The Role of Wavenumber One and Two
in the Development of Sudden Stratospheric Warmings

by

Andreas Wolfgang Miller

Submitted to the Department of Earth, Atmospheric, and Planetary Science in partial
fulfillment of the requirements for the degree of
Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

© 2017 Massachusetts Institute of Technology. All rights reserved.

Signature redacted
Signature of Author: __

Department of Earth, Atmospheric, and Planetary Science
August 31, 2017

Signature redacted
Certified by: __________
R. Alan Plumb
Professor of Atmospheric Science
Thesis Supervisor

Signature redacted
Accepted by: __________
Robert D. van der Hilst
Schlumberger Professor of Earth and Planetary Sciences
Head of Department
The Role of Wavenumber One and Two in the Development of Sudden Stratospheric Warmings
by
Andreas Wolfgang Miller

Submitted to the Department of Earth, Atmospheric, and Planetary Science on August 31, 2017 in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

In this thesis, we investigate the effects of planetary waves one and two on the polar stratosphere during boreal winter. We use MERRA reanalysis data and the FMS shallow-water model to compare and contrast their propagation into the stratosphere, their interactions within the stratosphere, and their effects on the polar vortex. The results have implications for the predictability of sudden stratospheric warmings (SSWs), theories on the developments of vortex splits and the role of zonal winds in the tropics.

In Chapter 2, we use correlations and regressions to demonstrate that the tropopause affects wavenumber one amplitudes more than wavenumber two. Thus, the statistical predictability of SSWs, based on synoptic events in the mid-troposphere (e.g. blockings), is limited. Composites of extreme heat fluxes reveal that they are likely caused by linear interference of the climatology and anomalies. The phases of anomalous planetary waves align with the climatology only during the largest heat fluxes.

In Chapter 3, the effect of wave-wave interactions within the stratosphere is quantified by analyzing eddy energy budgets. The energy transfer from wavenumber one toward wavenumber two plays a key role in the vortex split in January 2013 and several other SSWs. This mechanism might explain the growth of wavenumber two in the stratosphere in non-resonant conditions. However, wave-wave interactions are small in averages over all splits since 1979 suggesting that different processes can lead to vortex splits and that the common SSW definitions do not capture the timing of planetary wave growth.

In Chapter 4, we employ a shallow-water model to isolate the effects of wave one and two on the polar vortex over a large range of forcing amplitudes and vortex strengths. We are able to simulate SSW splits, which are unequivocally caused by wave-wave interactions. Furthermore, the initial response of the polar vortex depends strongly on the wavenumber of the forcing.

Thesis Supervisor: R. Alan Plumb
Title: Professor of Atmospheric Science
Acknowledgements

First, I want to thank Alan Plumb for his support throughout the years. He has been a great, patient and knowledgeable advisor. I was privileged to learn from him.

I also want to thank my committee members. Paul O’Gorman provided constructive critique from my general exam paper to the corrections of this thesis. I thank him for answering many questions and for chairing the defense. I thank Thomas Birner for his enthusiasm about my research and the detailed feedback on my work. I appreciated that he was able to attend my defense in person. John Marshall helped me distill my results before my pre-defense seminar and answering his questions critically improved this thesis. Furthermore, my research benefited from the discussions with Susan Solomon and the members of Alan’s research group. I am grateful for the conversations with Aditi, Daniela, Erik and Marianna. I thank them and many others in the PAOC community for listening my practices talks and providing feedback on my work. This work would not have been possible without the support of the administrators in PAOC and EAPS. Christine, Faith, Roberta and Vicky reimbursed me for expenses, guided me through the rules of MIT and the department, and were instrumental in the organization of the GCC conferences and many other events. Greg and Linda solved many technical issues and kept my computers running. The staff at the International Student Office, in particular Janka, kindly helped me navigate immigration laws and regulations.

This department has been more than a workplace for me. I look back at many funny, sad, thoughtful, enthusiastic, quiet, exhausting, and ironic moments. Alex, Annie, Ben, Ben, Bryan, Colin, Dan, Dan, Deepak, Emily, Erik, Kane, Marty, Michael, Mike, Sophie, Stephanie and Vince are only a few of the people, who supported me and made EAPS such a special place to me. I also want to thank everyone, who served with me on committees at MIT, the members of the Science Policy Initiative and my teammates on the many intramural teams.

It can be difficult to find work-life balance during graduate school. I was fortunate to find amazing friends outside of EAPS, who kept me grounded and sane. I am grateful for the friendships with Asa, Candy, Cory, Cecilia, Colleen, Fran, Hak, Ivana, Kelly, Laura, Liz, Maria, Markus, Sylvia. I am particularly thankful to have met Dan, Elise and Nate.

I am deeply thankful to my parents, Eva and Stefan. They inspire me and their support easily bridges the Atlantic.

Finally, I thank Kat -for her support, her trust in us, for everything.
"Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth"

–Jules Verne
# Contents

1 Introduction .............................................. 11
   1.1 Stratospheric climate .................................. 12
   1.2 Stratospheric variability .............................. 13
   1.3 Focus of this Thesis .................................... 17

2 Planetary Wave Climatology and large heat flux events ......... 19
   2.1 Introduction .......................................... 19
   2.2 Planetary Wave Climatology ........................... 22
   2.3 Extreme Heat Flux Events .............................. 33
   2.4 Conclusions ........................................... 43

3 The effect of wave-wave interactions on polar vortex splits .... 47
   3.1 Introduction .......................................... 47
   3.2 Energy budget equations ............................... 50
   3.3 Case study ............................................ 53
       3.3.1 Synoptics ....................................... 53
       3.3.2 Energy budgets .................................. 58
   3.4 Composites of Sudden Stratospheric Warmings ............ 65
   3.5 Conclusions ........................................... 67

4 Shallow-water model simulations of SSWs ...................... 71
   4.1 Introduction .......................................... 71
   4.2 Shallow-Water model ................................... 73
   4.3 Results ................................................ 78
       4.3.1 Model evaluation ................................. 78
       4.3.2 Transient Vortex Response ...................... 82
   4.4 Discussion and Conclusions ............................ 99

5 Conclusions and Outlook .................................. 103

Appendix A Additional Figures to Section 2.3 .................. 107
Chapter 1

Introduction

The stratosphere plays an important role for the life on Earth and the details of tropospheric climate. Initially thought of as a passive layer above the troposphere, scientists discovered that the stratospheric impact on the surface goes beyond the radiative properties of ozone. In the beginning of the century, Baldwin and Dunkerton (2001) and others demonstrated that the stratosphere is dynamically coupled to the troposphere on time scales of up to two months. In particular, the state of the winter stratosphere influences weather patterns at the surface. While planetary waves have been shown to play a critical role in communicating signals between the two layers, understanding the details of the coupling mechanism has proven to be difficult. This thesis aims to illuminate the processes, which determine the state of the stratosphere during Northern hemisphere winter. More specifically, the research utilizes reanalysis data and a simplified model to address the difference between planetary waves of zonal wavenumber one and two. Our analysis compares and contrasts these two most important actors in the stratosphere-troposphere coupling with respect to:

- Their climatologies and their average propagation into the stratosphere;
- The development of extreme meridional heat fluxes;
- The role of wave-wave interactions within the stratosphere during vortex splits;
- Their impact on the polar vortex in a shallow-water model.

The purpose of this chapter is to make the results in Chapters 2 to 5 accessible to a broader audience. Experts in stratospheric dynamics may want to skip to the following chapters, which include introductions to the specific research questions. Many of the well-known facts in this chapter and additional information can be found in Andrews et al. (1987); Labitzke and van Loon (1999); Haynes (2005); Waugh and Polvani (2010) and Gerber (2015). We avoid repeated references to these general publications to improve readability. The content of this chapter represents the author’s perspective on the fundamental processes underlying the research in the subsequent chapters, but it should not be considered original
research.

**Stratospheric climate**

The lower atmosphere receives energy from the Earth surface, which is heated by solar radiation. As a result the temperature of the lower atmosphere generally decreases with height. However, the relationship of temperature and height is reversed at higher levels. This change is due to the absorption of UV-light by ozone molecules. It is therefore useful to divide the atmosphere into different layers. The lowest layer, the troposphere, is characterized by decreasing temperatures with height and contains nearly all of the water vapor. Thus, most processes that control the weather at the surface are confined to this layer. The stratosphere is located above the troposphere and the two layers are separated by the tropopause. Climatologically, the tropopause is located between 10-12 km at high latitudes and as high as 18 km in the tropics. The stratosphere extends to about 50 km. In atmospheric science, it is conventional to use pressure instead of height as the vertical coordinate. In these coordinates the tropopause is usually detected between 300 hPa-100 hPa. The upper boundary of the stratosphere is found around 1 hPa.

Historically, the focus of stratospheric research has been on ozone chemistry and how it interacts with solar radiation. This work discovered that stratospheric ozone protects life on Earth from energetic UV-light and culminated in the work of Molina and Rowland (1974) and Crutzen (1970), who were able to explain the low ozone concentration with chemical reaction involving chlorofluorocarbons (CFCs). Their insights ultimately led to the Montreal protocol[^1], which limited the emission of CFCs. As the troposphere contains about 85% of the mass of the atmosphere, conventional wisdom was that the state of the stratosphere (which contains almost all of the remaining 15%) does not affect the state of the troposphere outside of ozone chemistry. The fact that the stratospheric circulation can influence the climate and weather of the troposphere was established in the early 2000s (e.g. Baldwin and Dunkerton (2001)).

The climate of the extra-tropical stratosphere is dominated by the strong contrast between summer and winter. During summer, the absorption of solar radiation by ozone leads to a relatively warm stratosphere and a small temperature gradient between low and high latitudes, which results in weak easterly flow. When the solar radiation decreases in fall, the stratosphere cools quickly because of its minimal heat capacity. The large temperature gradient between low latitudes and the winter pole leads to a strong westerly jet and low pressure at the pole. This large low pressure system is referred to as the polar vortex. The strong zonal (along latitudes) flow around the vortex prevents mixing in and out of the vortex, which isolates the cool winter pole from the warmer surrounding air and further reduces polar temperatures. The low temperatures play a key role in stratospheric ozone chemistry.

At the end of the long winter season, the westerlies become weaker before the vortex breaks up and easterly flow is reestablished. This breakup is called the final warming.

A second important feature of the extra-tropical stratosphere is the difference between the Southern and Northern hemisphere. While the summer conditions are relatively calm in both hemispheres, the Northern hemispheric winter is much more variable than its Southern counterpart. This thesis focuses on the processes that cause the increased variability and is therefore limited to the winter in the Northern hemisphere.

**Stratospheric variability**

There is consensus that the stratospheric variability during Northern hemisphere winters is due to vertically propagating waves. Atmospheric waves can be caused by a number of processes, which range from compressing air on small scales (sound waves) to the differential heating of land and ocean surfaces. The effect of atmospheric waves is mostly confined to the layers where they break. Eliassen and Palm (1961) and Andrews and McIntyre (1976) developed a useful framework to think about linear wave propagation. They defined Eliassen-Palm fluxes (short EP-fluxes), which point in the direction of the wave activity flux. The fact that the vertical EP-fluxes are proportional to the meridional heat flux motivates our analysis in Chapter 2. Furthermore, the divergence of the Eliassen-Palm flux quantifies the effect of eddy-fluxes on the zonal mean flow. There are two main mechanisms, which cause waves to break. First, waves break if their amplitude is comparable to their wavelength. Second, waves break on critical lines (also critical layer) where the phase speed of the wave is identical to the mean flow. The two most important classes of vertically propagating waves are gravity and planetary waves. Gravity waves are usually generated by flow over topography and by convection. Their typical wavelengths are relatively small and their phase speeds are large. Gravity waves tend to propagate through most of the stratosphere and break in the upper stratosphere and the mesosphere where they dominate the momentum budget Andrews et al. (1987). As their amplitude increases with decreasing density (increasing height), these waves often break because their amplitude becomes large compared to their wavelength. In contrast, planetary waves tend to be excited by large scale features like mountain ridges and the contrast between oceans and continents. They propagate vertically and westward in the absence of mean flow. Charney and Drazin (1961) developed the seminal theorem, which identifies the conditions for vertical propagation of planetary waves. Assuming a quasi-geostrophic beta-plane, and constant stratification and zonal mean wind, the theorem can be written as (Andrews et al. 1987):

$$0 < \bar{u} - c < u_c = \frac{\beta}{(k^2 + l^2) + \frac{f_0^2}{4H^2N^2}},$$  \hspace{1cm} (1.1)

Where $\bar{u}$ is the zonal mean wind, $c$ is the phase speed, $u_c$ is the critical zonal wind, $\beta = \partial f/\partial y |_{45^\circ N} = const$, $f_0$ is the Coriolis parameter at $45^\circ N$, $H$ is the characteristic scale height, $N$ is the buoyancy frequency, which includes the vertical temperature gradient, $k$ and $l$ are
the zonal and meridional wavenumbers. In general, stationary waves are unable to propagate through easterly flow (where $\bar{u} < 0$), which explains the small variability of the stratosphere during summer in both hemispheres. The key to understanding the Charney-Drazin theorem is that $u_c$, depends on the zonal wavenumber $k$. Thus, given the same atmospheric conditions, vertical wave propagation depends on the zonal wavenumber $k$. During winter, we expect waves with longer zonal wavelength (with larger critical velocity) to break at greater altitudes. Even the longest wave (wavenumber 1), usually breaks in the mid stratosphere around 10 hPa. The planetary wave breaking explains the larger variability of the westerly winds during winter. Furthermore, the Southern hemisphere contains fewer prominent mountain ridges and less land mass in the mid-latitudes, which leads to reduced planetary wave activity for wavenumbers that typically reach the stratosphere. Thus, the polar vortex in the stratosphere is less disturbed than the one in the North. The stability of the polar vortex in the Southern hemisphere leads to extremely cold temperatures, which are necessary for the formation of polar stratospheric clouds. These clouds unlock the mixed-phase chemistry, which explains the dramatic ozone loss over the Antarctic (ozone hole). In summary, planetary waves represent the dominate coupling mechanism between the troposphere and the stratosphere. However, the critical layers also depend on the mean flow in the stratosphere.

The most prominent representation of planetary wave breaking in the stratosphere is the sudden weakening of the polar vortex. These events are referred to as sudden stratospheric warmings (SSWs), because polar temperatures can rise by 50 K over a short period of time in response to the vortex collapse. SSWs were first described by Scherhag (1952). Since then a substantial body of literature has focused on these extremes events. Considering the large variability of the Northern polar vortex, SSWs were believed to occur exclusively in the Northern hemisphere. However, in 2002, a SSW was detected in the Southern hemisphere, when the Antarctic polar vortex split in two. The unanticipated event lead to a number of publications focusing on the dynamics of the Southern hemisphere. At this point, it is still unclear if the SSW in the Southern hemisphere was the result of extreme planetary wave forcing, resonant wave interaction or an outlier in vortex behavior.

Describing the complicated dynamics during SSWs has led to a number of different definitions for the sudden weakening of the polar vortex. The different definitions try to capture the vortex behavior in various datasets and models of varying complexity while focusing on distinct features (wave amplitudes, critical lines, vortex shape, etc). Butler et al. (2016) recently showed that the differences between SSW definitions can lead to significant differences in the results. Unfortunately, there is no uniformly accepted definition of SSWs at this point, which complicates the comparison between studies. One of the most popular measures of vortex breakups is a change of the zonal mean zonal wind at 60°N and 10 hPa from westerly to easterly. The composite study by Charlton and Polvani (2007) found about six SSWs per decade based on this definition. It is important to keep in mind that the low number of well observed SSWs limits the possibilities to test theories and models. Independent of the definition, SSWs are consistently found to be preceded by anomalously large
wave activity fluxes from the troposphere into the stratosphere. Supported by early modeling studies (Matsuno 1970, 1971; Holton and Mass 1976) a theory emerged that relatively small variations in the planetary wave activity could lead to abrupt instabilities of the polar vortex.

More reliable observations of SSWs in the 1970s sparked the idea that fundamentally different classes of events exist. The most common categories of SSWs are splits and displacements. Unsurprisingly, the exact definition of these terms is still debated. Displacements are generally characterized by a large high pressure system in the mid-stratosphere that coincides with a movement of the polar vortex towards lower latitudes. The high pressure has been found to develop preferentially over the Aleutian Islands and a vortex position over the Eurasian continent. In the vertical, displacements are associated with consistent westward tilt, which is conducive to vertical wave propagation. In contrast, the polar vortex is separated into two smaller distinct vortices during stratospheric splits. As displacements, splits seem to have a preferred orientation, with vortex centers over North America and Siberia. In general, splits are more barotropic than displacements. However, Matthewman et al. (2009) showed a westward tilting polar vortex at the central date of their composite of splits.

The discovery of different SSW types was important for two reasons. First, a number of studies (Matthewman et al. 2009; Mitchell et al. 2013) found statistically significant differences in the stratosphere during SSWs. More importantly, new theories on the mechanisms causing SSWs were developed. In particular, the traditional model for SSWs required significant westward tilt throughout the atmosphere to explain the upward propagation of planetary waves. The work of Plumb (1981), Esler and Scott (2005) and Matthewman and Esler (2011) presented theoretical and modeling evidence that resonance could occur in barotropic atmospheres and lead to the growth of wavenumber two amplitudes before split events. The idea was further supported by the observation that the stratosphere is usually dominated by wavenumber one Dunn-Sigouin and Shaw (2015). However, it is difficult to test atmospheric data for resonant behavior and thus prove the aforementioned theories. Furthermore, it became clear that the stratosphere is not simply responding to tropospheric forcing. Models of varying complexity show SSW-like behavior as part of internal stratospheric variability (Holton and Mass 1976; Scott and Haynes 2000; Scott and Polvani 2004). Additionally, the work of Scott and Polvani (2004, 2006); Hitchcock and Haynes (2016) presented evidence that the stratosphere partially controls the wave flux from the troposphere into the stratosphere. Thus, it became clear that SSWs are the result of a complex coupling between the two layers.

