of the topography experiments reveals differences in the positive feedback cycle across the different integrations. Pulse durations and amplitudes maximize in the region of topographic heights where the maximum in SSWs lies. Analyzing pulse characteristics and zonal wind for these increasing pulses shows that the strength of the positive feedback varies strongly with topographic height. At moderate topography, both the zonal wind and the wave amplitudes are within a parameter space such that the positive feedback is notably, and often significantly, strong. This finding is consistent with the high frequency of SSWs in this topographic height range. Our pulse analysis also shows that the typical time scale of pulses falls near to 10 days, supporting the prior results of a preferential wave forcing time scale.

Results are presented from an experiment in which the stratification of the tropopause inversion layer was adjusted through manipulation of the radiative equilibrium temperature.

The modeled changes in TIL stratification are roughly a factor of three smaller than those expected from the changes in radiative equilibrium, suggesting that waves may reduce the stratification by propagating energy out of this region. We find that the raised maximum stratification of the TIL results in frequent instability and corresponds to enhanced poleward wave propagation in the upper troposphere and lower stratosphere. These sources of lower stratospheric, upward-propagating waves may provide a nontrivial influence on wave-mean flow interactions which thus far has not been resolved. Additional study is needed here to draw out the applicability of these results to the real world.

CHAPTER 5

CONCLUSIONS

5.1. SUMMARY

A hierarchical approach was utilized to study wave-mean flow positive feedbacks associated with sudden stratospheric warmings (SSWs). These nonlinear feedbacks present a considerable challenge to the predictability of SSWs as small wave amplitudes may be significantly enhanced. Better determination of the wave-mean flow processes involved will lead to greater predictability not only through qualitative understanding but also through identification of unresolved physics which shape the evolution of these feedbacks.

The first component in the hierarchy is the relatively simple Holton-Mass model. This one-dimensional wave-mean flow interaction model has been shown to be remarkably applicable to the SSW problem and thus is a useful tool here. We present a reformulation of the bottom boundary condition which allows for specification of the upward wave activity fluxes there. Using experiments of idealized upward wave fluxes, it is found that bistability is a robust feature of this model while the widely known “vacillation cycles” are not. The latter is found to rely on a representation of the positive feedback which is likely too strong in the simple model.

Transient forcing experiments support previous results that there is a preferential wave forcing time scale on the the order of 10 days for forcing SSWs. Forcing the model at an internal wave time scale most efficiently drives the positive feedback. Finally, we find that an idealized tropopause inversion layer (TIL) stratification may result in wave generation in the lower stratosphere due to baroclinic instability.

Reanalysis data is studied for supporting evidence to the simple model results. Analysis of lower stratospheric heat fluxes and Eliassen-Palm (EP) fluxes reveals that the typical durations of waves prior to SSWs is on the order of 10 days. Though this time scale holds little predictive power, it confirms that there are order 10 day internal stratospheric wave time scales upon which incoming waves will most efficiently force the positive feedback.

Pulses of EP flux are analyzed, and a set of pulses which are a proxy for strong wave-mean flow positive feedback are found. This proxy set – in which the pulse durations increase with height – displays the expected characteristics of the positive feedback: strong precedent zonal winds and strong, deep subsequent wind decelerations. This proxy is thus a useful diagnostic for the wave-mean flow positive feedback.

A general circulation model (GCM) is employed to perform high order experiments. This GCM is a dry dynamical core integrating the spherical primitive equations. The model is forced by specified radiative damping to a specified radiative equilibrium and by idealized zonal topography of wavenumber 2. By stepwise increasing the height of the topography, we find that the frequency of SSWs within the model increases nonlinearly to a maximum and then strongly jumps down to a lower, steady value. While upward EP fluxes in the stratosphere are found to mimic this behavior, tropospheric upward EP fluxes are reduced for small to moderate topographic heights and remain reduced for large topography. This indicates 1) that large changes in wave-wave interactions occur in the troposphere across the range of topographic heights and 2) that the wave-mean flow positive feedback must be the result of increasing upward EP fluxes in the small to moderate topography range.

The latter point is shown through a pulse analysis similar to that undertaken for the reanalysis data. Characteristics of pulses with increasing duration with height change across the ranges of topography, showing that in the moderate topography range, the positive

feedback is strongest. This range comes about because neither the upward wave amplitudes nor the stratospheric zonal winds are too weak, as in the small and large topographic ranges, respectively. Additionally, these increasing pulses have duration on the order of 10 days, further showing the stratospheric internal wave time scale.

While supporting evidence of TIL-based wave generation was not found in the reanalysis data, a set of experiments with the GCM do show that the processes identified in the simple model are possible. By increasing the maximum stratification of the TIL, upward EP fluxes in the stratosphere increased in part due to lower stratospheric wave generation by baroclinic instability.

Overall, this dissertation contains four principal findings. Firstly, constant incoming wave activity flux leads to a constant zonal wind response. Secondly, the stratosphere has an internal wave forcing time scale. Wave forcing matching or near to this time scale most strongly activates the positive feedback cycle, resulting in large zonal wind decelerations and SSWs. Thirdly, increasing pulse duration with height is a useful diagnostic for analyzing characteristics of stratospheric wave-mean flow positive feedbacks. Finally, the TIL may provide a nontrivial source of wave amplitudes in the lower stratosphere which may then interact with the positive feedbacks higher up.