The importance of understanding the polar vortex and its variability was dramatically enhanced when Baldwin and Dunkerton (2001) and Thompson et al. (2002) demonstrated that the state of the polar vortex has statistically significant and long-lasting effects on the troposphere. Most notably, the front between polar and mid-latitude air is displaced following SSWs for up to two months. These time scales are crucial for seasonal weather predictions, which are important to industries in the logistics, energy and agriculture sectors. Due to the small mass of the stratosphere and the lack of evidence for large intrusions of strato-
spheric air during SSWs, it was clear that stratospheric air was not affecting tropospheric
temperatures and winds directly.

It makes sense intuitively that the abrupt changes in zonal mean flow affect the critical
lines of planetary waves and thus where planetary waves break. A change in planetary wave
breaking might then be able to affect the tropospheric circulation. However, the detailed
processes linking stratospheric and tropospheric variability are still not completely under-
stood. The most popular theory is called “downward control”. The basic idea is that, due
to continuity, the EP-flux divergence in the stratosphere itself causes a change in the flow
at lower levels. Wave reflection could be an alternative explanation (Perlwitz and Harnik
2004). The stratospheric influence on the troposphere appears to depend on the details of
the vortex breakup (Maycock and Hitchcock 2015; Runde et al. 2016) and involve smaller
tropospheric waves (Song and Robinson 2004). Matthewman et al. (2009); Mitchell et al.
(2013) found statistically significant different effects of splits and displacements on the tro-
posphere. However, Coughlin and Gray (2009) found no such differences and Karpechko
et al. (2017) determined that other factors are more predictive of the surface impact of
SSWs. In summary, stratosphere and troposphere affect each other during winter in com-
plex ways forming a two-way coupled system, which influences the ozone concentrations
in the stratosphere and the weather and climate patterns in the troposphere.

Studies of the atmosphere always rely on observations - either as the primary subject of an
analysis or as the basis and test for theories and models. The first systematic observations
of the stratosphere were recorded in the early 20th century, significantly later than tro-
pospheric measurements. However, observations remained extremely sparse in space and
time for decades after the initial measurement campaigns. Scherhag (1952) describes in
the aforementioned publication that the weather station in Berlin had started daily balloon
measurements of the stratosphere in 1951. Trying to compare his observation of the sudden
warmings, he was unable to find any balloon measurements of the same heights and only
five at lower stratospheric levels. This description illustrates that it is impossible to perform
modern atmospheric analysis on data as recent as the 1950s. Since there are no recon-
structions of stratospheric observations and, outside of few volcanic outbreaks, no proxy
data of stratospheric climate exist, the observational record is short. The situation funda-
mentally changed with the development of space technology. Rocket missions required a
much deeper understanding of the higher atmosphere and allowed measurements at high
levels. Most importantly, satellites revolutionized atmospheric observations by dramati-
cally increasing the number of observations - especially where traditional measurements
had been sparse. This thesis, and many other studies today, are only based on observations
after the regular satellite observations began in 1979. It is worth mentioning that satellites
also marked the change from in-situ to remote sensing observations. While remote sens-
ing technology has progressed rapidly over the last decades, the observations still have
larger uncertainties than most in-situ technologies. However, this disadvantage is greatly
outweighed by the amount of available information.

Independent of the type and density of observations, it is not trivial to combine measure-
ments into a coherent estimate of the state of the atmosphere at a given point of time. Today, thousands of measurements are recorded with instruments ranging from buoys and weather stations to radiosondes and satellites. Each observation is associated with different uncertainties, represents different parts of the atmosphere and is performed at different intervals. Furthermore, new instruments are constantly deployed and old ones lost. As a result, major research and logistics efforts are necessary to reconstruct the atmosphere from available measurements. Several centers around the world build numerical models to assimilate the information and produce their best guess of the state of the atmosphere over the last decades. These reanalysis products have many advantages. They combine different observations, eliminate unrealistic outliers, account for changes in instrumentation and convert the data to a regular grid. Furthermore, the model fills gaps in time and space with dynamically consistent results. As most observations only measure a small number of variables, reanalysis models also estimate variables, which were not observed, using available observations as constraints. Reanalysis datasets have become the standard in stratospheric dynamics and are often treated like observations. However, it is important to keep in mind that the reanalysis always includes model corrections and errors, which change over time, space and with the variable of interest. Fortunately, this thesis is focused on large-scale processes, which develop relatively slowly. Additionally, observations of the pressure field and the horizontal flow field are generally considered reliable and well constrained by the model dynamics. However, we use vertical velocities in Chapter 3, which are usually associated with larger uncertainties. In summary, reanalysis data have led to much more reliable access to observations and are adequate for the research of large scale stratospheric dynamics.

Focus of this Thesis

In this thesis, we focus on the influence of planetary waves on the stratosphere during Northern hemisphere winter. In particular, we investigate the differences between zonal wavenumbers one and two. We use reanalysis data and a shallow-water model to explore the details of wave propagation from the troposphere into the stratosphere, extreme events in meridional heat fluxes of either wavenumber, the role of wave-wave interactions within the stratosphere and the effect of the waves on the polar vortex in idealized configurations. Previous studies have often assumed that the only difference between the vertical propagation of wavenumber one and two is captured by the Charney-Drazin theorem. However, the aforementioned work on the development of splits and displacements suggests that, at least for extreme events, things are more complicated.

In Chapter 2.2, we determine the climatological differences in heat fluxes of the two wavenumbers. Our results show that the propagation of the large planetary waves is less predictable than previously assumed. We believe the results serve as important context to the study of extreme events and the numerous attempts to predict SSWs from tropospheric

---

2 more information on reanalysis data in general and specific products can be found at https://reanalyses.org/
events. Our analysis shows that wavenumber one dominates the geopotential and heat flux at most latitudes and levels during the winter, although the inter-annual variability is large. More surprisingly, the correlation between wave one heat fluxes below the tropopause and in the stratosphere is negligible and therefore has no effect on the strength of the polar vortex. It is only above the tropopause that wavenumber one can readily propagate into the stratosphere and influence vortex behavior. In contrast, wavenumber two correlates well across most levels, but the wave amplitudes are too small to affect the polar vortex. We then build composites of extreme heat fluxes of either wavenumber and find that the phases only align for a few days through most of the atmosphere when the largest heat fluxes occur (Chapter 2.3). This result supports the idea that planetary waves grow as the result of linear interference between the climatology and the anomaly.

Motivated by the analysis of extreme heat flux events, we investigate the role of wave-wave interactions within the stratosphere as a possible cause for the growth of wavenumber two (Chapter 3). We use eddy energy budgets of the stratosphere to show that the growth of wavenumber two during the split in January 2013 was the result of energy transfer from wavenumber one to wavenumber two within the stratosphere. This result challenges the assumption that the type of warming is solely defined by the dominant wavenumber of the tropospheric forcing. However, our results can not easily be generalized to all splits in the reanalysis data. We discuss the impact of sample size, SSW definitions and characterizations on our results. We suggest that multiple mechanisms exist, which cause the polar vortex to split.

Finally, we modify a shallow water model to investigate the growth of wavenumber one and two as a function of surface topography and polar vortex strength. The highly simplified setup allows us to perform over 2500 simulations over a large range of parameters. The experiments with wave two forcing expand on the recent results of (Liu and Scott 2015) by investigating the predictability of the vortex behavior using ensembles for every parameter combination. Furthermore, we show the relationship between the development of chaos and the formation of wavenumber one amplitudes for sub-critical forcing amplitudes. The sensitivity of the polar vortex to the details of the planetary wave forcing are illustrated by the different responses to the same topography with and without tropical easterlies. We observe no SSW-like vortex collapse if a representation of solid body rotation is added to the model setup. Our investigation shows that, outside of the quasi-resonant growth of wavenumber two, the maximum wave amplitudes increase similarly with surface topography and vortex strength. However the wave amplitudes result in fundamentally different vortex responses. Chapter 5 summarizes the findings of the thesis and outlines the new questions resulting from this work.
Chapter 2

Planetary Wave Climatology and large heat flux events

Introduction

In Chapter 1, we introduce the concept that the two-way stratosphere-troposphere coupling is largely explained by wave-mean flow and wave-wave interactions of vertically propagating planetary waves. The most prominent feature of this coupling are SSWs and their effect on the circulation in the troposphere. As a consequence, the main focus of stratospheric dynamics has been on composites of extreme events. These studies can be coarsely divided into three groups based on how the events are defined.

Baldwin and Dunkerton (2001) established that extreme states of the polar vortex are statistically significantly correlated to the state of the troposphere two months after the event is detected. In the following years, a number of studies explored composites of events based on the state of the polar vortex Thompson et al. (2002); Limpasuwan et al. (2004); Charlton and Polvani (2007). This line of work established that planetary waves affect the state of the polar vortex and are also responsible for the downward propagation of stratospheric information. As a result of changes in stratospheric circulation, planetary waves break at different heights, which affects the heat and momentum fluxes and ultimately the circulation in the atmosphere. However, Hitchcock et al. (2013), Butler et al. (2017) and others also found large variability among events in the stratosphere and troposphere. Furthermore, composites based on the state of the stratosphere explored a number of different event definitions (see Chapter 3) and developed at least two types of SSWs (splits and displacements). The question if different event types have significantly different effects on surface climate remains an active field of research (Coughlin and Gray 2009; Cohen and Jones 2012; Seviour et al. 2013; Mitchell et al. 2013; Maycock and Hitchcock 2015).

As a result of the inconsistent tropospheric effects of events defined in the stratosphere, a
second group of studies has focused on composites based on the surface impact on the troposphere (e.g. Zhou et al. 2002). Using reanalysis data and state of the art GCMs, Maycock and Hitchcock (2015) and Runde et al. (2016) found that the persistence of anomalies in the lower stratosphere is significantly correlated to the downward coupling of stratospheric events. In contrast, the amplitude of the stratospheric signal is not very predictive of surface impact. Investigating a large number of SSWs with a dry-dynamical core model, Jucker (2016) found no significant differences in the stratosphere between events with and without influence on the troposphere. This result suggests that downward propagation may depend on tropospheric processes and the details of numerical simulations. The complicated mechanism of downward propagation is further illustrated by the work of Polvani and Kushner (2002) and Song and Robinson (2004) who highlight the role of synoptic and planetary waves respectively.

Our work is part of the third group of studies, which focuses on the processes leading up to vortex events. There are a number of studies that built composites on events based on different measures of wave propagation from the troposphere to the stratosphere (e.g. vertical EP-flux, meridional heat flux, geopotential wave amplitudes at different levels). Polvani and Waugh (2004) were able to reproduce many characteristics of SSWs in composites based on the meridional heat flux at 100 hPa instead of the state of the vortex. They suggested that it was in fact the tropospheric signal that acted as a forcing on the polar vortex and finally affected the troposphere. However, Scott and Polvani (2004) and Hitchcock and Haynes (2016) found that the flux into the stratosphere depends in turn on the state of the polar vortex. Thus, a picture of a fully coupled system emerged. The vertical wave propagation depends on both the tropospheric forcing and the state of the vortex and the same is true for the effect of the stratosphere on the troposphere. A logical consequence of the established link between tropospherically generated planetary waves and the state of the polar vortex were attempts to link tropospheric events to large heat fluxes and SSWs. In her early work Labitzke (1977) identified the Aleutian as an important region in the development of SSWs. More recently, Martius et al. (2009); Colucci and Kelleher (2015); Ayarzagüena et al. (2015) established that almost all SSWs are preceded by tropospheric blocking patterns and that certain blocking locations are more likely to be associated with splits or displacements. However, the vast majority of blockings does not lead to extreme vortex events. If blockings are associated with large heat fluxes from the troposphere into the stratosphere, they are more likely to precede stratospheric vortex events. However, it is unclear if the blockings are the cause of the anomalously large heat fluxes. Furthermore, there is no established method to predict the heat flux, and thus the stratospheric impact, associated with individual blocking events. Finally, Watt-Meyer and Kushner (2015) found that large heat fluxes are largely explained by linear interference of the climatological wave and the anomalous wave at the time of the largest heat fluxes.

Relating tropospheric signals to stratospheric responses is further complicated by the finding of Dunn-Sigouin and Shaw (2015) that the heat flux is dominated by its wavenumber one component at most times. This result supports the ideas of Esler and Matthewman (2011)
and Matthewman and Esler (2011) that splits and displacements might be caused by different processes. While the propagation of wave one into the stratosphere is probably sufficient to explain the development of displacements, it is not well understood how small wave two heat fluxes lead to splits of the polar vortex. The two most prominent theories explain the amplification of the small tropospheric wave two signal with non-linear resonance. Plumb (1981) suggested that the atmosphere could create a wave cavity that traps planetary waves near the pole and leads to downward reflection, which can result in resonant interference. Esler and Scott (2005) showed that simple models can produce split warmings as the result of resonance of the barotropic mode. Plumb (2010) demonstrated that, in practice, it is not trivial to distinguish the two forms of resonance. The role of wave-wave interactions during SSWs was discussed by Scott and Dritschel (2006) and will be the focus of Chapter 3.

As discussed in Chapter 1, we rely on reanalysis data to study the development of the stratosphere since 1979. There are four major sources for reanalysis products: the European Center for Medium-Range Weather Forecasts (ECMWF), the National Centers for Environmental Prediction in collaboration with the National Center for Atmospheric Research (NCEP/NCAR), the Japanese Meteorological Agency (JMA) and National Aeronautics and Space Administration (NASA). Each organization has published several reanalysis products. For recent comparisons see Butler et al. (2017) and Fujiwara et al. (2017). Martineau and Son (2010) and Martineau and Son (2015) found that stratospheric variability and troposphere-stratosphere coupling are reasonably well represented in all datasets; presumably because of the large scale nature of most underlying processes. These processes can usually be observed from space and rely less heavily on sub-grid parameterizations than many tropospheric processes. The exception is the vertical velocity field, which is sensitive to smaller scale processes like convection and gravity waves. Larger biases can also be expected above 10 hPa (Manney et al. 2003) and before the satellite area where observations are sparse. At the time we had to choose a dataset MERRA (version 1) was a state of the art product to study stratospheric dynamics. MERRA is developed by NASA and has since been updated to version 2. We use the three-dimensional wind field, temperature and geopotential height on 42 pressure levels between the surface and 0.1hPa. The reanalysis model output is saved every 3 hours between 1979 and 2014 with a horizontal resolution of 288 longitudes and 144 latitudes, which translates to 1.25° × 1.25° grid boxes. In summary, there is no perfect dataset to study troposphere-stratosphere coupling, but we are confident that the results in the following two chapters are robust and not an artifact of our choice for MERRA.

The two-way coupling between the troposphere and stratosphere is strongest during extreme events. As a result, the literature has focused almost exclusively on events exceeding certain thresholds. Most publications on the climatological state of the stratosphere are either outdated (Mcinturff 1978; Harvey and Hitchman 1996) or primarily compare different datasets Martineau and Son (2010). Dunn-Sigouin and Shaw (2015) and Watt-Meyer and Kushner (2015) used the full winter seasons to quantify the importance of wavenumber

---

1http://disc.sci.gsfc.nasa.gov/daac-bin/DataHoldings.pl
one and linear interference to explain meridional heat fluxes respectively. However, few publications have focused on the climatology of wave amplitudes and heat fluxes from the troposphere to the stratosphere. Analyzing the climatology of heat fluxes provides important context to the composites of extreme events. The definition of these extremes necessarily involves choices on thresholds and where they should apply. Our results can inform these choices and help understand the differences between composites. Additionally, we calculate correlations between different atmospheric levels and variables to inform the search for tropospheric precursors of stratospheric vortex events described above. Thus, the first part of this chapter will investigate planetary wave amplitudes and meridional heat fluxes for wavenumber one and two at all levels of the MERRA reanalysis data without defining any events. In the second part, we compare and contrast composites of large wave one and wave two heat fluxes around the tropopause level. We investigate the interplay of wave one and two heat fluxes during extreme events, which synoptic features are statistically correlated to large heat fluxes, and the vertical structure of the waves over the period of the composites. In particular, we address the question why previous efforts to predict extreme stratospheric heat fluxes from tropospheric observations have been unsuccessful and what mechanisms could lead to the growth of wavenumber two within the stratosphere.

**Planetary Wave Climatology**

As the propagation of planetary waves into the stratosphere strongly depends on the mean flow, it is useful to limit a climatology to the season when planetary waves can reach the mid-stratosphere. We define an extended winter or cold season following Charlton and Polvani (2007) from the beginning of October through the end of April. Figures 2.1 and 2.2 are based on this dataset. As Mitchell et al. (2011); Dunn-Sigouin and Shaw (2015) and others have used shorter winter seasons, we repeated the analysis in this section with a winter season from November to March. While we find no qualitative differences based on the length of the winter season, we use this shorter season in Figures 2.3, 2.4 and 2.5 for easier comparison to recent publications.

Figure 2.1 quantifies the importance of wavenumbers one to three at different levels and latitudes averaged from 1979 to 2014. For every time step and at every level and latitude, we calculate how much of the variance in geopotential is explained by the first three harmonics. Figure 2.1 shows histograms of how often a given wavenumber explains a certain percentage of the variance for a given level and latitude. At 50°N and 500 hPa, no wave explains less than 10% or more than 40% of the variance (see Figure 2.1a). We know from observations and from the Charney-Drazin theorem Charney and Drazin (1961) that long waves propagate higher into the stratosphere. The result is the large variance explained by wave-number one in Figures 2.1c and 2.1d compared to 2.1a and 2.1b. A less appreciated fact is that smaller wavenumbers also become increasingly important towards the pole (Hoskins and Karoly 1981). The effect becomes clear when histograms at the same level but at different latitudes are compared. For example, wavenumber one almost never explains more
than 80% of the variance at 50°N and 10 hPa (Figure 2.1c), but does so almost all the time at the same level at 80°N (Figure 2.1d). Our results agree qualitatively with Birner and Albers (2017), who found that the wavenumber one component of the upward planetary wave activity flux becomes more important with height.

Figure 2.1 shows the average state of the atmosphere. However, as explained in Chapter 2.1, much attention has been paid to extreme events and the variability of planetary wave activity. Thus, Figure 2.2 depicts timeseries of the number of timesteps at which the amplitude of wavenumber two exceeds the amplitude of wavenumber one at 5 different levels and 2 latitudes. It is clear from Figure 2.2a that the amplitude of wavenumber two is larger than wavenumber one almost 50% of the time in the upper troposphere and even into the stratosphere. However, the inter-annual variability is large (compare years 1982 and 1985). The increased importance of wavenumber one with increasing latitude and height becomes apparent when Figure 2.2a is compared to Figure 2.2b. At 80°N, wavenumber two is still larger than wavenumber one about half the time at 250 hPa, but the importance of wavenumber two decreases more quickly with height closer to the pole. On the other hand, the inter-annual variability is smaller at 80°N, indicating that most of the extreme events are due to planetary wave activity near the vortex edge as described by Polvani et al. (1994). While the results in Figures 2.1 and 2.2 are not necessarily surprising, they serve as important context to interpret the following results and other publications that only report amplitudes during certain times. For example, extreme events are often defined based on the anomalous wave amplitudes over certain latitude bands and sometimes classified by the relative amplitudes of wavenumber one and two. Our results show that the choice of these criteria can have large effects on the selected events.

As SSWs are consistently found to be preceded by anomalously large wave amplitudes, vertical EP-fluxes and EP-flux divergence in the stratosphere (Charlton and Polvani 2007), it was natural to look for links between tropospheric features that could cause the anomalous planetary wave activity. While blockings and other weather patterns in the troposphere are correlated to the state of the stratospheric polar vortex, the correlation coefficients are small and it remains unclear what distinguishes tropospheric features that do or do not affect the stratosphere (Colucci and Kelleher 2015). Furthermore, these studies usually ignore the state of the stratosphere in which the waves propagate. We are using the MERRA reanalysis climatology to understand how planetary waves propagate into the stratosphere and how they affect the polar vortex. Correlating planetary wave amplitudes, meridional heat fluxes and the state of the vortex over all time steps and at various levels provides an upper bound on the predictability of stratospheric dynamics from the troposphere.

For our analysis, we integrate the square of the first two harmonics of geopotential over the Northern hemisphere and correlate the time series at three different levels with itself, the zonal mean zonal wind at 60°N and meridional heat flux of the same wavenumber integrated over the Northern hemisphere. Figures 2.3, 2.4 and 2.5 show the correlations based on the time series at 500 hPa, 100 hPa and 10 hPa respectively. We calculate statistical significance using the auto- and cross-correlations of each variable pair at each level. The shading indi-
Figure 2.1: Histograms of percent of variance of geopotential height explained by wavenumbers 1 (black), 2 (red) and 3 (blue) based on an extended winter season from October to April (212 days). a) 50°N and 500 hPa b) 80°N and 500 hPa c) 50°N and 10 hPa d) 80°N and 10 hPa
Figure 2.2: Time series of the number of 3-hour occurrences per extended winter (212 days or 1696 occurrences) when the amplitude of wavenumber two exceeds the amplitude of wavenumber one at 5 levels between 250 hPa and 1 hPa. a) 50°N b) 80°N

cates times and levels where the correlations are not statistically significant at the 95% level. Figure 2.3a depicts the correlation of the geopotential wave amplitude of wavenumber one at 500 hPa against itself. We find that the correlation of wavenumber one declines sharply for lags beyond ±5 days and levels above 100 hPa. The correlation with the amplitudes in the stratosphere are below 0.2 at all times. The correlation between the wave amplitude and the heat flux (Figure 2.3c) has a very similar structure although the values are below 0.5 at all times and levels. The mechanism, by which blockings are speculated to affect the polar stratosphere, is based on amplified planetary wave amplitudes during blocking events. Our results show why predicting planetary wave propagation from the mid-troposphere into the stratosphere is difficult. First, blockings are an imperfect predictor of planetary wave amplitudes in the troposphere and stratosphere (Martius et al. 2009; Colucci and Kelleher 2015). Additionally, the correlation of planetary wave one amplitudes within the troposphere fall off quickly. Thus, forecasts of tropospheric wave amplitudes might be limited to time scales of less than one week, which is a typical time scale for the prediction of many tropospheric phenomena. Furthermore, the correlation between wavenumber one amplitudes and heat fluxes is below 0.5. Most strikingly, the amplitude of wavenumber one in the troposphere is almost uncorrelated to the wave amplitude or the heat flux in the stratosphere indicating that processes above 500 hPa are crucial to understand planetary wave propagation into the stratosphere. This result is underlined by Figure 2.3b which depicts the correlation of the wavenumber one amplitudes with the zonal mean zonal wind at 60°N. We use the zonal mean wind as a crude measure for the strength of the polar vortex and the circulation in the troposphere. We find that, statistically, the wavenumber one amplitudes at 500 hPa have almost no correlation with the zonal mean zonal wind at any level. This is not surprising in the troposphere where the zonal mean is influenced by many different waves. The small correlations in the stratosphere illustrate the complex relationship between tropospheric waves and the state of the stratospheric polar vortex.

Figures 2.3d, 2.3e and 2.3f are equivalent to Figures 2.3a, 2.3b and 2.3c for wavenumber
Figure 2.3: Correlation of the square of geopotential wave amplitude at 500 hPa integrated over the Northern hemisphere (area-weighted) with itself (left column), the zonal mean zonal wind at 60°N (middle) and the meridional heat flux integrated over the Northern hemisphere (area-weighted) (right) from 1000 hPa to 1 hPa and lags from -30 to +30 days. a) and c) show the wavenumber one components; d) and f) show wavenumber two. Shaded values are not statistically significant at the 95% level.
two instead of one. Figure 2.3d shows that, on average, wavenumber two correlates better with levels in the stratosphere than wavenumber one. While the correlation of the wavenumber one amplitude with itself falls to 0.2 around 100 hPa, the same values is reached just below 1 hPa for wavenumber two. The correlation falls off at about the same lags for either wavenumber. However, while the correlations based on wavenumber one are statistically significant at almost all lags and levels, the significant correlations for wavenumber two are mostly confined to the lags between ±7.5 days. As for wavenumber one, the heat fluxes show smaller correlations but the same overall structure as the geopotential wave amplitudes (see Figure 2.3f). We conclude that, statistically, the propagation of wavenumber two into the stratosphere could reasonably be predicted from the troposphere. However, the correlation between the wave amplitude and the zonal mean zonal wind is small and negative in upper troposphere and below 0.2 at all lags and levels above 100 hPa (Figure 2.3e). We believe that the relatively small wavenumber two amplitudes in the stratosphere (see Figures 2.1 and 2.2) are insufficient to affect the polar vortex, which also agrees with Dunn-Sigouin and Shaw (2015) who found that the overall heat flux into the stratosphere is dominated by the wavenumber one component. Thus, wavenumber two flux into the stratosphere might be more predictable than the wavenumber one component, but it is not useful to determine the behavior of the polar vortex and unlikely to cause SSWs outside of resonant amplification.

We are interested to determine the level at which waves one and two reliably propagate into the stratosphere and affect the polar vortex. Thus, we repeat the calculations leading to Figure 2.3 for 100 hPa (Figures 2.4) and 10 hPa (Figure 2.5). At 100 hPa, the wavenumber one amplitude is strongly correlated to itself and the first harmonic of the meridional heat flux throughout most of the stratosphere at lags of about a week in either direction. As expected from the previous discussion the correlation to the mid-troposphere is weak. Most importantly, Figure 2.4b shows negative correlations with the zonal wind at 60°N, confirming the result of Polvani and Waugh (2004) that the heat flux at 100 hPa correlates significantly with the strength of the polar vortex. The signal is largest at the top of the dataset and propagates downward over time. The effect of planetary waves on the zonal mean wind lasts substantially longer than the peak in the heat fluxes or wave amplitudes. In contrast, wave two is still almost uncorrelated to the polar vortex (Figure 2.4e). In fact, the largest correlation indicates a slightly stronger zonal mean wind in the stratosphere at negative lags. Unsurprisingly, the correlations of wave amplitudes and heat fluxes expand throughout the stratosphere and even extend close to the surface. Comparing Figures 2.4b and Figures 2.4e confirms the previous results in this chapter and in Dunn-Sigouin and Shaw (2015) that the stratosphere is usually dominated by wavenumber one.

The correlations based on the 10 hPa level (Figure 2.5) show that the planetary waves in the mid-stratosphere are most strongly correlated to heat fluxes around 30 hPa. As most planetary waves dissipate within the stratosphere, it is not surprising that the vertical propagation of wave one decreases above 10 hPa. Furthermore, the correlation with the zonal wind further confirms the strong influence of wave one on the polar vortex. Figure 2.1 shows that the wavenumber one explains most of the planetary wave amplitude in the polar vortex. Thus,
Figure 2.4: Correlation of the square of geopotential wave amplitude at 100 hPa integrated over the Northern hemisphere (area-weighted) with itself (left column), the zonal mean zonal wind at 60°N (middle) and the meridional heat flux integrated over the Northern hemisphere (area-weighted) (right) from 1000 hPa to 1 hPa and lags from -30 to +30 days. a) and c) show the wavenumber one components; d) and f) show wavenumber two. Shaded values are not statistically significant at the 95% level.
Figure 2.5: Correlation of the square of geopotential wave amplitude at 10 hPa integrated over the Northern hemisphere (area-weighted) with itself (left column), the zonal mean zonal wind at 60°N (middle) and the meridional heat flux integrated over the Northern hemisphere (area-weighted) (right) from 1000 hPa to 1 hPa and lags from -30 to +30 days. a) and c) show the wavenumber one components; d) and f) show wavenumber two. Shaded values are not statistically significant at the 95% level.
the negative lags in Figure 2.5b likely coincide with less overall wave activity and stronger winds. In contrast, the winds are negatively correlated to wavenumber one at positive lags demonstrating the weakening effect of wavenumber one on the strength of the polar vortex. More surprisingly, the second harmonic of the geopotential is not large enough to affect the strength of the polar vortex even at 10 hPa (Figure 2.5e). It is clear from this analysis that the average state of the vortex is almost entirely explained by wave one despite the strong vertical correlation of wave two amplitudes with itself and the meridional heat flux.

Predicting geopotential height fields in the troposphere is part of operational weather forecasts. Thus, many studies have focused on predicting planetary wave propagation into the stratosphere from phenomena associated with amplified geopotential amplitudes of small wavenumbers. However, we explain in Chapter 1 that the vertical wave propagation is not directly related to the geopotential wave amplitude but to the meridional heat flux. To address the question if predicting heat fluxes in the troposphere is a more promising path to forecast heat fluxes into the stratosphere, we repeat our analysis by correlating the variables against the meridional heat flux instead of the geopotential amplitude at the same levels. The qualitative results remain the same (not shown) and confirm that the results above are not sensitive to the measure of planetary wave activity.

The results in this section may be confusing as we showed that the behavior of the polar vortex is dominated by wavenumber one, but that the wavenumber one amplitude is almost uncorrelated to the amplitude in the stratosphere. A partial explanation can be found in Figure 2.6, where we repeat the calculations above for regressions instead of correlations. The regression coefficients are scaled by the standard deviation of the time series at any given level, so the results are no longer between -1 and 1. Figure 2.6 should be compared to plots a) and d) in Figures 2.3, 2.4 and 2.5. While the overall structure of correlation and regression coefficients is similar, there are important differences. In particular, Figure 2.6a shows regression coefficients above 2 throughout most of the stratosphere and reaches values above 6 above 10 hPa. This result helps to interpret the previous results. Wavenumber one amplitudes are large in the stratosphere and we know that waves propagate from the troposphere into the stratosphere. The small correlation coefficients in Figure 2.3a is therefore the result of the large standard deviation of the wave one amplitude at this level. However, understanding the propagation of wave one remains complicated, as the regression coefficients decrease between 500 hPa and the lower stratosphere. In fact, the regression coefficients are negative for small negative lags between about 100 hPa and 20 hPa, which indicates the importance of upper levels in the propagation of wavenumber one. At 100 hPa (Figure 2.6c), wavenumber one amplitudes show positive regression coefficients up to 4 throughout the stratosphere for lags up to about 10 days. The negative coefficients for larger lags are probably the result of reduced planetary wave activity in response to the zonal wind response in Figure 2.4b. As in the previous analysis, the amplitudes in the lower stratosphere are almost unconnected to the troposphere. The regression based on 10 hPa peaks at higher levels than the correlation (compare Figures 2.5d and 2.6e), which is most likely due to large variability at the highest levels.
Figure 2.6: Dimensionless regression coefficient for the regression of the square of geopotential wave amplitude at 500 hPa (top), 100 hPa (middle) and 10 hPa (bottom) integrated over the Northern hemisphere (area-weighted) with itself between 1000 hPa to 1 hPa and lags from -30 to +30 days. The left (right) column shows the regression coefficients based on wavenumber one (two). Note the non-linear contour intervals.
The regressions for wavenumber two mostly confirm the results based on the correlation coefficients. The regression coefficients at 500 hPa (Figure 2.6b) show a strong connection to the stratosphere, which appears unaffected by the tropopause. The negative values in the stratosphere at lags of about ±15 days further emphasizes the shorter timescale of wave two propagation into the stratosphere. The positive coefficients at all lags and almost all stratospheric levels in Figure 2.6d illustrate that, on average, wave two amplitudes are too small to affect the polar vortex and therefore future wave propagation. The largest difference in the comparison of correlations and regression of wavenumber two amplitudes appears at 10 hPa (compare Figures 2.5d and 2.6d). The correlation coefficients reach statistically significant values above 0.4 in the troposphere while the regression coefficients are below 0.1 in the troposphere at all lags. This result suggests that, due to the standard deviation of the compared time series, it is not trivial to track wave events in the stratosphere to the troposphere even if the regression of lower levels to higher levels is large and positive.

In summary, we built a climatology of MERRA data from October to April and from 1979 to 2014 to quantify the average propagation of planetary waves from the troposphere into the stratosphere. We show that wavenumber one becomes more dominant with height and latitude. However, the relative importance of wavenumber one is subject to large inter-annual variability. Our results indicate that the waves one and two propagate fundamentally different into the stratosphere and have different effects on the polar vortex. Wave one amplitudes above the tropopause correlate well with the strength of the polar vortex. However, the wave one amplitude in the stratosphere is almost uncorrelated to the mid-troposphere. The picture is further complicated by our analysis of regression coefficients, which show the strong influence of tropospheric wave one amplitudes on the stratosphere, but further emphasize the effect of the tropopause on wavenumber one. This study can help to understand the limits of predicting stratospheric processes from tropospheric levels. In contrast, wavenumber two propagates readily from the troposphere into the stratosphere without affecting the polar vortex. As a result, identifying tropospheric processes that lead to enhanced planetary wave amplitudes is insufficient to predict stratospheric dynamics. In particular, the zonal mean zonal wind in the stratosphere is almost uncorrelated to the amplitude of geopotential or meridional heat flux of either wavenumber.

Two major question emerge from this work. First, which processes control the propagation of wavenumber one from the troposphere into the stratosphere at levels between 300 hPa and 100 hPa? Both, Polvani and Waugh (2004) and Birner and Albers (2017), identify the layer below the tropopause as crucial to understand the planetary wave activity in the stratosphere as well. Our work shows that this “communication layer” is particularly important to the propagation of wave one disturbances. Given the same background conditions, the Charney-Drazin theorem (Charney and Stern 1962) predicts wavenumber one to propagate higher into the stratosphere than wavenumber two. While the overall heat flux in the stratosphere is indeed dominated by wavenumber one, it is unclear what leads to the strong decay of wavenumber one around the tropopause whereas the correlations of wavenumber two are seemingly unaffected by this barrier. Furthermore, the difference between regression and correlation coefficients illustrates that the standard deviation of wave amplitudes
strongly depends on height. Tripathi et al. (2015) suggests that models might be unable to answer this questions as they do not accurately capture the growth of wave two within the stratosphere leading up to splits.

Second, the presented results are unable to explain the existence of polar vortex splits. As wave two has almost no average influence on the polar vortex, splits are likely the result of special circumstances that lead to the growth of wavenumber two for a short period of time. We will revisit these questions in the second part of this chapter and in Chapter 3.

**Extreme Heat Flux Events**

The climatology of planetary waves in the previous section can not answer the questions under which circumstances wave two grows in the stratosphere and what allows wavenumber one to propagate from the troposphere to 100 hPa. In this section, we build composites for large wave one and two heat fluxes at 100 hPa respectively to address these issues. We choose this level for easy comparison with Polvani and Waugh (2004) and because we show in the previous section that both waves propagate through most of the stratosphere from this height. Other studies have mostly used levels at or above 100 hPa Dunn-Sigouin and Shaw (2015); Watt-Meyer and Kushner (2015). We define events based on two standard deviations of the time series of the meridional heat flux integrated over the Northern hemisphere for the first two harmonics. There is no general agreement if heat fluxes have to be sustained over a certain time to affect the polar vortex. Polvani and Waugh (2004) used a 40-day running mean in their heat flux composites, while Garfinkel et al. (2010) averaged the 100 days with the largest anomalous heat fluxes. Sjoberg and Birner (2014) found in data and models that fluxes between 10 and 20 days are most likely to affect the polar vortex. Figure 2.7 shows the detected events for both wavenumbers as a function of the length of the running mean. It is noteworthy that the number of extremes declines quickly with increasing running mean length despite the fact that the standard deviations are calculated after applying the running mean. Thus, the variability of the meridional heat flux is largest at short time scales. We decide to define events based on a 30-day running mean. Beyond this threshold, the detected events remain almost unchanged, which indicates that we capture large sustained heat fluxes at almost all measures. Furthermore, we avoid including marginal events that are visible only at specific running mean values. We are confident that wavenumber one heat fluxes, sustained over this time period, have a significant effect on the polar vortex. Watt-Meyer and Kushner (2015) showed that the heat fluxes are dominated by linear interference between the climatology and anomalies from the climatology. Thus, composites anomalies in this chapter are calculated by subtracting the multi-year winter average from individual events before averaging over all events of a given wavenumber.