5.2. OUTLOOK

A number of avenues of additional study in all components of the modeling hierarchy are available. In the case of the simple model, the reformulated bottom boundary is an ideal way to study preconditioning of the polar vortex. Rather than considering a transient pulse in isolation, two pulses, separated by some distance could be considered. When the pulses are not separated by long amounts of time, we would expect the positive feedback to be able

to amplify the zonal wind effect of the second pulse. Then a question of interest here is, how separated must these pulses be before the second pulse acts independently? Further, what more strongly controls this time of separation: the background state or the radiative damping?

For the reanalysis study, future work should continue towards understanding how the vertically-coherent pulses increase in duration between the upper troposphere and lower stratosphere. Postulation on the role of high wavenumber filtering was given earlier, but this remains unfounded without further consideration. Understanding would likely be bolstered greatly from analyses of individual vertically-coherent pulse events. Of particular interest here would be those pulses which are associated with (i.e. precede) SSWs.

An issue raised in analyzing these vertically-coherent pulses was on the influence of lower stratospheric radiative damping. Though we argued that this effect would be small, it would be beneficial to understand the relative magnitude. For this, the GCM would be an appropriate tool to apply. Changes in lower stratospheric radiative damping could be included in integrations with set topography. How the statistics of pulses in this region change would better illustrate the impact from radiative damping.

Additional future work with the GCM would be on analyzing the changes in synoptic scale wave amplitudes in the troposphere. The reasons for this were not well explored here, but may be important for understanding of stratosphere-troposphere coupling. It is possible that these amplitudes are reduced by increasing frequency of anomalous downward wave events as topography is increased. If this is so, it presents evidence in support of previous studies that stratospheric reflections of waves are a pathway for stratosphere-troposphere coupling.

REFERENCES

- Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics*. Academic Press, 489 pp pp.
- Andrews, D. G. and M. E. McIntyre, 1976: Planetary waves in horizontal and vertical shear: the generalized Eliassen-Palm relation and the mean zonal acceleration. *J. Atmos. Sci.*, **33**, 2031–2048.
- Baldwin, M. P. and T. J. Dunkerton, 1989: The stratospheric major warming of early December 1987. *J. Atmos. Sci.*, **46**, 2863–2884.
- Baldwin, M. P., D. B. Stephenson, D. W. J. Thompson, T. J. Dunkerton, A. J. Charlton, and A. O’Neill, 2003: Stratospheric memory and skill of extended-range weather forecasts. *Science*, **301**, 636–640.
- Birner, T., 2006: Fine-scale structure of the extratropical tropopause region. *J. Geophys. Res.*, **111**, D04104, doi:10.1029/2005JD006301.
- Birner, T., D. Sankey, and T. G. Shepherd, 2006: The tropopause inversion layer in models and analyses. *Geophys. Res. Lett.*, **33**, L14804, doi:10.1029/2006GL026549.
- Birner, T. and P. D. Williams, 2008: Sudden stratospheric warmings as noise-induced transitions. *J. Atmos. Sci.*, **65**, 3337–3343.
- Boville, B. A., 1984: The influence of the polar night jet on the tropospheric circulation in a GCM. *J. Atmos. Sci.*, **41**, 1132–1142.
- Butler, A. H., D. J. Seidel, S. C. Hardiman, N. Butchart, T. Birner, and A. Match, Submitted: Defining sudden stratospheric warmings. *Bull. Am. Meteorol. Soc.*
- Celani, A., S. Musacchio, and D. Vincenzi, 2010: Turbulence in more than two and less than three dimensions. *Phys. Rev. Lett.*, **104**, 184506–1.

- Chao, W. C., 1985: Sudden stratospheric warmings as catastrophes. *J. Atmos. Sci.*, **42**, 1631–1646.
- Chapman, W. A. and T. Miles, 1981: Planetary-scale wave guides in the troposphere and stratosphere. *Nature*, **293**, 108–112.
- Charlton, A. J., A. O’Neill, W. A. Lahoz, and P. Berrisford, 2005: The splitting of the stratospheric polar vortex in the Southern Hemisphere, September 2002: dynamical evolution. *J. Atmos. Sci.*, **62**, 590–602.
- Charlton, A. J. and L. M. Polvani, 2007: A new look at sudden stratospheric warmings. Part I: climatology and modeling benchmarks. *J. Clim.*, **20**, 449–469.
- Charney, J. G. and P. G. Drazin, 1961: Propagation of planetary-scale disturbances from the lower to the upper atmosphere. *J. Geophys. Res.*, **66**, 83–109.
- Charney, J. G. and M. E. Stern, 1962: On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.*, **19**, 159–172.
- Chen, P. and W. A. Robinson, 1992: Propagation of planetary waves between the troposphere and stratosphere. *J. Atmos. Sci.*, **49**, 2533–2545.
- Coy, L., E. R. Nash, and P. A. Newman, 1997: Meteorology of the polar vortex: spring 1997. *Geophys. Res. Lett.*, **24**, 2693–2696.
- Dickinson, R. E., 1968: Planetary Rossby waves propagating vertically through weak westerly wind wave guides. *J. Atmos. Sci.*, **25**, 984–1002.
- Dickinson, R. E., 1969: Theory of planetary wave-zonal flow interaction. *J. Atmos. Sci.*, **26**, 73–81.
- Edmon, H. J., B. J. Hoskins, and M. E. McIntyre, 1980: Eliassen-Palm cross sections of the troposphere. *J. Atmos. Sci.*, **37**, 2600–2616.