The large variability of meridional heat fluxes on short time scales can be seen in the composite of heat flux anomalies in Figure 2.8. It shows the heat flux anomaly of the same harmonic, which was used to detect the events. We find 19 events based on wavenumber one and 13 based on wavenumber two. Day 0 is defined as the first day the 30 day running
mean at 100 hPa exceeds two standard deviations. The composites for both wavenumbers are dominated by large bursts at the end of the averaging period. As expected from the results above, the heat flux anomalies for both waves are positive at all stratospheric levels and the wavenumber two anomalies extend lower into the troposphere at the time of the largest anomalies. However, the positive wave one anomalies seem to be connected to slightly positive heat fluxes at the surface at lags as much as 20 days before the largest anomalies at 100 hPa. A small correlation between wave amplitudes at 100 hPa and heat fluxes is also apparent in the climatology (Figure 2.4c). In general, the wave one heat flux anomalies develop over a larger period of time while the wave two events appear to capture two separate bursts. We speculate that peaks in wave two are shorter than in wave one and therefore result in fewer extremes when averaging over 30 days (also compare Figures 2.4c and 2.4f). In the future, it might be worth exploring different running means for the extremes in waves one and two respectively.

It is well known that waves one and two tend to be anti-correlated in the stratosphere (Labitzke 1981). Figure 2.9 shows the heat flux anomalies of the wavenumber that was not used to define the events. While it is not clear from wave propagation theory that one wave amplitude has to be small during extremes of the other wavenumber, we see negative anomalies at almost all lags and levels. Interestingly, we find a large, short burst of wave one heat flux centered at 10 hPa and just before day zero of the wave two composite (Figure 2.9b). The
Figure 2.8: a) Composite of wavenumber one heat flux anomalies of events based on the wavenumber one heat flux component as a function of time and height. Day 0 represents the first day the centered running mean exceeds 2 standard deviations at 100 hPa. Anomalies are integrated over the Northern hemisphere and given in standard deviations with respect to the climatology. b) as a) but for wavenumber two component of wavenumber two events.
wave one signal is located just above the first of the two bursts in Figure 2.8b. However, it is unlikely that the weak ($\leq 1std$) wave two anomaly could cause the strong wave one signal. On the other hand, it is unclear if the wave one peak in Figure 2.9b is related to the stronger burst in wave two heat fluxes around day 13. The wave one signal appears to be confined to the mid-stratosphere, while the largest wave two heat fluxes are found at lower levels with large, positive anomalies reaching the surface. We will revisit the question of wave-wave interactions during extreme events in Chapter 3.

We now show the influence of the heat flux anomalies on the strength of zonal mean zonal wind at 60°N (Figure 2.10). Figure 2.10a shows the downward propagation of weak zonal mean winds following day zero of the wave one composite. The signal reaches the surface around day 15. At negative lags, the zonal mean wind is slightly stronger than normal. Unsurprisingly, this result is very similar to Polvani and Waugh (2004) who did not distinguish between wave one and two heat fluxes (see their Figure 4 and note that the NAM signal is of opposite sign compared to the zonal mean wind). It also agrees with the results of Dunn-Sigouin and Shaw (2015) and the results above that the wave one component dominates the heat flux in the stratosphere and strongly affects the polar vortex. The zonal mean

36
Figure 2.10: a) Composite of zonal mean zonal wind anomalies at 60°N for events based on the wavenumber one heat flux component as a function of time and height. Day 0 represents the first day the centered running mean exceeds 2 standard deviations at 100 hPa. Anomalies are given in standard deviations with respect to the climatology. b) as a) but for wavenumber two events.

Wind anomalies of the wave two composite are smaller. However, Figure 2.10b shows that the largest wave two heat fluxes can affect the polar vortex in the stratosphere despite the small correlation in Figure 2.4e. It is worth noting that the zonal wind anomaly changes sign later for wavenumber two than for wavenumber one. The signal aligns almost perfectly with the maximum heat flux in Figure 2.8b and the change in sign occurs almost simultaneously throughout the stratosphere. This behavior agrees with many descriptions of stratospheric splits or events based on wavenumber two amplitudes that identify a sudden almost barotropic onset of these events (Matthewman et al. 2009; Esler and Scott 2005).

We have shown that the defined heat flux events affect the polar vortex and even the zonal mean wind at the surface. We investigate the geopotential patterns in three dimensions over the course of the 60 day composites to identify when and where coherent structures develop that can help explain how planetary waves reach the 100 hPa level and ultimately the stratosphere. Figure 2.11 shows the geopotential anomalies associated with the wave composite at 6 different levels and averaged over days 11-15, when the flux is largest. The figures of the other timesteps can be found in the Appendix A. Note that the color bar is adjusted for each level. Figure 2.11 shows that there are statistically significant signatures (shaded white) in the geopotential fields at all levels from 850 hPa to 10 hPa. In particu-
Figure 2.11: Geopotential anomalies for the 5-day average around day 13 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White hatching indicates statistical significance at the 95% level. Note that the color bar is different for each level.

The anomalies are negative above eastern Asia up to 50 hPa. The negative anomalies tilt from Western Europe at low levels to West Canada and the Aleutians in the stratosphere. The orientation of the anomalies compares well to the results of Matthewman et al. (2009) and Garfinkel et al. (2010) despite their event definitions being based on the state of the vortex and daily mean heat fluxes respectively. Thus, we are confident that our results contribute to the understanding of SSW displacements and are not particularly sensitive to the chosen running mean time scale. Statistically significant features begin to develop in the troposphere almost four weeks earlier at the beginning of the period covered by the running mean, where we see slightly positive heat fluxes in Figure 2.8a. At 10 hPa, positive anomalies, indicating a weakening of the vortex, appear around day zero. After day 15, the anomalies in the stratosphere are positive everywhere until the end of the composite. In the troposphere, the anomalies of both signs become weaker and less organized.
The geopotential anomalies associated with the composite of wave two events are shown in Figure 2.12. As for wavenumber one, we detect statistically significant signatures at all levels. Most prominently, we find positive anomalies over the Aleutian Islands at all levels between 500 hPa to 10 hPa. This region has been found to be prominent in the development of SSWs (Labitzke 1977; Harvey and Hitchman 1996). The rest of the Western hemisphere is characterized by a complex tripole structure with negative anomalies above the North Atlantic and positive anomalies around the North pole and South of about 40°N. Finally, negative anomalies appear over large parts of Asia in the troposphere. In the stratosphere, a wavenumber two pattern emerges, with positive anomalies over the Aleutian Islands and Europe, and negative anomalies over Canada and a weaker one over Siberia. Similar vortex orientations were found in Matthewman et al. (2009) and Garfinkel et al. (2010). The anomalies of wavenumber two events tilt less with height than the wavenumber one events suggesting a more barotropic structure consistent with polar vortex splits. Our composite shows positive anomalies over the Aleutians at least 4 weeks before the peak of the event, but they are not statistically significant at the 95%-level until about 5 days prior to the largest heat fluxes at 100 hPa. A larger dataset or an event definition that results in a larger number of events could address if the geopotential anomaly consistently changes this early. The second consistent feature during the development of the largest wave two heat fluxes are the negative geopotential anomalies over the North Atlantic. We find smaller scale features to be statistically significant weeks before day 13, but the anomalies are small and the location changes over time. These signature may be the result of the relatively small sample size and are also not found in the previously cited composite analyses. In accordance with the small surface impact of wave two events in Figure 2.10b, the tropospheric anomalies after day 13 become small and are not statistically significant. In the stratosphere, the pattern develops from a quad-pole to a dipole pattern following the largest heat fluxes. It is not surprising that the weakened vortex is characterized by positive anomalies at the pole and over most of the Eastern hemisphere. However, in contrast to the wavenumber one composite, we find significant negative anomalies over North America for at least two weeks following the event. This result has two implications. First, the largest wave two heat fluxes are short as the prominence of wavenumber one reemerges quickly after the wave two bursts. Second, the effect of wave two on the polar vortex is less pronounced compared to wavenumber one where all anomalies are positive following the peak heat flux.

We now analyze the wave amplitudes of the geopotential anomalies in Figures 2.11 and 2.12. Figure 2.13 depicts the wavenumber one amplitudes of the geopotential anomalies at 60°N for each of the 19 wavenumber one events (colored lines) between the surface and 1 hPa. The black solid line represents the average over all events and the black dashed line indicates the spread of one standard deviation in either direction. The overall increase of amplitudes with height is expected as the amplitude of planetary waves increases with decreasing density. Comparing the amplitudes for the three different times it becomes clear that wavenumber one heat fluxes are long-lasting events. The amplitudes are substantial for all events during most of the investigated period of time. In particular, the geopotential anomalies are of comparable magnitude even four weeks before the largest heat fluxes at
Figure 2.12: Geopotential anomalies for the 5-day average around day 13 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White hatching indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Anomalous geopotential amplitudes of wavenumber one at 60°N as a function of height for wavenumber one events for 5-day averages centered on days -13, 13, 27. Each colored line represents one event. The black solid line represents the average over all events and the black dashed lines show the spread (± 1 standard deviation).

100 hPa. The recovery phase of SSWs is usually characterized by small horizontal potential vorticity gradients and smaller waves are found to be important in the tropospheric response following vortex events (Limpasuvan et al. 2004). However, the amplitudes in our composite are almost as large on day 27 as on day 13 for most events and at most levels. This result indicates that heat flux events are not necessarily followed by the reduced stratospheric planetary wave activity that is generally observed after the SSWs. The spread among the events is relatively small but increases with height.

The small variability among events and the long-lasting nature of wave one heat fluxes become apparent when Figure 2.13 is compared to Figure 2.14. Before day 13, the wavenumber two amplitudes are substantially smaller in the stratosphere than for wavenumber one events. In the troposphere, we find few events with amplitudes as large as the wavenumber one counterparts on day -13. In accordance with the short heat flux bursts in Figure 2.8b, the amplitudes increase quickly around day 13 and decrease more rapidly afterwards compared to Figure 2.13. The most prominent difference between the two composites is the spread among events. While wave one events show little variation below 50 hPa, the wavenumber two amplitudes are much less consistent. In particular, we find events with very small amplitudes at most levels. We believe that the large spread is less indicative of the varying intensity of wavenumber two events, but represents the varying timing of the short wave two bursts. The relatively long running mean of 30 days picks up wave two heat fluxes that peak at different times during the investigated 60-day period. Conversely, the typical development of wavenumber one heat fluxes lasts longer, so the timing aligns.

Finally, Figures 2.15 and 2.16 show the phases associated with the wave amplitudes discussed above. The phase is defined as the longitude of the first maximum of a given wavenumber. We plot the phases on a scale from -360° to +360° to avoid arbitrary jumps when
Figure 2.14: Anomalous geopotential amplitudes of wavenumber two at 60°N as a function of height for wavenumber two events for 5-day averages centered on days -13, 13, 27. Each colored line represents one event. The black solid line represents the average over all events and the black dashed lines show the spread (±1 standard deviation).

phases oscillate around 0°. We subtract 360°(180°) of a given wavenumber one (two) phase if it reduces the difference between levels. Wavenumber one tilts westward with height for most events and at most times. However, the phases of the 19 events are almost randomly distributed on days -13 and 27. Conversely, the phases align at the time of the largest heat fluxes. As a result, the spread among events is substantially reduced around day 13. We have no physical explanation for the curious alignment of many events at day 27 and 1 hPa. As the quality of reanalysis data is worse above 10 hPa (Manney et al. 2003), this result should be interpreted cautiously.

Figure 2.14 shows surprisingly similar evolutions for the wave two events. There is little structure to the phase profiles before and after the largest heat fluxes, but the events are closely aligned during the peak of the composite. While heat flux events are different from events based on the state of the polar vortex, it is useful to compare our results to composite of vortex splits. Split events are often described as barotropic Esler and Scott (2005) and we see little westward tilt on days -13 and barotropic structures on day 27 for the wavenumber two events. However, we emphasize that the largest wave two heat fluxes are associated with substantial westward tilt among all events. This result is similar to the splits composited in Matthewman et al. (2009). Plumb (2010) showed that it is difficult to distinguish between non-linear resonance of the barotropic mode and resonance based on wave reflection. Thus, the westward tilt at the time of the largest heat fluxes alone can not exclude barotropic mode resonance as the cause for the detected extremes. As discussed above, the effect of wavenumber two heat fluxes on the polar vortex is relatively small.

The bold lines in Figures 2.15 and 2.16 show the average over the climatological phases at the time of the events. The climatology aligns closely with the vents at the time of the largest heat fluxes. This result agrees with the work of Watt-Meyer and Kushner (2015) who found
that large heat fluxes are mainly the result of linear interference between the climatological mean and anomalies caused by standing waves. Our analysis emphasizes the timing of this effect as the phases only align for a few days during the composite and separates extremes by wavenumber.

Conclusions

We have calculated a climatology of planetary wave activity in the Northern hemisphere winter from 1979 to 2014. While theory predicts that wavenumber one is more important with increasing height and latitude, quantifying these effects provides important context to the anomalies and event composites of previous studies. Additionally, we find that the relative importance of wavenumber one and two at different levels and latitudes varies.
strongly from year to year. As a result, studies that composite days above a certain threshold (Garfinkel et al. 2010) might only sample a few winter seasons. Using the same data, we calculate correlations between wave amplitudes, meridional heat fluxes and the zonal mean zonal wind at 60°N, at various levels. Our results show that neither the wavenumber one amplitude nor heat flux in the troposphere are well correlated to the same variables at 100 hPa and above. Consequently, they also have little correlation to the climatological state of the polar vortex. However, the regression coefficient between the tropospheric and stratospheric wave one amplitudes are large. Additionally, we find that wavenumber one correlations and regressions are small around the tropopause. This result provides a possible explanation why previous studies have not been able to link tropospheric processes (such as blocking or low pressure systems) to heat fluxes into the stratosphere convincingly. We believe more promising approaches might have to rely on variables that are directly linked to the propagation of planetary waves. Dunn-Sigouin and Shaw (2015) found vertical wind shear to be a key ingredient to explain large heat fluxes. Domeisen et al. (2017) explored the idea that the phase speed of tropospheric disturbances might predict their stratospheric impact. As the background flow and the phase speed are directly related to the Charney-Drazin theorem, these new approaches might help predict the stratospheric impact of tropospheric events and address the question why similar events do not always reach the stratosphere.

The correlations for wavenumber two suggest that the climatological, vertical propagation of the two wavenumbers can not be simply described by the Charney-Drazin theorem. The theory does not predict the stronger correlations between the tropospheric and stratospheric wave amplitudes for wavenumber two compared with wavenumber one. However, the regression analysis reflects the stronger influence of wavenumber one on the polar vortex. The difference between the types of analysis emphasizes the large standard deviation of stratospheric wave amplitudes. Furthermore, the potential predictability of wave two propagation is of limited use as we find no correlation between wave two activity and the state of the polar vortex. This result confirms that the behavior of the vortex is explained by wavenumber one at almost all times. Following the idea of Domeisen et al. (2017), it is possible that the average phase speed of wave two disturbances varies from those of wave one. Future work is necessary to explain the distinct propagation of wavenumbers one and two from the troposphere into the stratosphere.

In part two of this chapter, we investigate composites of large heat flux anomalies of wavenumber one and two. Our results on extreme heat fluxes at 100 hPa show similar connections between heat fluxes and geopotential anomalies as Polvani and Waugh (2004); Garfinkel et al. (2010). Furthermore, we find that wave two heat fluxes have shorter characteristic time scales, which might help explain the abrupt onset of sudden stratospheric splits reported by Matthewman et al. (2009). Our analysis of geopotential anomalies results in statistically significant patterns weeks before the largest heat fluxes at almost all levels. Cohen and Jones (2012) also found tropospheric precursors to SSWs weeks before the central date. As in our study, the signal is stronger for wave one events. However, it is unclear if there is real predictive value in these patterns. Further studies using different methods and larger
sample sizes will be necessary to determine the robustness of this signal.

The phase profiles at different times illustrate the development of extreme heat fluxes. First, large heat fluxes are associated with a remarkably consistent phase-height relationship throughout the atmosphere. As the anomalies align with the climatology, our results agree with Watt-Meyer and Kushner (2015) who showed that linear interference is responsible for a large fraction of heat fluxes into the stratosphere. Thus, one should look at extreme heat fluxes as an emphasized mean state rather than the appearance of a distinct wave pattern. Our work expands on the previous results in multiple ways. First, we show that extreme heat fluxes for both wavenumbers are likely governed by the same mechanism despite their different climatological effect on the stratosphere. We further show that large heat fluxes align at many levels simultaneously suggesting that predictions might require data from more than one level. Additionally, the development of the phase profiles over time suggests that the phases of the anomaly and the climatology only align for a few days while the largest heat fluxes occur. This is true even if the wave amplitudes are large over a longer period of time. Thus, extreme heat fluxes appear to be more sensitive to the timing of the phase relationship between anomaly and climatology than the magnitude of the geopotential amplitudes. The weak correlations of planetary wave one amplitudes between the troposphere and the stratosphere agree with the results of Polvani and Waugh (2004) and Birner and Albers (2017) that the tropospheric signal declines sharply between 300 hPa and 100 hPa. Based on the analysis of large wave one heat fluxes, we speculate that the propagation of wave one into the stratosphere might be less dependent on a particular process in the upper troposphere but simply related to a misalignment of the phase of the average and the anomaly. From this perspective, the more readily propagating wave two would be explained by the fact that climatology and anomaly are more likely to sync in the upper troposphere. Our research does not address the question why climatology and anomaly align at certain times and levels. More research is needed to support or disprove this theory.

In our study, extreme wavenumber two events are mostly barotropic outside of the time of the largest heat fluxes. However, during the peak of the event, almost all events show a clear westward tilt with height. A similar observation was found by Matthewman et al. (2009) during split events, but Esler and Scott (2005) have shown that barotropic resonance might be responsible for vortex splits. Plumb (2010) showed that westward tilt alone is not sufficient to distinguish between resonance of the barotropic mode and resonance based on wave reflection. Our work underlines the fact that barotropic structures are not necessary to produce large wave two heat fluxes from the troposphere into the stratosphere.

Finally, we observe a short but large wave one anomaly just before day zero of the wave two composite. This result raises questions about the importance of wave interactions within the stratosphere which is the focus of the following chapter.
Chapter 3

The effect of wave-wave interactions on polar vortex splits

Introduction

In the previous chapter, we focus on the growth of planetary waves and the associated heat flux. The work is motivated by the connection between large heat fluxes and the state of the polar vortex Polvani and Waugh (2004). This Chapter reverses the perspective in the sense that we detect extreme events based on the state of the polar vortex and then explore what caused the SSWs. The general features of SSWs are described in Chapter 1.

The study of SSWs is closely tied to an ongoing debate on how to properly define them (see Butler et al. (2015) for a detailed description of the criteria used now and in the past). One of the more popular definitions uses the sign of the zonal mean zonal wind at 60°N and was formalized by Charlton and Polvani (2007). Alternatives include, but are not limited to, algorithms based on the analysis of empirical orthogonal functions of various variables on different levels (Baldwin and Dunkerton 2001; Gerber and Polvani 2009; Hitchcock and Shepherd 2013), the statistical moments of the 2-D potential vorticity field (Mitchell et al. 2011) and k-mean cluster analysis (Coughlin and Gray 2009). The differences between definitions is not just a semantic one. Criteria based on temperature or zonal mean wind can more easily be applied to relatively sparse observations. However, algorithms based on the moments of potential vorticity may be more readily applied to PV dynamics. Additionally, definitions that rely on a fixed latitude might be useful to interpret reanalysis data, but are difficult to adapt to models where the size of the polar vortex may vary from the observed climatology. Butler et al. (2015) show that the number and the timing of detected events varies significantly between different definitions (see their Figure 2). The issue of inconsistent criteria also applies to the distinction of splits and displacements. Algorithms categorizing SSWs have been based on geopotential height amplitudes (Yoden et al. 1999),
PV gradients (Charlton and Polvani 2007), the moments of PV (Mitchell et al. 2011) and the distribution of geopotential height at 10 hPa (Miller et al. 2013). In summary, there is no consensus on how to define or categorize SSWs. This issue is emphasized by the relatively small sample size of about 25 events between 1979 and 2017 depending on the definition and the dataset. The implications are twofold. First, a small sample makes it harder to test or train algorithms to optimize parameter choices. Second, if two algorithms disagree on only a few events, it can have a substantial impact on the composite analysis. As a consequence it can be difficult to compare composites of SSWs. In this study, we use the method of Charlton and Polvani (2007) to detect SSWs and the algorithm of Miller et al. (2013) to distinguish splits and displacements.

Despite the uncertainty surrounding the definition and categorization of SSWs, the majority of publications agrees that different mechanisms are responsible for the splits and displacements. In Chapter 2.2, we show that, climatologically, the correlation between wave two amplitudes and the zonal mean wind in the stratosphere is small. Furthermore, we find in Chapter 2.3 that the wave one and two components of the meridional heat flux are anti-correlated in the stratosphere during extreme heat flux events. In particular, there is a distinct spike in the wavenumber one heat flux just before day 0 of the wave two composite (Figure 2.9b). Labitzke (1977) was among the first to note a similar behavior before the onset of stratospheric splits. Similarly, Bancala et al. (2012) found that a substantial fraction of splits are preceded by large anomalies in the wavenumber one heat flux at 100 hPa. However, it remains unclear if the large wave one heat fluxes are directly responsible for the sudden increase in wave two heat fluxes or if wave one preconditions the vortex for wave two growth. The process of preconditioning was described by Plumb (1981) who suggested that wave breaking at the edge of the polar vortex could lead to a resonant state even if the forcing itself was slightly off-resonant. Matthewman and Esler (2011) later expanded this idea to more realistic models of SSWs. On the other hand, one can imagine that the breaking of wavenumber one in the stratosphere converts energy towards larger wavenumbers. This process is also called wave-wave (or eddy-eddy) interaction as energy is fluxed from one wavenumber to another. Dunkerton et al. (1981) and references therein found that models pointed at the latter process as the probable cause for wave two growth during the 1979 event. However, they did not quantify the effect of wave-wave interactions in the stratosphere and both data and models were limited compared to modern reanalysis datasets. More recently, Scott and Dritschel (2006) studied vortex-vortex interactions in a highly idealized settings and found qualitative similarities with observed SSWs. Wave-wave interactions can occur independently of resonant behavior. As there is no accepted method to test observations or reanalysis data for resonance directly, quantifying the effect of wave-wave interactions may also help to understand the importance of non-linear resonance during splits. To our knowledge, there are no comprehensive studies of the role of wave-wave interactions during SSWs. Since Smith (1983); Smith et al. (1984) used satellite data to quantify wave-wave interactions, the process has not been studied in observations or reanalysis data.
SSWs have been analyzed using a variety of techniques from synoptic descriptions (McInturff 1978), to TEM analysis (Limpasuvan et al. 2004) and annular mode calculations (Baldwin and Dunkerton 2001). Performing a Fourier analysis on geopotential (height) or the upward EP-flux are the most common methods to quantify the effect of different waves on the polar vortex. While these calculations have contributed substantially to our understanding of SSWs, they are also limited in important ways. The geopotential amplitude of a given wave is a good measure of the state of the vortex, but it provides little insight into the underlying processes that shaped the vortex. The analysis of vertical EP-fluxes and their divergence quantifies the upward propagation and the breaking of planetary waves. However, no information about the interaction of waves is obtained by this method. In this study, we use kinetic and potential eddy energy to gain insight into the processes leading to SSWs and the growth of planetary waves. In particular, the equations include the effect of every wave-wave interaction on the energy of a particular wavenumber (details in Chapter 3.2). Few studies have used energetics to study stratospheric dynamics (Chandra 1980) and Plumb (1983) has shown that the results of such analyses may be ambiguous if they are interpreted locally. We avoid these caveats by integrating the eddy energy equations over most of the Northern hemisphere stratosphere.

The previous Chapter and the discussion above focus on the effect of planetary waves on the stratosphere. This perspective has become increasingly popular to study SSWs, because there is a large body of literature with theories that lead to quantifiable results (e.g. Andrews et al. (1987)). However, most theories of wave dynamics assume a limited size of the wave amplitudes and the magnitude of wave breaking. It is not obvious that these conditions are met during SSWs, during which waves grow fast and wave breaking is enhanced. In fluid dynamics, the separation of variables into zonal mean and deviations from that mean is less common. “Vortex dynamics” generally refers to methods that describe the synoptic, fully non-linear behavior of a vortex. Due to the complexity of the troposphere-stratosphere system vortex dynamics approaches are usually harder to quantify and have therefore been less popular in the literature. A description of this approach is Polvani and Plumb (1992) who used a contour dynamics model to study the filamentation and breakup of the stratospheric polar vortex without explicitly separating eddies and mean flow. The results in this chapter can help to gauge the validity of applying wave-dynamics to describe SSWs. If the energy of the polar vortex is dominated by wave-wave interactions within the stratosphere rather than wave propagation from the troposphere into the stratosphere, one might question the application of theories which ignore the interplay of waves. In this chapter, we combine the synoptic analysis of the SSW from January 2013 with energy budgets of wavenumber one and two. The fact that the interpretation of the fully non-linear synoptic development agrees with the results of the energy budget calculations, strengthens the confidence in our conclusions.

In this chapter, we use the reanalysis data described in Chapter 2.1 to explore the wave growth around the central date of split events. In Section 3.2, we derive equations to calculate eddy potential and eddy kinetic energy budgets of the stratosphere which allow us to
Energy budget equations

In this section, we are deriving equations for the eddy-energy budgets of kinetic and potential energy. The focus is to show how they can be used to quantify energy transport between waves of different wavenumbers within the stratosphere. We are using the notation of Plumb (1983) where it is applicable and define the eddy available potential energy (from now on just potential energy) as

$$P_E = \frac{-R}{2p} \left( \frac{p}{p_0} \right) \frac{\partial^2 \theta}{\partial \theta / \partial p}.$$  \hspace{1cm} (3.1)

We derive an equation for the change in potential energy by expanding the total derivative of $\theta'$ and partitioning the terms into the zonal mean ($\overline{\cdot}$) and deviations from the zonal mean ($\cdot'$).

$$\frac{d\theta'}{dt} = \frac{\partial \overline{\theta}}{\partial t} + \frac{\partial \theta'}{\partial t} + \overline{u} \cdot \nabla \theta + \overline{u} \cdot \nabla \theta' + u' \cdot \nabla \theta + u' \cdot \nabla \theta' = \overline{J} + J'$$ \hspace{1cm} (3.2)

Where $J$ represents the sum of all non-conservative processes, such as radiation, which tend to be poorly constrained by observations. Vectors are printed in bold. $u$ is the three-dimensional velocity in pressure coordinates. We eliminate the zonal mean terms by subtracting the equation for the zonal mean

$$\frac{\partial \overline{\theta}}{\partial t} + \overline{u} \cdot \nabla \overline{\theta} + \overline{u} \cdot \nabla \theta' = \overline{J}$$ \hspace{1cm} (3.3)

from equation 3.2. The result is then multiplied by $\theta'$. Finally, taking the zonal mean and multiplying by $\left(-R/2p (p/p_0)^{\kappa}\right)$ leads to an equation for the time derivative of potential energy as the sum of four terms:

$$\frac{dP_E}{dt} = \frac{-R}{2p} \left( \frac{p}{p_0} \right)^{\kappa} \frac{1}{\partial \theta / \partial p} \left[ \begin{array}{c} \text{term 1} \\ \text{term 2} \\ \text{term 3} \\ \text{term 4} \end{array} \right]$$ \hspace{1cm} (3.4)

This derivation neglects the time variations of $\partial \overline{\theta} / \partial p$, which has almost no effect on the results presented in this chapter (not shown). The vertical resolution of MERRA data in the stratosphere might be too small to resolve the changes in stratification. Terms 1 describes the advection of potential energy by the mean flow. Term 2 is the work of the heat flux up
or down the gradient of zonal mean potential temperature. It includes the exchanges with both mean potential energy and eddy kinetic energy. Term 3 is the triple correlation between \( \theta' \), \( \nabla \theta' \) and the three-dimensional flow field. This term describes wave-wave interactions. It is generally neglected as the sum over all wavenumbers and the whole domain should be zero. However, it is crucial to determine the energy transport between different wavenumbers. Quantifying the aforementioned non-conservative processes requires additional assumptions. To first order \( J' \) is assumed to be proportional to \( \theta' \) as shown in term 4. \( \tau \) is a characteristic time scale of the non-conservative processes. In our analysis, including this crude approximation of non-conservative processes with a time scale of about 16 days is necessary to balance the left- and right-hand side of equations 3.4. 16 days is reasonably compared to radiative time scales for the height range between 100hPa and 1hPa (Andrews et al. 1987). However, it is important to keep in mind that term 4 is not an explicit representation of radiation or other non-conservative processes. It is a first order approximation for a number of processes that are not well constrained by observations. In order to describe the effect of terms 1-4 on the growth and decline of wavenumber one and two we perform a Fourier-analysis on equation 3.4.

\[
\left( \frac{dP_E}{dt} \right)_k = \frac{-R}{2p} \left( \frac{p}{p_0} \right)^k \frac{1}{\partial \theta / \partial p} \left[ \underbrace{\text{term 1}}_{1/2} - \underbrace{\text{term 2}}_{\theta'_{\text{zonal}}} - \underbrace{\text{term 3}}_{\theta'_{\text{eddy}}} - \underbrace{\text{damp}}_{1/2 \theta'_{\text{radiation}}} \right]
\]

where \( k \) indicates the k-th wavenumber. As we are calculating zonal wavenumbers, the Fourier-analysis has no effect on the zonal mean variables. Using the Parsifal theorem (Peixoto and Oort 1992), it can be shown that the zonal mean of the product of two Fourier components is only different from zero for the terms of equal wavenumber. For example, the only contributions to the term \( u' \theta'_k \) come from \( u'_l \theta'_l \), \( u'_m \theta'_m \), ..., which is the product of the two Fourier series and should not be confused with Einstein summation. In equation 3.5, term 3 refers to the overall effect of all wave-wave interactions on the potential energy budget of wavenumber \( k \). It does not identify which wavenumbers interact to cause the change in potential energy. It is useful to rewrite term 3 as

\[
\overline{\theta'_{k}}(u' \cdot \nabla \theta')_{-k} = \sum_{l=-\infty}^{\infty} \overline{\theta'_{k}}(u'_{l} \cdot \nabla \theta'_{m})
\]

\( k = l \pm m \). (3.6)

Equation 3.6 can be used to calculate the effect of any particular wave-wave interaction. The potential energy budget of wavenumber \( k \) is affected by the interaction of the wind field of wavenumber \( l \) with the temperature gradient of wavenumber \( m \). As the stratosphere is dominated by wavenumbers one and two, it is feasible to calculate term 3 for all meaningful wave-wave interactions.
The derivation for the eddy kinetic energy

\[ KE = \frac{1}{2} (u'^2 + v'^2) \]  

(3.7)

(from now on just kinetic energy) follows analogous steps to the potential energy budget. Instead of expanding the total derivative of \( \theta' \), the kinetic energy budget is derived from the momentum equation. The time derivative of kinetic energy can then be expressed as the sum of 6 terms

\[
\frac{dKE}{dt} = \underbrace{-u'u' \cdot \nabla u} - \underbrace{\nabla \cdot (u'\Phi')} - \underbrace{\frac{R}{p} \omega k^2} - u'(u' \cdot \nabla u') - v'(u' \cdot \nabla v') - u'L'. 
\]  

(3.8)

\( \Phi \) is the geopotential and \( \omega \) the vertical velocity in pressure coordinates. In term 2, \( u \) represents the horizontal velocities. Term 1 describes the work of the momentum flux against the zonal mean shear. We find that an equivalent term for the meridional velocity is negligible. Energy flux convergence due to waves is represented in term 2. Term 3 is the change of kinetic energy due to rising warm air or sinking cold air. Terms 4 and 5 are two triple correlations that, analogous to term 3 in equation 3.4, describe wave-wave interactions. Term 6 is the non-conservative term, which incorporates all missing processes in one unknown term \( L \). We will neglect this term in all future calculations and discussions because our results show that non-conservative processes are much more important to close the potential energy budget. A Fourier-analysis of equation 3.8 leads to

\[
\left( \frac{dKE}{dt} \right)_k = -u'_k (u' \cdot \nabla u')_k - \nabla \cdot (u'_k \Phi'_k) - \frac{R}{p} \omega'_k T'_{-k} - u'_k (u' \cdot \nabla u')_{-k} - v'_k (u' \cdot \nabla v')_{-k}. 
\]  

(3.9)

As discussed above, it is useful to rewrite the triple correlations as follows

\[
\begin{align*}
    u'_k (u' \cdot \nabla u')_{-k} &= \sum_{l=-\infty}^{\infty} u'_l (u'_l \cdot \nabla u'_m) \\
    v'_k (u' \cdot \nabla v')_{-k} &= \sum_{l=-\infty}^{\infty} v'_l (u'_l \cdot \nabla v'_m)
\end{align*}
\]  

(3.10)

In summary, the equations above describe the potential and kinetic energy of eddies overall and for a given wavenumber \( k \). It is worth noting that no assumption was made that restricts the application of these equations to the stratosphere. However, two complicating factors might limit their value in the troposphere. First, non-conservative processes, especially those related to water, are more prominent in the troposphere. The simple parametrization in Equation 3.4 may have to be replaced with a more explicit representation of the physical
processes affecting the potential energy budget. Similarly, the kinetic energy budget might no longer close sufficiently well without the consideration of non-conservative processes. Second, the application of Equations 3.6 and 3.10 is particularly attractive in the stratosphere where most of the energy is contained by wavenumber one and two (see Chapter 2). Calculating the effect of each wave-wave interaction on the energy of a specific wavenumber becomes increasingly difficult at lower levels, as the energy is more evenly distributed over many different waves. Nonetheless, Boer and Shepherd (1983) showed that studying the role of wave-wave interactions in the troposphere can contribute to our understanding of atmospheric turbulence on macroscopic scales.