- Esler, J. G., L. M. Polvani, and R. K. Scott, 2006: The Antarctic stratospheric sudden warming of 2002: a self-tuned resonance? *Geophys. Res. Lett.*, **33**, L12 804, doi:10.1029/2006GL026034.
- Fels, S., 1985: Radiative-dynamical interactions in the middle atmosphere. *Advances in Geophysics*, **28A**, 277–300.
- Finger, F. G., H. M. Woolf, and C. E. Anderson, 1966: Synoptic analyses of the 5-, 2-, and 0.4-millibar surfaces for the IQSY period. *Mon. Weather Rev.*, **94**, 651–661.
- Garcia, R. R. and D. L. Hartmann, 1980: The role of planetary waves in the maintenance of the zonally averaged ozone distribution of the upper stratosphere. *J. Atmos. Sci.*, **37**, 2248–2264.
- Geisler, J. E., 1974: A numerical model of the sudden stratospheric warming mechanism. *J. Geophys. Res.*, **79**, 4989–4999.
- Geisler, J. E. and R. E. Dickinson, 1974: Numerical study of an interacting rossby wave and barotropic zonal flow. *J. Atmos. Sci.*, **31**, 946–955.
- Gerber, E. P., 2012: Stratospheric versus tropospheric control of the strength and structure of the Brewer-Dobson circulation. *J. Atmos. Sci.*, **69**, 2857–2877.
- Gerber, E. P. and L. M. Polvani, 2009: Stratosphere-troposphere coupling in a relatively simple AGCM: the importance of stratospheric variability. *J. Clim.*, **22**, 1920–1933.
- Hardiman, S. C. and P. H. Haynes, 2008: Dynamical sensitivity of the stratospheric circulation and downward influence of upper level perturbations. *J. Geophys. Res.*, **113**, D23 103, doi:10.1029/2008JD010168.
- Hare, F. K., 1960: The disturbed circulation of the arctic stratosphere. *J. Meteor.*, **17**, 36–51.
- Harnik, N., 2009: Observed stratospheric downward reflection and its relation to upward pulses of wave activity. *J. Geophys. Res.*, **114**, D08 120, doi:10.1029/2008JD010493.

- Harnik, N. and R. S. Lindzen, 2001: The effect of reflecting surfaces on the vertical structure and variability of stratospheric planetary waves. *J. Atmos. Sci.*, **58**, 2872–2894.
- Held, I. M., 2005: The gap between simulation and understanding in climate modeling. *Bull. Amer. Meteor. Soc.*, **86**, 1609–1614.
- Held, I. M. and M. J. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Am. Meteorol. Soc.*, **75**, 1825–1830.
- Hirooka, T., T. Ichimaru, and H. Mukougawa, 2007: Predictability of stratospheric sudden warmings as inferred from ensemble forecast data: intercomparison of 2001/02 and 2003/04 winters. *J. Meteor. Soc. Japan*, **85**, 919–925.
- 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.
- Holton, J. R. and C. Mass, 1976: Stratospheric vacillation cycles. *J. Atmos. Sci.*, **33**, 2218–2225.
- Hoskins, B. J., 1983: Dynamical processes in the atmosphere and the use of models. *Q. J. R. Meteorol. Soc.*, **109**, 1–21.
- Jucker, M., S. Feuglistaler, and G. K. Vallis, Submitted: Sudden stratospheric warmings in and idealized GCM. *J. Geophys. Res.*
- Julian, P. R. and K. B. Labitzke, 1965: A study of atmospheric energetics during the January–February 1963 stratospheric warming. *J. Atmos. Sci.*, **22**, 597–610.
- Kuroda, Y., 2002: Relationship between the polar-night jet oscillation and the annular mode. *Geophys. Res. Lett.*, **29**, 1240, doi:10.1029/2001GL013933.
- Limpasuvan, V., D. W. J. Thompson, and D. L. Hartmann, 2004: The life cycle of the Northern Hemisphere sudden stratospheric warmings. *J. Clim.*, **17**, 2584–2596.

- Mahlman, J. D., 1969: Heat balance and mean meridional circulations in the polar stratosphere during the sudden warming of January 1958. *Mon. Weather Rev.*, **97**, 534–540.
- Martius, O., L. M. Polvani, and H. C. Davies, 2009: Blocking precursors to stratospheric sudden warming events. *Geophys. Res. Lett.*, **36**, L14 806, doi:10.1029/2009GL038776.
- Matsuno, T., 1970: Vertical propagation of stationary planetary waves in the winter Northern Hemisphere. *J. Atmos. Sci.*, **27**, 871–883.
- Matsuno, T., 1971: A dynamical model of the sudden stratospheric warming. *J. Atmos. Sci.*, **28**, 1479–1494.
- McInturff, R. M., 1978: *Stratospheric warmings: Synoptic, dynamic and general-circulation aspects*, Vol. 1017. NASA Ref. Publ., 50 pp. pp.
- McIntyre, M. E., 1982: How well do we understand the dynamics of stratospheric warmings? *J. Meteor. Soc. Japan*, **60**, 37–65.
- McIntyre, M. E. and T. N. Palmer, 1983: Breaking planetary waves in the stratosphere. *Nature*, **305**, 593–600.
- Miyakoda, K., R. F. Strickler, and G. D. Hembree, 1970: Numerical simulation of the breakdown of a polar-night vortex in the stratosphere. *J. Atmos. Sci.*, **27**, 139–154.
- Mukougawa, H. and T. Hirooka, 2004: Predictability of stratospheric sudden warming: a case study for 1998/99 winter. *Mon. Wea. Rev.*, **132**, 1764–1776.
- Newman, P. A. and E. R. Nash, 2000: Quantifying the wave driving of the stratosphere. *J. Geophys. Res.*, **105**, 12 485–12 497.
- Newman, P. A., E. R. Nash, and E. R. Rosenfield, 2001: What controls the temperature of the Arctic stratosphere during the spring? *J. Geophys. Res.*, **106**, 19 999–20 010.
- Perlwitz, J. and N. Harnik, 2003: Observational evidence of a stratospheric influence on the troposphere by planetary wave reflection. *J. Clim.*, **16**, 3011–3026.