Case study

Synoptics

In this section, we analyze the SSW of January 4 2013. As discussed in Section 3.1, it is useful to investigate the development of the polar vortex as a whole before separating the zonal mean and individual wavenumbers. The synoptic description of the vortex is necessarily less quantitative but serves as important context to interpret the results of the following energy budget calculations. A major advantage of this exercise is that no assumptions have to be made and all physical processes and non-linear effects are, by definition, included. Furthermore, the importance of large scale processes in the stratosphere facilitates the visualization and interpretation of the polar vortex during SSWs. Figure 3.1 shows stereopolar maps of geopotential at 10 hPa at 10 day intervals between 20 days before and 10 days after the central date of the SSW. Leading up to the central date, the polar vortex is not centered on the pole but it is displaced toward the Eurasian continent. This observation is in line with a number of studies which have found the polar vortex to be preferentially disturbed by high pressure over the Aleutian Islands (Harvey and Hitchman 1996). The location is also consistent with the positive anomalies in Figures 2.11 and 2.12. Interestingly, the Aleutian high is found in the composites based on either wavenumber and thus might be of limited use to distinguish the growth of wavenumber one and two. The characteristic split of the polar vortex can only be identified at and after the central date. While Figure 3.1 shows only one high pressure system over the Aleutian Islands, the absence of low geopotential values around 0°W is consistent with the positive anomalies in the previous chapter where the climatological mean is removed. Even 10 days before the central date the vortex is seemingly intact despite being smaller and centered further off the pole. However, a distinct area of high geopotential can now be identified over East Asia. This location seems to be a specific feature of this case study as neither our analysis of large heat fluxes nor the composites in the literature identify this region to be particularly important. On the day of the central date, the polar vortex has split into two distinct vortices. Their orientation aligns with the results based on the composite analysis in Chapter 2.3 and the work of Matthewman et al. (2009). Finally, the slow recovery of the polar vortex at 10 hPa is illustrated by the weak vortices
Figure 3.1: Geopotential at 10 hPa at 4 times between 20 days before and 10 days after the central date of the SSW in January 2013.
10 days after the central date and the high geopotential centered on the pole.

Figure 3.2 depicts the absolute temperature distribution for the same times and level as Figure 3.1. 20 days before the SSW, cold temperatures below 190 K are co-located with the center of the vortex. This is to be expected as the stratosphere cools radiatively during winter and mixing across the edge of the polar vortex is small outside of SSWs (Polvani et al. 1994). Over the course of the next 30 days, the correlation of cold temperatures with low geopotential becomes weaker. As the vortex splits into two, the vortex over North America is still associated with cold air, but 10 days after the central date both vortices are at least partly co-located with temperatures above 240 K. While the stratosphere warms rapidly, the horizontal gradients in geopotential and temperature decrease as the vortex edge is disturbed and no longer inhibits mixing. It is important to note that the majority of the warming during SSWs is not due to horizontal advection, but caused by vertical, adiabatic processes (Funke et al. 2010).

To gain more insight into the vertical structure of the SSW, we calculate potential vorticity on three isentropes (PV). Figure 3.3 shows the potential vorticity fields at 465 K, 515 K and 565 K 10 days before, at, and 10 days after the central date. As PV is proportional to the second derivative of geopotential height, small scales are emphasized and the filamentation of the vortex due to planetary wave breaking becomes more apparent compared to Figure 3.1. The most prominent feature of the PV analysis is the similarity between the different levels. The shape of the PV maxima is remarkably constant at all levels even as the vortex splits. Splits have been described as predominately barotropic Charlton and Polvani (2007); Matthewman et al. (2009) and Esler and Scott (2005) showed model results where SSWs resulted mainly from resonant growth of the barotropic mode. From the synoptic analysis of the SSW, we expect wavenumber one to be prominent leading up to the central date. There is no sign of a split vortex 10 days before the central date. Thus the growth of wavenumber two, that is traditionally associated with the split of the polar vortex, should occur during the 10 days before the event is detected. We anticipate that, the reduced differences in geopotential after the central date go along with smaller overall wave amplitudes. As the vortex is still split 10 days after the central date, wave two is presumingly still prominent throughout the stratosphere.

A first step to quantify the influence of wavenumber one and two is to perform a Fourier analysis on the geopotential field. Figure 3.4 shows the planetary wave amplitudes for wavenumber one and two at 10 hPa as a function of time and latitude. The central date of the SSW is indicated by the red line. As expected from Chapter 2.2, the maximum of wavenumber one is slightly North of the maximum in wavenumber two. The displacement of the vortex before the central date in Figures 3.1 and 3.3 is associated with an increase in the amplitude of wavenumber one between 50°N and 60°N. This region aligns with the edge of the vortex where planetary waves break predominately. Before the central date wavenumber one is between 5 and 10 times as large as wavenumber two. The amplitude of wavenumber two becomes significant at the central date. This is because of a slight decrease in the wavenumber one amplitude and sharp increase in wavenumber two. In fact, wavenumber two peaks a few
Figure 3.2: Absolute temperature in K at 10 hPa at 4 times between 20 days before and 10 days after the central date of the SSW in January 2013.
Figure 3.3: Isentropic potential vorticity in PVU at 465 K (a), (b), (c), 515 K (d), (e), (f) and 565 K (g), (h), (i) between 10 days before and 10 days after the central date of the January 2013 SSW. Note that the color bar was adjusted to fit the range of values at each level.
days after the central date of the SSW is detected. This behavior exemplifies the difficulty of characterizing SSWs as splits and displacements. Although it might seem obvious from Figures 3.1 or 3.3 that the event should be referred to as a split, a quantitative analysis of wave amplitudes may not necessarily confirm these subjective impressions. Similarly, the growth and decay of the two wavenumbers does not directly affect the detection of SSWs. We will discuss the implications of this issue in further detail in Section 3.4. Following the central date, both wave amplitudes decrease, which agrees with the reduced gradients 10 days after the central date described above. The slow recovery of the polar vortex manifests itself in small wavenumber one amplitudes even 30 days after the central date.

In summary, the synoptic analysis of the SSW from January 2013 classifies this event as a textbook example of a split compared to findings of case studies (Manney et al. 1994; Andrews et al. 1987) and composites (Charlton et al. 2007). We identify many features, such as the barotropic structure, the high geopotential over the Aleutian Islands and the two vortices over North America and Siberia, that have been reported as typical for this type of event (Charlton and Polvani 2007; Esler and Scott 2005). Most importantly, we find large wavenumber one amplitudes before and the SSW as described by Labitzke (1977) and Bancala et al. (2012). The rest of this chapter will focus on explaining the rapid growth of wavenumber two at stratospheric levels using the energy budget equations derived in Section 3.2.

Energy budgets

In this section, we are quantifying the developments of the polar stratosphere as described above. In particular, we are using the energy equations derived in Section 3.2 to investigate the split of the vortex which is associated with a sudden increase in the amplitude of
wavenumber two (see Figure 3.4). The following figures show time series of the change in potential and kinetic energy between 30 days prior to 30 days after the central date of the SSW. The values are obtained by integrating the equations in Section 3.2 over the entire Northern hemisphere and from 100 hPa to 1 hPa. Thus, if wavenumber two grows in the troposphere and propagates into the stratosphere, it manifests itself in the terms which do not flux energy from one wavenumber into another. In contrast, large wave-wave interactions indicate that the growth of wavenumber two occurs within the stratosphere. The qualitative results are not sensitive to the choice of the intervals in latitude or height (not shown).

The first step in analyzing energy budgets is to ensure that the budget is closed, i.e. that the sum of the individual terms explains most if not all of the actual energy tendencies. As we described in the derivation above this is not an issue for the kinetic energy budget where the sum of terms 1-5 of Equation 3.8 matches the change in kinetic energy calculated from Equation 3.7. However, non-conservative processes are important to balance the potential energy budget. Figure 3.5 depicts the potential energy tendency calculated from Equation 3.1 compared to the sum over terms 1-3 and 1-4 in Equation 3.4 respectively. It is clear that the sum over the three conservative terms overestimates the potential energy substantially. However, including the crude parametrization of non-conservative processes outlined in Section 3.2, balances the budget almost perfectly. We conclude that radiative and chemical processes can not be neglected in the potential energy budget and proceed by including the damping term in any future calculations. The budgets are equally well balanced after separating individual wavenumbers using Fourier analysis (not shown).

Figure 3.6a separates the five terms on the right-hand side of Equation 3.8. As expected, terms 4 and 5 nearly vanish at all times. These terms describe the transformation of energy from one wavenumber to another one but have no net effect on the energy summed over all wavenumbers. The same is true for term 3 in Equation 3.4 (see Figure 3.6b). While we are using the three-dimensional wind field and nabla-operator to calculate the kinetic energies, the longitudinal derivatives of the zonal mean values vanish and the terms including the vertical velocities (except for term 3) do not contribute significantly to the budget (not shown). Thus, terms 1 and 2 are dominated by the derivatives with respect to latitude. We find that the kinetic energy tendency is the result of a rather complicated interplay of terms 1-3. The energy flux convergence (Term 2) is almost always positive and increases significantly around 20 days prior to the central date. At this time, term 2 is dominating the kinetic energy budget and we observe an increase in the amplitude of wavenumber one (see Figure 3.4). Interestingly, this process seems to be unaffected by the SSW - at least integrated over all wavenumbers. In the days leading up to the central date, terms 1 and 3 increase in magnitude, but with opposite signs. As a result, the time derivative of kinetic energy remains large and positive until the central date of the SSW. As the central date approaches the vertical heat flux changes signs and the momentum flux divergence nearly vanishes for the next 30 days. The decrease in term 1 is likely explained by the the collapse of the polar vortex and the associated reduction in shear of the zonal mean zonal wind. After the central date the change in kinetic energy is small as the positive momentum flux divergence and
Figure 3.5: Potential eddy energy tendency budget in J/(kg·s) for the SSW on January 6 2013. Integral over Northern hemisphere and from 100 hPa to 1 hPa of Equation 3.4. The integral of the left hand side is shown in red, the sum over terms 1-3 in purple and the sum over terms 1-4 in blue.

The potential energy budget (Figure 3.6b) is easier to understand. In addition to the vanishing wave-wave interactions, the advection (Term 1) does not contribute significantly to the potential energy budget at any time. The small values are expected as the main contribution to term 1 is the advection by the zonal mean of the meridional flow, which is small when integrated over the hemisphere. Thus, the potential energy budget simplifies to the balance between the positive work against the gradient of zonal mean potential temperature (term 2) and the parameterized damping. The potential energy peaks about 5 days before the central date and declines over the following days to reach its most negative value shortly after the central date. Comparing the sum over all terms in Figures 3.6a and 3.6b shows that the changes in kinetic energy are about twice as large as the changes in potential energy. Furthermore, we see positive values leading up to the central date in both budgets. An increase in eddy energy is consistent with enhanced planetary wave activity in the stratosphere, which is the primary cause for SSWs. While these results help to identify important processes during SSWs, performing a Fourier analysis on the energy budgets is necessary to explain the growth of wavenumber two around the central date.

Figure 3.7 shows the contribution of wavenumbers 1-3 to the overall energy tendencies. The increase in kinetic energy before the central date is almost completely explained by the wavenumber one component. Wavenumber two is negligible before the central date. At the time of the central date, wavenumber one collapses and becomes strongly negative, while wavenumber two increases quickly. The result is a small overall change in kinetic energy at the time of the vortex break up. The energy of neither wavenumber changes substantially
Figure 3.6: a) Kinetic eddy energy tendency budget in J/(kg s) for the SSW on January 6 2013. Integral over Northern hemisphere and from 100 hPa to 1 hPa of Equation 3.8. The blue line represents the sum over all terms on the right hand side, while the green, yellow, black, purple and red line show contributions of the individual terms. b) Analogous to a) but for the potential energy tendency shown in Equation 3.4.

Figure 3.7: a) Kinetic eddy energy tendency budget in J/(kg s) for the SSW on January 6 2013. Integral over Northern hemisphere and from 100 hPa to 1 hPa of Equation 3.8. The blue line represents the total kinetic energy tendency, while the yellow, purple and black line show contributions of wavenumber one, two and three. b) Analogous to a) but for the potential energy tendency shown in Equation 3.4.
after day 5. The small values agree with our synoptic analysis where we found the vortex to be split for at least 10 days after the central date. It is worth remembering that the budget describes the time derivative of energy, so values around zero correspond to small changes in kinetic energy. Around day 20, there is a small and short increase in wavenumber one, which is again balanced by a drop in the wave two component so that the total change in energy remains small.

The potential energy budget (Figure 3.7b) is dominated by wavenumber one which matches the overall energy almost perfectly. The contribution of wavenumber two to the potential energy budget is minimal at all times, but reaches its maximum around the central date. As the tendencies of the potential energy budget are small and because there is no notable development of the wavenumber two component, we focus on the kinetic energy budget for the rest of this chapter. In agreement with the results in Chapter 2.2, wavenumbers larger than two can be neglected in both budgets.

We are left with the question which processes lead to the sudden increase in the kinetic energy budget of wavenumber two. To answer this question, we integrate Equation 3.9, with \( k = 2 \). However, we avoid the complex balance of 5 terms. Instead, the sum over all terms is separated into the sum over terms 1-3 and the sum of terms 4-5. The former sum is labeled “linear terms” as these terms do not transfer energy between wavenumbers. We refer to the latter sum as “nonlinear terms”, because these terms quantify wave-wave interactions, which flux energy from one wave to another. Figure 3.8a shows that the increase in wave two energy is almost exclusively explained by the sum of the two wave-wave interaction terms. In fact, the nonlinear terms are responsible for most changes in the kinetic energy budget over the 60 day period around the central date. The same qualitative result is found to explain the collapse of wavenumber one described above (not shown). The non-linear terms in Figure 3.8a represent the effect of all wave -wave interactions on the energy budget of wavenumber two. The anti-correlation of waves one and two in Figure 3.7a suggests that wave-wave interactions of these waves are responsible for the vortex split. This hypothesis is tested by recalculating the nonlinear terms from Figure 3.8a according to Equation 3.10, which allows us to quantify the effect of any given wave-wave interaction on a specific wavenumber. In our case, we choose \( l = m = 1 \) and \( k = 2 \) to determine the effect of wavenumber one on wavenumber two. Terms 4 and 5 can also be interpreted as the effect of wave one breaking on the wave two budget. It is clear from Figure 3.8b that the energy conversion from wavenumber one toward wavenumber two is the major reason for the increase in wave two energy and therefore the split of the polar vortex. All other wave-wave interactions are negligible in comparison.

We have shown that the split of the polar vortex in January 2013 was most likely caused by wave-wave interactions within the stratosphere and not by an influx of wave two energy from below. We will now try to expand our analysis to the full length of the reanalysis dataset.
Figure 3.8: a) Kinetic eddy energy tendency budget in J/(kg s) for the SSW on January 6 2013. Integral over Northern hemisphere and from 100 hPa to 1 hPa of Equation 3.9 for zonal wavenumber k=2. The blue line represents the sum over all terms on the right hand side, while the yellow combines the contributions of terms 1-3, which do not transfer energy between wavenumbers. The purple line shows the sum of terms 4 and 5 which represent the effect of all wave-wave interactions on the wavenumber 2 budget. b) Integral over same time and volume but for equation 3.10, where k=2 and l=m=1. The blue is identical to a). The red line represents the transfer of kinetic energy from wavenumber one to two.
Figure 3.9: a) Kinetic eddy energy tendency budget in J/(kg s) for the all SSWs between 1979 and 2014. Integral over Northern hemisphere and from 100 hPa to 1 hPa of Equation 3.9 for zonal wavenumber $k=2$. The blue solid line represents the sum over all terms on the right hand side, while the yellow combines the contributions of terms 1-3, which do not transfer energy between wavenumbers. The purple line shows the sum of terms 4 and 5 which represent the effect of all wave-wave interactions on the wavenumber 2 budget. The dashed lines indicate a spread of plus/minus one standard deviation with respect to the solid blue line. b) Analogous to a) but for splits only.
Composites of Sudden Stratospheric Warmings

The results of the previous section present a mechanism to explain the occurrence of split SSWs despite the small correlation between the wave two heat flux at 100 hPa and the state of the polar vortex (see Chapter 2.2). In this section, we build a composite of all split SSWs in the MERRA reanalysis to test how consistently wave-wave interactions are key to explain these events. As described in Section 3.1, one has to make a choice on how to define SSWs. We use the popular criteria defined in Charlton and Polvani (2007), which define the central date of a SSW as the time at which the zonal mean zonal wind at 10 hPa and 60° N changes from westerly to easterly. SSWs have to be separated by at least 20 consecutive days of westerly flow. Additionally, the condition that the zonal wind has to return to positive values for at least 10 days after the central date, excludes final warmings. Splits and displacements are categorized using an algorithm described in Miller et al. (2013), which uses geopotential at 10 hPa at the time of the central date to determine the type of the SSW. If the global minimum (the center of the vortex) is sufficiently connected to every local minimum in geopotential between the pole and 30°N, the SSW will be categorized as a displacement. In total, we detect 24 SSWs - 14 splits and 10 displacements (see table in the Appendix B). We repeat the analysis described in Section 3.3 for each event and compare the averages over all events and over split events only. The composites are centered on the central date of the SSWs, because it is an objective measure of the timing of SSWs and the wave-wave interactions are large at that time in the case of the 2013 warming.