- Plumb, R. A., 1981: Instability of the disturbed polar night vortex: a theory of stratospheric warmings. *J. Atmos. Sci.*, **38**, 2514–2531.
- Plumb, R. A., 1989: On the seasonal cycle of stratospheric planetary waves. *Pure Appl. Geophys.*, **130**, 233–242.
- Plumb, R. A., 2010: Planetary waves and the extratropical winter stratosphere. *The Stratosphere: Dynamics, Transport, and Chemistry*, L. M. Polvani, A. H. Sobel, and D. W. Waugh, Eds., AGU, Geophysical Monograph Series, Vol. 190.
- Polvani, L. M. and P. J. Kushner, 2002: Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. *Geophys. Res. Lett.*, **29**, 1114, doi:10.1029/2001GL014284.
- Polvani, L. M. and D. W. Waugh, 2004: Upward wave activity flux as a precursor to extreme stratospheric events and subsequent anomalous surface weather regimes. *J. Clim.*, **17**, 3548–3554.
- Ruzmaikin, A., J. Lawrence, and C. Cadavid, 2003: A simple model of stratospheric dynamics including solar variability. *J. Clim.*, **16**, 1593–1600.
- Scherhag, R., 1952: Die explosionsartige Stratosphärenenerwärmung des Spätwinters 1951/1952. *Ber. Deut. Wetterdienstes*, **6**, 51–63.
- Schoeberl, M. R., 1978: Stratospheric warmings: observations and theory. *Rev. Geophys. Space Phys.*, **16**, 521–538.
- Scott, R. K. and L. M. Polvani, 2004: Stratospheric control of upward wave flux near the tropopause. *Geophys. Res. Lett.*, **31**, L02 115, doi:10.1029/2003GL017965.
- Simpson, I. R., M. Blackburn, and J. D. Haigh, 2009: The role of eddies in driving the tropospheric response to stratospheric heating perturbations. *J. Atmos. Sci.*, **66**, 1347–1365.

- Sjoberg, J. P., 2010: Low-order models of sudden stratospheric warmings. M.S. thesis, Dept. of Atmospheric Science, Colorado State University, 89 pp.
- Sjoberg, J. P. and T. Birner, 2012: Transient tropospheric forcing of sudden stratospheric warmings. *J. Atmos. Sci.*, **69**, 3420–3432.
- Sjoberg, J. P. and T. Birner, In Press: Stratospheric wave-mean flow feedbacks and sudden stratospheric warmings in a simple model forced by upward wave-activity flux. *J. Atmos. Sci.*
- Teweles, S. and F. G. Finger, 1958: An abrupt change in stratospheric circulation beginning in mid-January 1958. *Mon. Wea. Rev.*, **86**, 23–28.
- Thompson, D. W. J., M. P. Baldwin, and J. M. Wallace, 2002: Stratospheric connection to Northern Hemisphere wintertime weather: implications for prediction. *J. Clim.*, **15**, 1421–1428.
- Waugh, D. W., W. J. Randel, S. Pawson, P. A. Newman, and E. R. Nash, 1999: Persistence of the lower stratospheric polar vortices. *J. Geophys. Res.*, **104**, 27 191–27 201.
- Yoden, S., 1987: Bifurcation properties of a stratospheric vacillation model. *J. Atmos. Sci.*, **44**, 1723–1733.