The focus of this study is to quantify the impact of wave-wave interactions, so Figure 3.9 shows the same terms as Figure 3.8a for the composite means. The additional dotted lines represent the spread among the group of events averaged in the plot (+1 standard deviation from the composite mean). Averaging over all SSWs in our dataset leads to very small tendencies in the kinetic energy budget for wavenumber two (Figure 3.9a). Furthermore, the standard deviation of the distribution is large compared to the average. This result is not surprising as the number of events is limited (24) and almost half of the composite (10) is characterized as displacements, which are thought to have small or no signatures in wavenumber two. The effect of separating splits from displacements is depicted in Figure 3.9b. The energy tendency is positive around 10 days before the central date. However, linear and non-linear terms are comparable during split events. The main conclusion from the split composite is that the energy tendencies are small at all times. There is no sign for growth of wavenumber two following the central date and the spread across the 14 events is much larger than the signal at all times. There are two main reasons why the composite results are so different from the case study and not statistically significant. First, the number of splits is very small leading to a very large standard variation. Comparing Figures 3.9a and 3.9b on day -23 exemplifies this issue as the spike in the standard deviations is due to a single event. Figure 3.10 shows the effect of wave-wave interactions on the wavenumber two energy budget for four more SSWs. In all cases, the transfer of energy from wavenumber one to wavenumber two explains most of the signal in the wavenumber two budget. Studying
Figure 3.10: Same as in Figure 3.8b for the SSWs 5, 6, 9 and 19. Events number 5 and 19 are classified as displacements. Number 6 and 9 are classified as splits. The central dates and classifications of all events are listed in Table B.1.
the budget for every SSW, we find that non-linear processes play a substantial role in the
energy budgets of about a third of all SSWs. However, the wave-wave interactions can be
large and negative as well (see Figure 3.10c). The small sample is insufficient to overcome
the large variability between events. In particular, the timing of wave-wave interactions is
very inconsistent as the wave two peak is almost unrelated to the detection of the central
date.
Second, the definition of SSWs and how to categorize them into splits and displacements
strongly affects our composites. Butler et al. (2015) demonstrated the large differences be-
tween various SSW definitions. It is fair to assume that the differences between algorithms
to distinguish splits and displacements is at least as large, because there is no consensus in
the literature how to define the two groups. In fact, two of the events in Figure 3.10 are clas-
sified as splits despite the large signal in the wave two energy budget. Butler et al. (2017)
and Charlton and Polvani (2007) even identify different central dates using one definition
on different reanalysis products. We further explore this issue by comparing the magnitude
of the non-linear terms for all time steps in our extended winter dataset with SSWs from
different definitions and categorizations (not shown). As expected, we find large variability
in the importance of wave-wave interactions in the stratosphere. However, there is no clear
correlation between the magnitude of wave breaking and the occurrence of SSWs. The
tested definitions detect a number of SSWs that are associated with large wave-breaking
events, but wave-wave interactions can be large outside of SSWs and SSWs can occur de-
spite small wave-wave interactions. More importantly, large wave-wave interactions do not
always predict if an event is classified as a split or displacement and the timing of the wave-
wave interactions seems to be uncorrelated to the detection of the central date. As a result,
even averages over splits with large wave-wave interactions lead to small signals. This is-
 sue can be understood from Figures 3.4 and 3.3. The vortex splits at the time the SSW is
detected, but the vortex remains split for at least 10 more days. A slightly earlier central
date could have detected the SSW as wavenumber one was dominant in the stratosphere.
Conversely, the wave-wave interaction could have occurred 10 days before the central date
without changing the categorization of the event. In short, we find that the timing of wave-
number two growth affects our results but is not sufficiently captured by a particular defini-
tion of stratospheric split events. Thus, the presented analysis of energy budgets is useful to
understand the processes during particular SSWs, but they may not easily be expanded to
all events.

Conclusions

In this chapter, we present a new method to analyze processes within the stratosphere during
SSWs. The set of equations in Section 3.2 is not new, but the wave-wave interactions are
often neglected as they do not affect the overall energy budget. In general, energetics have
rarely been used to understand stratospheric dynamics. Chandra (1980) used energetics to
study the thermal structure of the middle atmosphere, but, to our knowledge, Smith (1983);
Smith et al. (1984) are the only studies quantifying wave-wave interactions in the stratosphere explicitly from observations or reanalysis data. Their focus was on the enstrophy budget and relied on a single season of poorly resolved data. We demonstrate that the combination of synoptic analysis and energy budgets can be a powerful tool to identify dominant processes during SSWs and to explain the interplay of waves within the stratosphere.

The synoptic description of the SSW of January 6 2013 identifies the event as a split that develops from a displaced vortex. The orientation of the initial vortex, the two smaller vortices after the central date and the development of high geopotential over the Aleutian Islands, agree with most climatologies of splits (Charlton and Polvani 2007; Matthewman et al. 2009). Furthermore, the growth of wavenumber one preceding the split of the polar vortex agrees with the findings of Labitzke (1977) and Bancala et al. (2012), who analyzed SSWs in observations and models. We find that the potential energy budget can be closed using a simple parametrization of non-conservative processes and a reasonable timescale of 16 days. The energy budgets, integrated over all wavenumbers, reveal a complicated balance of multiple processes over the course of the SSW. However, these processes only affect one wavenumber and do not contribute to the transfer of energy between waves. Our results suggest that the split of the polar vortex can be explained by the transfer of energy from wavenumber one toward wavenumber two within the stratosphere. Upward propagation of wave two from the troposphere into the stratosphere would manifest itself in the linear terms as the non-linear terms only capture the wave-wave interactions above 100 hPa. We interpret our results as a possible answer to the question why we observe splits despite the small effect of wavenumber two on the polar vortex described in Chapter 2.2. Wave-wave interactions within the stratosphere are a plausible explanation for splits in the absence of large upward wave two propagation and fit the observation of Cohen and Jones (2012) that the upward wave activity flux before splits is less connected to the troposphere than leading up to displacements. This theory is independent, but not incompatible with the resonance theories of Plumb (1981); Matthewman and Esler (2011); Esler and Matthewman (2011). Wave breaking of wavenumber one might be able to shape the polar vortex toward a resonant state without transferring substantial energy into wave two as described in Albers and Birner (2014), but the flux of energy from wave one into wave two is not necessarily associated with resonant behavior. In fact Bancala et al. (2012) found splits with and without a preceding wavenumber one signal suggesting that there might be more than one mechanism causing these events. Neither theory relies on large wave two propagation from the troposphere into the stratosphere, which we rarely observe.

In the introduction, we raise the question if theories based on the separation of mean flow and eddies are a good approximation to describe SSWs, which are characterized by substantial wave breaking. We find wave-wave interactions in the stratosphere to be small at most times indicating that analyzing waves independently is a reasonable simplification to describe the stratospheric dynamics from a climatological perspective. However, wave interactions affect the stratosphere; in particular during the extreme events that the field has focused on over the last two or three decades. Our results suggest that quantifying wave-
wave interactions can help to understand the fully non-linear synoptic analysis and serve as important context to theories that ignore this process. We believe that vortex dynamics and wave dynamics can compliment each other in that way.

There are important caveats to our results. The composite of all splits in the dataset do not show the pronounced wave two signal we find for the case study. If splits are caused by more than one mechanism, one would not expect to find a clear signal in the composite. Furthermore, the observed wave-wave interactions are short-lived, so that small differences in the timing of wave two growth, with respect to the central date, can blur the effects in the composite analysis. We find large differences among events and between different definitions for events. The small sample size leads to a non-satisfying signal to noise ratios. In summary, wave-wave interaction may play a larger role in the stratosphere than was previously assumed, but more work is necessary to establish the link between this process and the occurrence of SSWs.

Our results raise questions that future research could address. In Chapter 2.2, we show that wavenumber one usually dominates the development of the Northern hemisphere stratosphere during winter. We also know that planetary waves with small wavenumbers usually dissipate in the stratosphere (Baldwin and Holton 1988). Future research could address the question under which conditions large wave one amplitudes lead to a transfer of energy toward wave two instead of causing displacements of the polar vortex. Furthermore, we see large variability within the SSWs that are categorized as splits. Similarly, Bancala et al. (2012) finds that large wavenumber one amplitudes do not precede all splits. The reanalysis dataset is too short to divide the number of splits into multiple sub-types. However, as the representations of SSWs in GCMs improves, one could contrast splits that are caused by wave breaking with those that are not using model data.
Chapter 4

Shallow-water model simulations of SSWs

Introduction

The previous two chapters are based on reanalysis data, which is produced by complex models, which assimilate many different observations to produce an estimate for the most realistic reconstruction of the atmosphere over the last decades. Most of the variables in this thesis can be observed directly or are well constrained by the dynamical process at the dominant length and time scales in the stratosphere. While these datasets are important and useful, they are necessarily limited to the time during which high quality measurements of the stratosphere exist. As a result, the interpretation of extreme events in stratospheric reanalysis data suffers from small sample sizes. Furthermore, the complexity of the real atmosphere makes it almost impossible to isolate individual mechanisms and test theories on reanalysis data. Thus, there is a long history to use numerical simulations to expand and test our understanding of the climate system in general and the stratosphere in particular.

There are two fundamentally different approaches to modeling the atmosphere.

The most intuitive models incorporate all known and quantifiable processes to produce the most realistic simulations of the atmosphere (or the whole climate system). These Global Circulation Models (GCMs) are powerful tools to describe the current understanding of the climate system to produce estimates of global climate change, and to serve as input for decision makers. However, GCMs are extremely complex and therefore computationally very expensive. As a result, GCMs have only recently become a reasonable option to expand the limited sample size of observations. Additionally, these models are typically too expensive to explore the vast parameter space that controls their behavior. Finally, their complexity makes it almost impossible to isolate cause and effect of a certain phenomenon.

The second class of models simplifies the atmosphere to varying degrees leading to more
cost-effective simulations and more intuitive results. The purpose of these models is not to produce the most realistic representation of the atmosphere, but to provide a toolbox to understand individual mechanisms. Ideally, the insight from studying simplified models can feed back on the development of GCMs and the interpretation of real world data. The range of simplified models is large and usually described as a hierarchy with extremely simple models (e.g. one-box energy models) at the bottom and full-complexity GCMs at the top. The choice of model depends on the investigated questions. In general, it is often useful to employ the simplest model, which captures the processes necessary to investigate a research question. The reduced complexity helps to isolate cause and effect and numerically cheaper models can be used to create larger samples. However, choosing a model makes inherent assumptions about the nature of the problem and possible explanations, since many processes may not be accurately represented in simplified models.

Stratospheric variability has been studied with a variety of models. Matsuno (1971) and Holton and Mass (1976) showed that extremely simple models can produce SSW-like variability. For example, the model in Holton and Mass (1976) is one-dimensional. More complicated models of the stratosphere include PV-contour models (Polvani and Plumb 1989), purely barotropic models (Esler and Scott 2005), stratosphere-only models (Scott and Haynes 2000), simplified three-dimensional models (e.g. dry-dynamical core models) (Gerber and Polvani 2009; Sheshadri et al. 2015) and full GCMs Bancala et al. (2012). The representation of the stratosphere in many full GCMs was not sufficient to study stratospheric dynamics until the last generation of GCMs was published in advance of the most recent IPCC report (Stocker et al. 2013). In response to this issue several high-top GCMs were developed (Schmidt et al. 2006). Typically, these models were characterized by a reduced resolution at lower levels in order to allow the representation of a more realistic stratosphere. Today, these models are most important to study questions that involve processes above the stratosphere.

Until recently, shallow-water models had rarely been used to study the stratosphere (Polvani et al. 1994). This is possibly due to fact that shallow-water models have typically only one layer and a major focus has been to understand the upward propagation of planetary waves into the stratosphere. However, the results of Plumb (1981) and Polvani et al. (1994) (among many others) showed that the response of the polar vortex to planetary wave forcing is complicated even if the vertical propagation problem is ignored. The fact that wavenumber one and two have very different effects on the polar vortex support this idea (see Chapter 2). Recently, Liu and Scott (2015) studied the onset of stratospheric splits using a shallow water model with wave two surface topography. As in previous studies (Holton and Mass 1976; Esler and Scott 2005; Matthewman and Esler 2011), the wave amplitudes grow abruptly for certain vortices and forcing amplitudes indicating that splits are likely the result of resonant growth of planetary wavenumber two. Scott (2016) used the same setup to identify a parameter regime, which leads to constant splitting and merging of the polar vortex.

We are encouraged by these results and choose to employ a similar setup to study the different effects of wavenumber one and two topography on the polar vortex. As we saw in
Chapter 2, wavenumber one and two propagate into the stratosphere under different conditions and have very different average effects on the polar vortex. Additionally, we find that the growth of planetary wavenumber one could lead to subsequent growth of wavenumber two within the stratosphere (Chapter 3). However, the analysis of wave-wave interactions in the stratosphere was complicated by the small number of splits, the inconsistent timing of wave growth with respect to the central date of the SSWs and questions on how to define events. The shallow-water model allows us to run many experiments under idealized conditions to address some of the questions raised in the previous chapters. In particular, we are interested in the following questions:

- Does topographic forcing only lead to SSWs that are dominated by the same wavenumber?
- How do wave amplitudes and polar vortex response depend on the wavenumber of the forcing?
- Can we identify events that are likely caused by wave-wave interactions?

To answer these questions, we will introduce the model equations and the experimental setup (Section 4.2) before we evaluate the model setup using a series of longer simulations (Section 4.3.1). In Section 4.3.2, we discuss a large number of shorter experiments, which explore the interactions of different topographies and vortices. Finally, we summarize and interpret our results in Section 4.4.

**Shallow-Water model**

We use the one-layer shallow-water model developed in NOAA's Flexible Model System (FMS	extsuperscript{1}). The spectral model is solved on spherical coordinates. In our setup, it has 256 longitudes and 128 latitudes resulting in a T85 resolution or about 1.4°. While the results in Section 4.3.1 were calculated with a time step of 600s, we use 300s in Section 4.3.2 to ensure numeric stability of calculations with large forcing amplitudes or very strong polar vorticies. The figures are based on daily averages. The model solves the following set of equations (notation following Polvani et al. (1994)):

\[
\begin{align*}
\zeta_t &= -\nabla \cdot (v \zeta_a) \\
\delta_t &= -\frac{1}{2} \nabla^2 (v \cdot v) + k \cdot \nabla \times (v \zeta_a) - g \nabla^2 (h + h_b) \\
h_t &= -\nabla \cdot (hv) - \tau_E^{-1} (h - h_{eq}).
\end{align*}
\]

where \(\zeta, \delta\) and \(h\) are the vorticity, divergence and layer thickness. \(v = (u, v)\) is the two dimensional velocity field. \(\zeta_a\) is the absolute vorticity (sum of Coriolis parameter and relative vorticity). Note that the product of \(h\) with the gravitational acceleration \(g\) results in the

\(^{1}\)For further information visit http://www.gfdl.noaa.gov/fms
geopotential. The steady state solution is determined by Newtonian relaxation to the equilibrium height $h_{eq}$, which is a simple approximation of thermal relaxation. Additionally, the model includes a friction term in the momentum equation.

\[
\frac{\partial v}{\partial t} \sim \frac{v}{\tau_{fric}} \\
\frac{\partial h}{\partial t} \sim \frac{h}{\tau_{thermal}}
\]  

(4.4)

We find that the friction term is not needed to stabilize the model numerically. As it is challenging to derive a physically meaningful form of the friction term in a shallow-water model representing the stratosphere, we disable the friction equation. We choose an equilibrium height profile, which falls off towards the North pole, to represent a simplified polar vortex. We ignore zonal asymmetries of the polar vortex. To avoid flow across the pole, we use a profile with no latitudinal gradient at the pole. We add a term representing solid body rotation to establish easterly flow at low latitudes. The equilibrium height profile is given by

\[
\text{for } \varphi > h_{width} :

h_{eq} = h_0 - h_{amp} \sin^2 \left( \frac{\varphi - h_{width}}{1 - \frac{2h_{width}}{\pi}} \right) + s_{body.amp} \cos (2\varphi) 
\]  

(4.5)

\[
\text{else:}

h_{eq} = h_0 + s_{body.amp} \cos (2\varphi).
\]  

(4.6)

Where $\varphi$ is latitude. The equilibrium or steady state solution depends on five parameters: $h_0$, $h_{width}$, $h_{amp}$, $s_{body.amp}$, $\tau_{thermal}$. $h_0$ describes the depth of the fluid in absence of the solid body rotation term and the polar vortex. It has to be large enough to allow for the sharp drop-off near the North pole. Note, that the zonal wind is determined by the latitudinal gradient of the height field and not the absolute depth of the fluid. As a result, the impact of $h_0$ is relatively small. We choose $h_0 = 8\,\text{km}$ for all results presented in this chapter. $h_{width}$ controls the latitude at which the equilibrium height profile starts to fall off to simulate the polar vortex. Thus, $h_{width}$ has a strong effect on the position of the jet in the Northern hemisphere. The wave amplitudes in the model depend on the relative position of the topography and the jet. We keep $h_{width}$ fixed at $15^\circ\text{N}$. $s_{body.amp}$ represents the strength of the solid body rotation term. This term was introduced to produce a more realistic wind profile outside of the polar vortex. In particular, the term is necessary to create easterly flow at lower latitudes, which influences where planetary waves break. We choose a constant value of $s_{body.amp} = 1\,\text{km}$. $\tau_{thermal}$ controls how closely $h$ is tied to $h_{eq}$. We choose a thermal damping time scale of 10 days, which is in the range of radiative time scales in the stratosphere (Andrews et al. 1987). Finally, $h_{amp}$ controls the minimum equilibrium height at the North pole. We will also refer to $h_{amp}$ as the strength of the polar vortex. Larger values of $h_{amp}$ lead to a smaller equilibrium height at the North pole, a stronger vortex...
Figure 4.1: Equilibrium height profiles as a function of latitude for different values of $h_{amp}$. a) $h_{amp} = 0.5 \text{ km}$ b) $h_{amp} = 5.0 \text{ km}$ c) $h_{amp} = 7.0 \text{ km}$.

and a higher wind speeds of the jet. Figure 4.1 shows three different equilibrium height profiles for three different values of $h_{amp}$. The smallest value of $h_{amp}$ leads to a profile that is dominated by the solid body rotation term (Figure 4.1a). For larger values of $h_{amp}$, the profile becomes less symmetric relative to the equator and falls off sharply at the North pole. The minimal equilibrium height at the North pole in Figure 4.1c is probably not realistic, but one advantage of simplified models is that they can explore extremes in the parameter space. In general, larger values of $h_{amp}$ lead to stronger zonal winds winds around the vortex. We vary this parameter systematically and will discuss the results in the following sections.