APPENDIX A

ADDITIONAL FIGURES

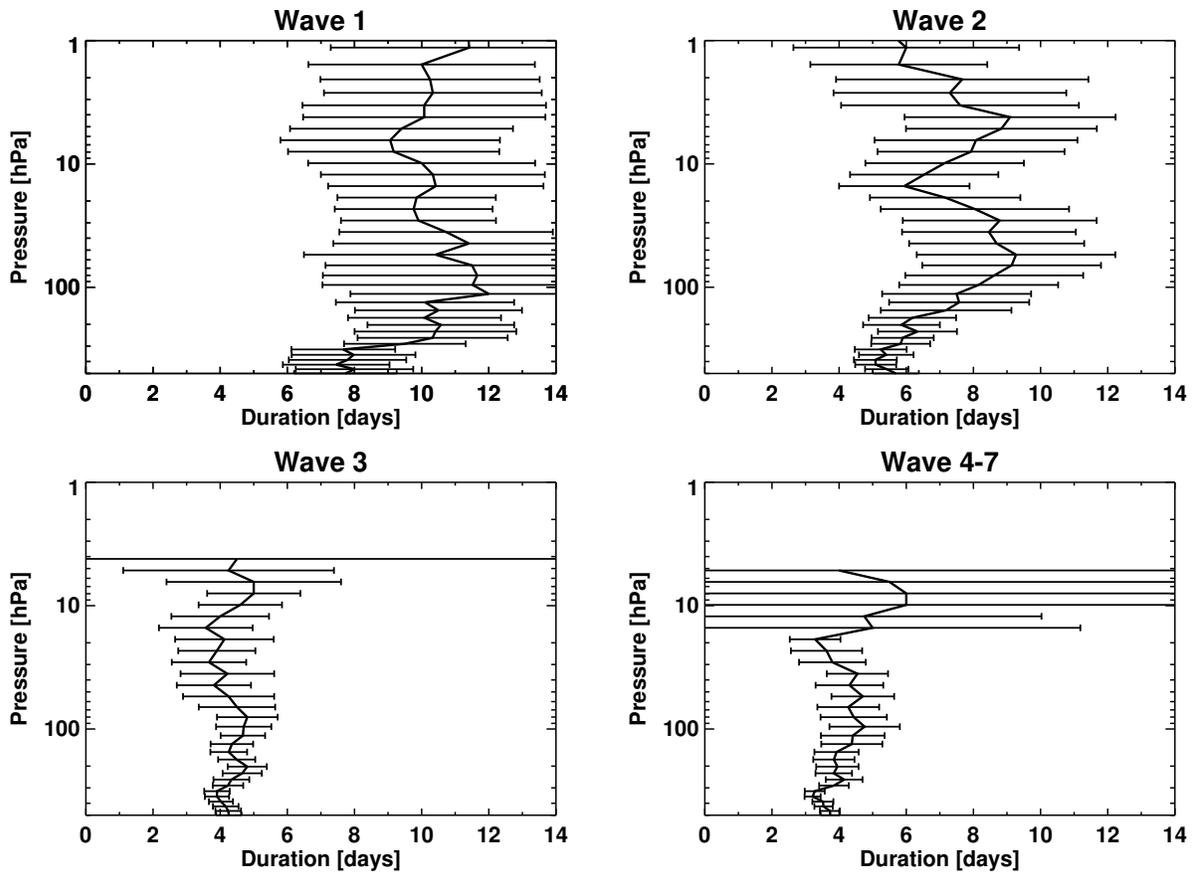


FIGURE A.1. Mean durations for vertically-coherent waves with 500 hPa starting level, here shown for wave 1 (top left), wave 2 (top right), wave 3 (bottom left), and waves 4-7 (bottom right). 95% significance bounds are shown.

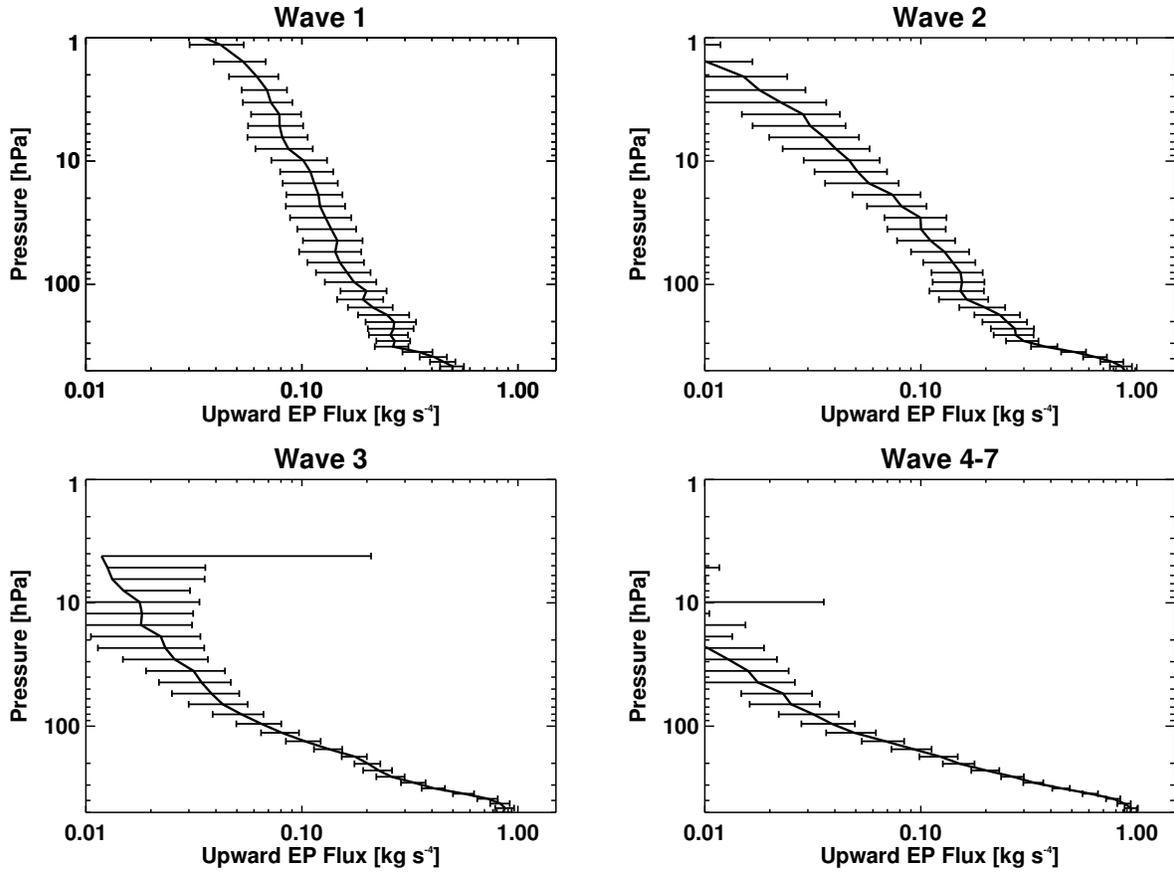


FIGURE A.2. Mean amplitudes for vertically-coherent waves with 500 hPa starting level, here shown for wave 1 (top left), wave 2 (top right), wave 3 (bottom left), and waves 4-7 (bottom right). 95% significance bounds are shown.

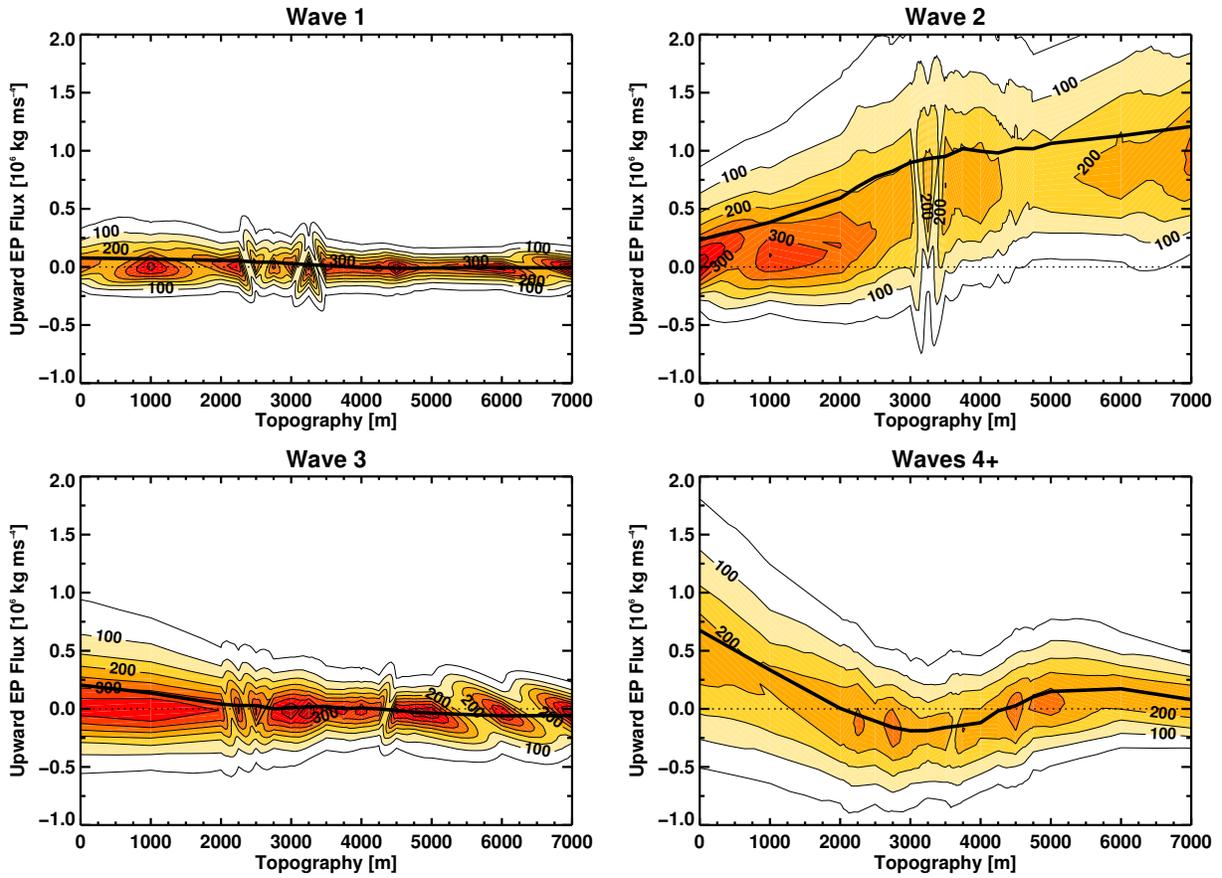


FIGURE A.3. Contours of total occurrences of a given value of upward EP flux at 300 hPa as a function of topographic height. Here, the EP fluxes are decomposed into wave 1 (upper left), wave 2 (upper right), wave 3 (lower left), and waves 4+ (lower right). The solid curve shows the mean upward EP flux for each wavenumber. The dotted line denotes zero upward EP flux. The contour spacing is 50 occurrences.

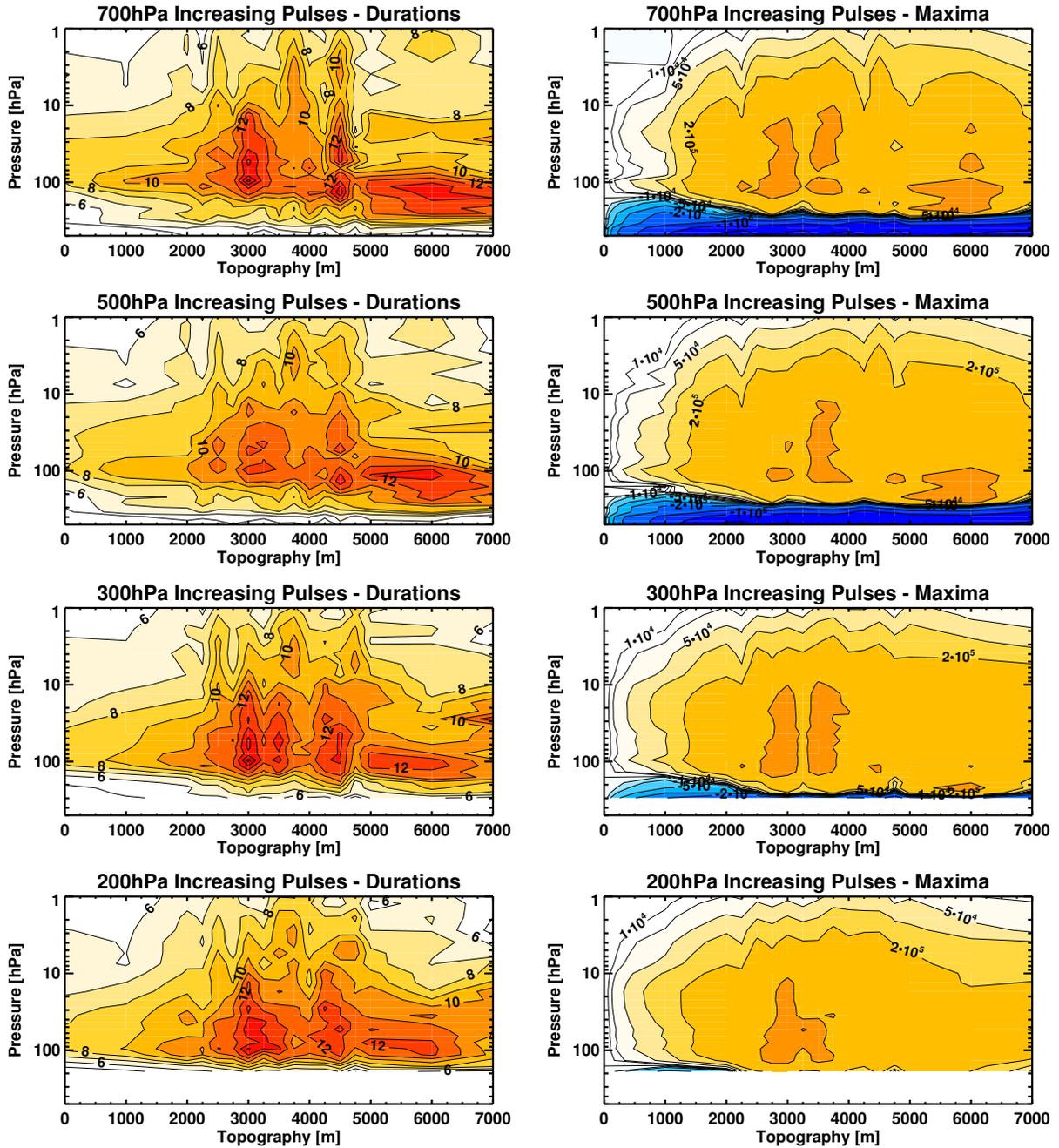


FIGURE A.4. Durations (left column) and upward EP flux maxima (right column) for increasing pulses with 700 hPa, 500 hPa, 300 hPa, and 200 hPa starting levels, in respective descending row order. The contouring is as in Figure 4.8

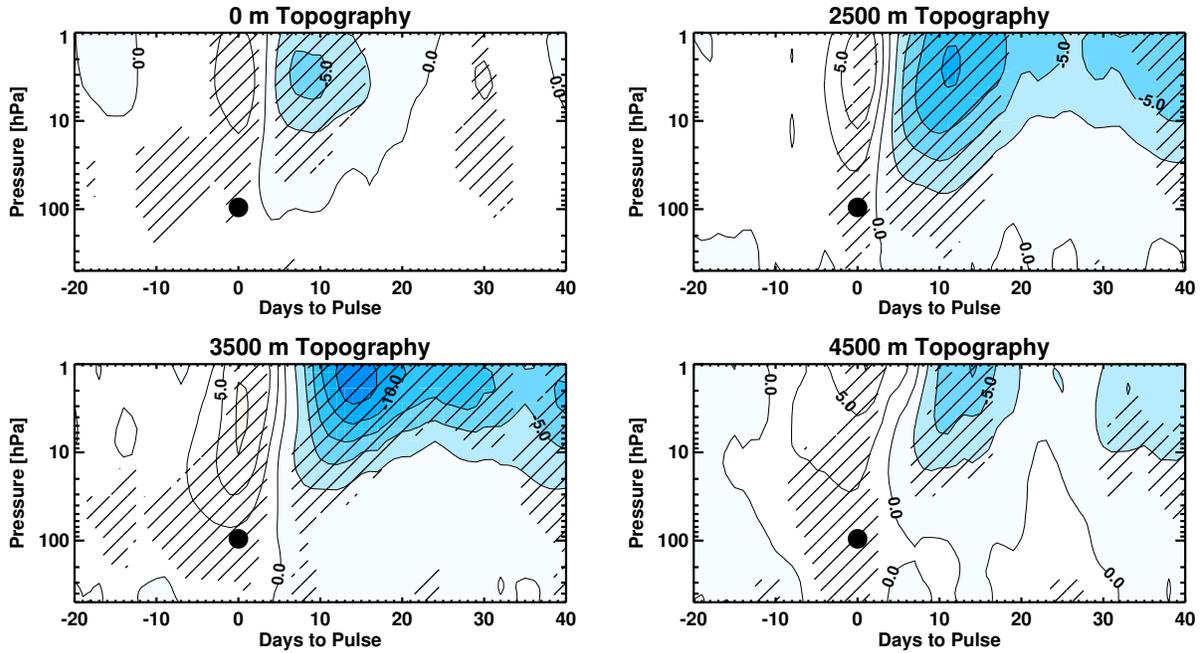


FIGURE A.5. Composite evolutions of the anomalous zonal winds with respect to increasing pulses with starting level of 500 hPa for four topographic heights: 0 m (top left), 2500 m (top right), 3500 m (bottom left), and 4500 m (bottom right). The central date (day 0) is the date at which the increasing pulses arrive at 10 hPa, marked by the black circle. Significance at the 95% level is hatched.

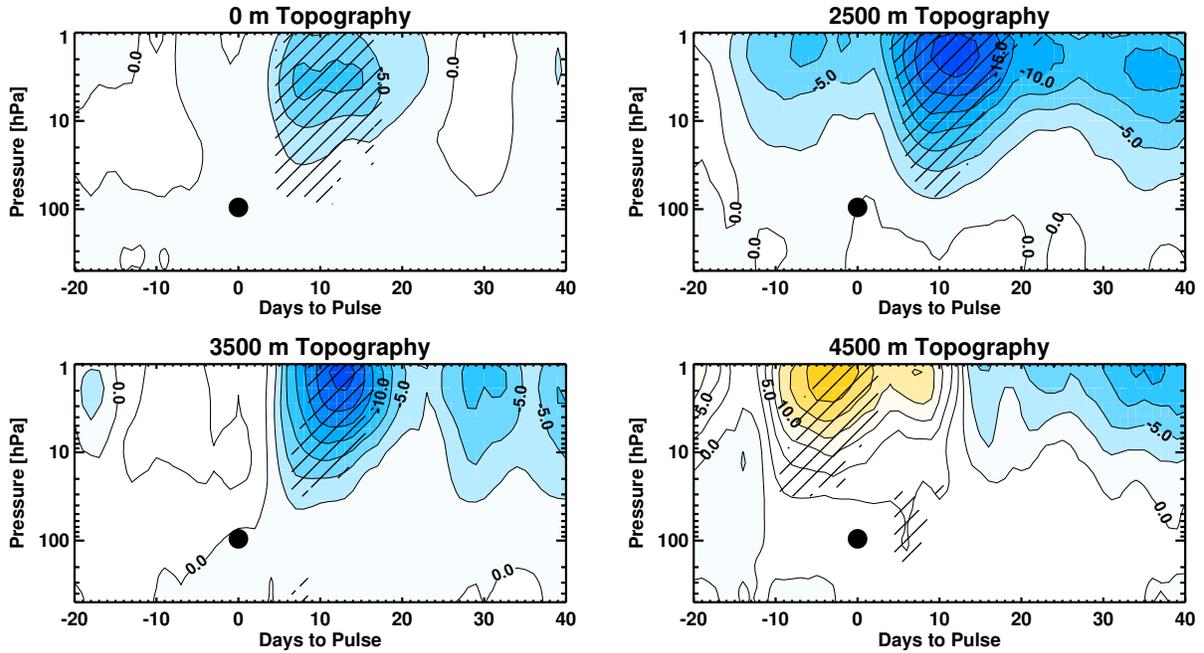


FIGURE A.6. Composite differences in zonal wind between increasing and decreasing pulses with starting level of 500 hPa, shown here for four topographic heights: 0 m (upper left), 2500 m (upper right), 3500 m (lower left), and 4500 m (lower right).

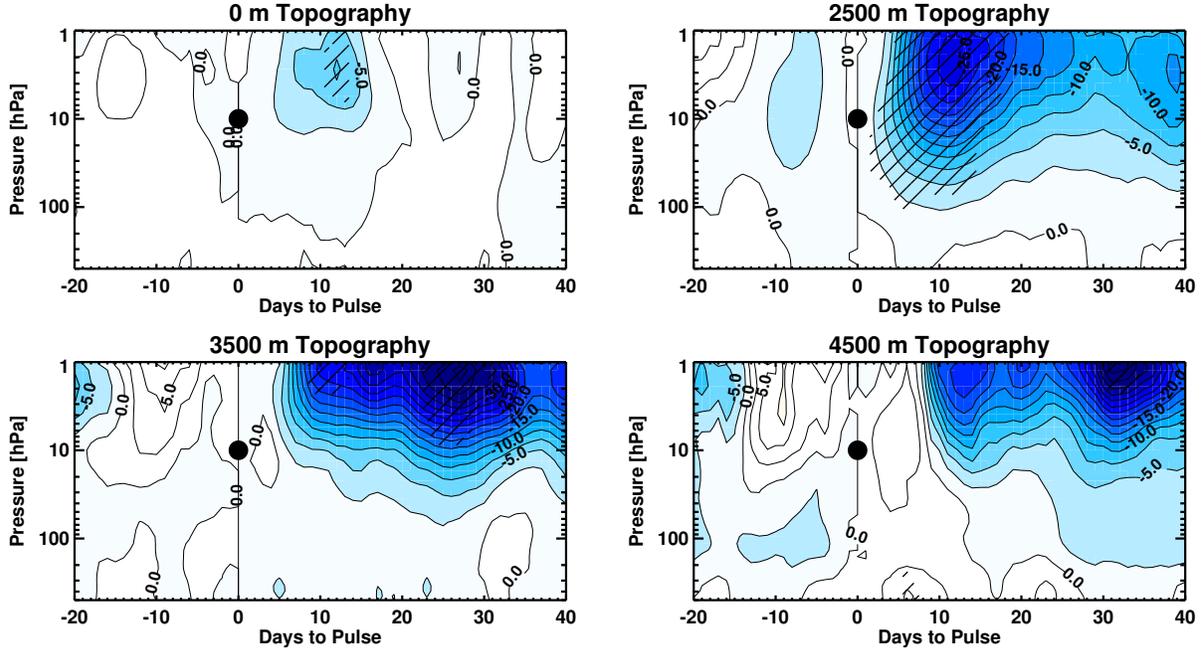


FIGURE A.7. Composite differences in zonal wind changes from the central date between increasing and decreasing pulses with starting level of 500 hPa, shown here for four topographic heights: 0 m (upper left), 2500 m (upper right), 3500 m (lower left), and 4500 m (lower right).