The original version of the FMS shallow-water model does not include an option to introduce bottom topography. Thus, we modify the code and create the new variable $h_b$ to represent topography (see Equation 4.3). The use of a static topography is somewhat different from the setup in Liu and Scott (2015), who, following Esler and Scott (2005), used a topography that is moving with an angular frequency representing solid body rotation. The topography can be interpreted as a forcing term, because the atmosphere relaxes to a steady state if $h_b = 0$ and no other time dependence is introduced. We use the same functional form for the bottom topography as Polvani et al. (1994):

$$h_b = H_b T(t) e^{-((\phi - \phi_0)/\Delta \phi)^2} \cos(k\lambda). \quad (4.7)$$

where $\lambda$ is longitude, $\Delta \phi$ determines the width of the bottom topography. Wider topography has a stronger effect on the vortex. We do not vary this parameter for the presented results. We choose $\Delta \phi = 15^\circ$. $\phi_0$ controls the latitude of the maximum height of the topography. The effect of the topography increases with latitude. This could be due to the way the waves and the jet align or because the layer thickness decreases with latitude, so that the topography represents a larger fraction of the layer depth. As discussed above, the position of the topography relative to the jet latitude affects the wave amplitudes. We keep $\phi_0$ constant at 45°N to isolate the effect of the height and wavenumber of the forcing. $T(t)$ is a dimension-
Figure 4.2: Surface topography in geopotential \([m^2/s^2]\) for wavenumber one (a) and two (b) with \(H_b = 2\, \text{km}\).

less time-dependent scaling factor. We use \(T(t)\) to introduce topography gradually to avoid the formation of large gravity waves, which can cause numeric instabilities. \(H_b\) is the amplitude and \(k\) the wavenumber of the forcing. Exploring the effect of both parameters is at the center of this study and will be described in detail in the next section. Figure 4.2 shows the surface topography in geopotential for wavenumbers one and two. We do not change the orientation of the topography, as the zonally symmetric vortex should not be affected by the longitude of the mountain. Note that the forcing is proportional to the cosine of longitude, which means that we impose equally shaped mountains and valleys.

The shallow-water model does not include land-sea contrast, an explicit representation of radiation or chemistry, moisture, clouds, a seasonal or daily cycle and many other factors that are important to model a realistic climate. However, we do not expect these processes to be of first order importance to the interactions of planetary waves and the polar vortex. The aforementioned studies by Polvani et al. (1994); Liu and Scott (2015) and Scott (2016) make us confident that the presented set of equations is useful to address the outlined questions.

We perform three sets of experiments. First, we run the model without topography but with varying vortex parameters for 500 days each. In this configuration the model reaches a steady state relatively quickly. We use this setup to evaluate our choices for the parameters in Equation 4.6. Furthermore, we use the equilibrated states as a starting point for the ensemble experiments in Section 4.3.2.

In a second set of experiments, we disturb the same vortex with wave one and two topography of varying amplitude. These simulations address the question if our model setup can produce realistic SSW frequencies. Additionally, we are interested in the question if wave one (two) forcing only produces displacements (splits). Gerber and Polvani (2009) and She-shadri et al. (2015) found in a dry dynamical core model, that certain forcings would only lead to one type of event and that only splits were produced at a realistic frequency. Our simulations were performed before we introduced the solid body rotation term in Equation 76.
4.6. We believe that the results are still useful to understand the general behavior of the model. All experiments are performed with $h_{amp} = 4 \text{ km}$, which results in a maximum jet of about 60 m/s. The model is forced with wave one topography between 1.0 km and 5.0 km and wave two topography between 0.75 km and 3.5 km. We then count the number of SSWs as defined in Charlton and Polvani (2007) and classify them into splits and displacements using the algorithm described in Chapter 3 and published in Miller et al. (2013).

Finally, we use the equilibrium states from the first set of experiments and use them as a starting point for a number of 250 day long runs. At the beginning of each run, we introduce topography using the following function for $T(t)$ in Equation 4.7

$$T(t) = 1 - \exp((-t)/\tau_{hb}).$$  \hspace{1cm} (4.8)

This equation introduces $\tau_{hb}$ as the characteristic time scale of the topography build-up. As vortex breakups can vary significantly, we vary $\tau_{hb}$ between 5.0 and 4.90 days in increments of 0.01 days to create ensemble members for each combination of topography and vortex. We find no systematic changes in vortex behavior for the small changes in $\tau_{hb}$. Thus, we view the ensemble members as equally valid representations of the wave-vortex interaction.

Figure 4.3 shows the maximum topography as function of time for one of the simulations. The topography increases slowly enough to avoid numerical issues, but fast enough to be considered constant following day 30. Table 4.1 summarizes the parameter combinations for this set of experiments.
Table 4.1: Summary of the parameter combinations for the ensembles presented in Chapter 4.3.2.

<table>
<thead>
<tr>
<th>k</th>
<th>$h_{\text{amp}}$ [km]</th>
<th>$H_b$ [km]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0</td>
<td>2.0, 2.5, 3.0, 3.5, 4.0, 4.5</td>
</tr>
<tr>
<td>2</td>
<td>0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0</td>
<td>0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.5</td>
</tr>
</tbody>
</table>

Results

Model evaluation

In a first step, we evaluate runs without topographic forcing to make sure that our “base climate” represents realistic stratospheric winter conditions -given the constrains of a shallow-water model. Figure 4.4 shows the geopotential, zonal and meridional wind at 60°N as a function of time. In this case, $h_{\text{amp}}$ is 7 km, which represents the strongest vortex (zonal wind of about 130 m/s) in this study. We find that stronger vortices equilibrate slower than weak ones (not shown). However, it is clear from Figure 4.4 that even the strongest vortex reaches steady state well before the end of the run on day 500. We expect the meridional wind to vanish due to the zonally symmetric equilibrium height profile. Figure 4.5 compares the zonal mean zonal wind of three runs with no topography and different values for $h_{\text{amp}}$. The small value of $h_{\text{amp}}$ in Figure 4.5a shows strong easterly winds at the equator and westerly jets in both hemispheres. This wind profile is dominated by the solid body rotation term. Increasing the value of $h_{\text{amp}}$ decreases the value of the equilibrium height profile north of $h_{\text{width}}$. As a result, the zonal mean wind increases in the Northern hemisphere. The Southern hemisphere is virtually unaffected by $h_{\text{amp}}$. As the geostrophic zonal wind is proportional to the latitudinal derivative of geopotential, the wind scales close to linear with $h_{\text{amp}}$. Our results differ from Liu and Scott (2015) and Scott (2016) in an interesting way. In these studies, an additional drag term is needed to prevent unrealistic growth of easterly winds at the equator. In our model, easterly winds do not grow over time, so no additional drag term is needed. In fact, introducing friction as described in Equation 4.5 can prevent the formation of easterlies entirely. At this point, the cause of the differences between these similar model setups is unclear. In summary, we find that the chosen model settings lead to a predictable and reasonable wind profile and that the model reaches steady state well before day 500. We proceed by exploring the effect of topography on the model.

As described above, the second set of experiments does not include easterlies at the equator and is based on a vortex with maximum winds around 60 m/s. Each simulation is 1000 days long. Table 4.2 shows the number and type of SSWs for each combination of wavenumber and height of the bottom topography. As in the previous chapter, we use the SSW definition
Figure 4.4: Geopotential (a)), zonal wind (b)) and meridional wind (c)) at 60°N as a function of time with $h_{amp} = 7\, km$.

Figure 4.5: Equilibrium zonal mean zonal wind as a function of latitude associated with the equilibrium height profiles in Figure 4.1.
from Charlton and Polvani (2007). Unsurprisingly, we find that the SSWs have preferred orientations, which is consistent with the constant orientation of the surface topography. For wavenumber one topography, the number of SSWs tends to increase with the height of the topography. The large number of splits in the run with $h_b = 5\, km$ is not physical. Closer examination shows that the vortex never fully recovers from the large wave forcing and is constantly broken up. The detected SSWs represent seemingly random vacillations of the zonal wind at 60°N around 0 m/s. We observe a similar behavior for wavenumber two topography. The number of SSWs increases with increasing topography for amplitudes below 2.25 km. At larger forcing amplitudes, the vortex is too distorted to recover to westerly zonal mean winds at 60°N, which explains the smaller number of SSWs. At this point the vortex and SSWs become ill-defined. Table 4.2 shows that the polar vortex is more sensitive to wave two forcing compared to wave one forcing of the same height. This behavior can possibly be explained with the larger geopotential gradients imposed by larger wavenumbers. However, Gerber and Polvani (2009) found that topography with wavenumber three and four leads to less stratospheric variability. The difference is likely due to the fact that these waves are unable to propagate into the stratosphere in the three-dimensional model.

We discuss the issues associated with SSW definitions in Chapter 3. While SSWs might not be the perfect measure of the vortex response, they are commonly used to compare model runs and reanalysis datasets. We will discuss this issue in more detail below. The SSW frequencies of the shallow-water model should not be compared to the climatological values (e.g. Charlton and Polvani (2007)) as our model setup is very simplified and does not include a seasonal cycle. It is however encouraging that we can find runs with frequencies close to the values reported by Gerber and Polvani (2009). As in the previous chapter, we use the algorithm described in Miller et al. (2013) to classify the SSWs as splits or displacements. However, the two parameters have to be adjusted to the higher resolution of the shallow-water model compared to the reanalysis data. The parameter choices are arbitrary but agree with our subjective classification reasonably well. Table 4.2 shows that displacements (splits) are more likely to occur in runs with wave one (two) topography. Only, the simulation with 3.25 km wave one topography produces a realistic ratio of splits to displacements. However, in contrast to Gerber and Polvani (2009), we find SSWs with dominant wavenumbers that are not equal to the wavenumber of the forcing. Figure 4.6 depicts four snapshots of the geopotential at the central date of different SSWs. The events in the left column are identified as displacements, the ones on the right as splits. The SSWs in the top row were produced with wavenumber two forcing, while the bottom row is the result of wavenumber one topography. The answer to our first question is that the model can produce either type of event with forcing of either wavenumber. By design, the simulations with wave one topography have no wave two component. Thus, split (and large wave two amplitudes) in simulations with wave one topography have to be unequivocally caused by wave-wave interactions. This is result further supports the idea that energy transfer within the stratosphere might play an important role in the development of stratospheric splits as discussed in Chapter 3. However, it is difficult to isolate individual SSWs in the presented simulations. Thus, the next Section focuses on experiments with one event per simulations.
Figure 4.6: Contour plot of geopotential \([m^2/s^2]\) at the central dates of 4 different SSWs. The events in the top (bottom) row were forced with wave two (one) topography. The left (right) column was characterized as displacements (splits).
We conclude that the shallow water model can produce reasonable zonal mean wind profiles, which we can control and predict using the equilibrium height profile. Furthermore, the model’s response to topography results in variability comparable to other simplified models and reanalysis data. Importantly, we find SSWs of either type with forcings of either wavenumber. Thus, it is a useful tool to explore the effect of wavenumber one and two on the polar vortex.

Transient Vortex Response

In this section, we present the results of the third set of experiments described in Section 4.2. We run the model into equilibrium using $h_{amp}$-values between 0.5 km and 7 km, which leads to jets around the polar vortex between about 30 m/s and 130 m/s. We then introduce wave one topography with $H_b$ ranging from 2 km to 4.5 km and wave two topography between 0.5 km and 2.5 km. We repeat each experiment 11 times with small variations of $\tau_{lb}$ to assess the variability in the vortex response. Figure 4.7 shows the zonal mean wind at 60°N for the ensembles with two different wave two topographies and two different vortex strengths. Despite the small difference in forcing amplitude, the vortex response in Figures 4.7a and 4.7c is qualitatively different. While the simulations with $H_b = 1$ km show no sign of a vortex collapse, the same vortex reacts strongly to topography of just 250 m more. The abrupt change is consistent with theories on non-linear resonance as the cause for the large wave two amplitudes in the stratosphere and split SSWs. However, our analysis is not suited to prove that resonance is causing the observed behavior. We therefore refer to the phenomenon as “quasi-resonant”. The vortex response is similar to the results in Liu and Scott (2015).

The ensemble members allow us to show the spread in the vortex response. For runs with sub-critical forcing, the 11 model runs are basically indistinguishable until around day 650 and the spread remains small until the end of the simulations (4.7a). Experiments with super-critical forcing diverge more strongly. The timing of the vortex collapse and the minimum zonal winds are more consistent than the vortex recovery. The threshold at which the vortex behavior changes depends on the vortex strength. Figures 4.7a and 4.7b show the zonal wind response to the same forcing for different vortices. One might assume that stronger winds are less affected by the same topography. However stronger winds also produce larger planetary wave amplitudes, if they interact with the same topography. Additionally, the stronger PV-gradient serves as a wave guide and affects wave propagation. In our experiments, stronger vortices consistently collapse at lower surface amplitudes. This behavior was also observed by Liu and Scott (2015). However, the oscillations in our model have a longer period, which is probably due to the different model setups. We find a second vortex breakup at the end of some simulations (e.g in Figure 4.7c). The calculations of Liu and Scott (2015) show vortex cycles on the order of tens of days.

Figure 4.8 shows the zonal wind response for simulations with wavenumber one topography. As for wavenumber two, the vortex response depends on the height of the topography.
Table 4.2: Overview of experiments with $h_{amp} = 4\, km$ and varying topographic forcings. SSWs are defined following Charlton and Polvani (2007) and categorized following Miller et al. (2013).

<table>
<thead>
<tr>
<th>$H_b$ [km]</th>
<th>Wavenumber</th>
<th># of SSWs</th>
<th>Splits/Displacements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>1</td>
<td>0</td>
<td>0/0</td>
</tr>
<tr>
<td>2.0</td>
<td>1</td>
<td>0</td>
<td>0/0</td>
</tr>
<tr>
<td>2.25</td>
<td>1</td>
<td>0</td>
<td>0/0</td>
</tr>
<tr>
<td>2.5</td>
<td>1</td>
<td>0</td>
<td>0/0</td>
</tr>
<tr>
<td>2.75</td>
<td>1</td>
<td>0</td>
<td>0/0</td>
</tr>
<tr>
<td>3.0</td>
<td>1</td>
<td>2</td>
<td>0/2</td>
</tr>
<tr>
<td>3.25</td>
<td>1</td>
<td>5</td>
<td>2/3</td>
</tr>
<tr>
<td>3.5</td>
<td>1</td>
<td>3</td>
<td>2/1</td>
</tr>
<tr>
<td>3.75</td>
<td>1</td>
<td>12</td>
<td>2/10</td>
</tr>
<tr>
<td>4.0</td>
<td>1</td>
<td>15</td>
<td>2/13</td>
</tr>
<tr>
<td>5.0</td>
<td>1</td>
<td>13</td>
<td>10/3</td>
</tr>
<tr>
<td>0.75</td>
<td>2</td>
<td>0</td>
<td>0/0</td>
</tr>
<tr>
<td>1.0</td>
<td>2</td>
<td>2</td>
<td>2/0</td>
</tr>
<tr>
<td>1.25</td>
<td>2</td>
<td>4</td>
<td>4/0</td>
</tr>
<tr>
<td>1.5</td>
<td>2</td>
<td>11</td>
<td>11/0</td>
</tr>
<tr>
<td>1.75</td>
<td>2</td>
<td>12</td>
<td>12/0</td>
</tr>
<tr>
<td>2.0</td>
<td>2</td>
<td>13</td>
<td>12/1</td>
</tr>
<tr>
<td>2.25</td>
<td>2</td>
<td>18</td>
<td>15/3</td>
</tr>
<tr>
<td>2.5</td>
<td>2</td>
<td>13</td>
<td>10/3</td>
</tr>
<tr>
<td>2.75</td>
<td>2</td>
<td>9</td>
<td>8/1</td>
</tr>
<tr>
<td>3.0</td>
<td>2</td>
<td>4</td>
<td>4/0</td>
</tr>
<tr>
<td>3.25</td>
<td>2</td>
<td>2</td>
<td>2/0</td>
</tr>
<tr>
<td>3.5</td>
<td>2</td>
<td>4</td>
<td>4/0</td>
</tr>
</tbody>
</table>
Figure 4.7: Zonal mean zonal wind at 60°N for 11 ensemble members with wave two topography. The forcing amplitude is $H_b = 1.0 km$ ($H_b = 1.25 km$) for the top (bottom) row. The vortex strength in the left (right) column is $h_{\text{amp}} = 4.5 km$ ($h_{\text{amp}} = 6.5 km$).
and the strength of the jet. However, the response changes gradually as a function of the forcing. Matthewman and Esler (2011) and Esler and Matthewman (2011) predict abrupt changes in wave activity between resonant and non-resonant regimes. We find no such threshold behavior for wavenumber one forcing (note that $H_b = 1 km$ and $H_b = 1.25 km$ in Figure 4.8, but $H_b = 2 km$ and $H_b = 4 km$ in Figure 4.8). This result is not necessarily surprising as displacements are rarely discussed as resonant phenomena. Esler and Matthewman (2011) present a theory for the formation of displacements as the result of resonance of the baroclinic mode. As one-layer shallow-water models are necessarily barotropic, our experiments are not designed to address the question if displacements are the result of baroclinic resonance.

The vortex responds very differently to the two wavenumbers. First, the zonal wind falls off sharply and recovers within a few days at the beginning of all runs forced with wavenumber one topography. The amplitude of the drop increases with increasing vortex strength and topography. The zonal wind is almost unaffected by the build-up of wave two topography. We have no explanation for this behavior at this time. Furthermore, the experiments with wavenumber one forcing show either very little change in the zonal wind at 60°N or approach a new quasi-equilibrium after the vortex strength is reduced. There are only a few simulations in which the vortex collapses and recovers within the 250 days. In these cases, the vortex collapse is usually not nearly as strong as for wavenumber two (see Figure 4.8b around day 580) or the recovery starts at the very end of the simulated time period (e.g. 4.8c). This behavior is very different from observations of SSWs, but comparable to results of the dry-dynamical core simulations in Gerber and Polvani (2009). However, Taguchi and Yoden (2002) found the most realistic variability in experiments with wave one forcing. As we observe SSW-like behavior with wave one forcing in the previous section, we perform a number of experiments to explore the underlying reasons for the difference in simulations with constant and switch-on wavenumber one topography. First, we test the effect of the different equilibrium height profiles. In a limited number of experiments, we find that the strong initial response of the vortex is weaker in simulations with the simple equilibrium height profile described in Section 4.3.1. The main difference between the two profiles is the formation of easterly winds in the tropics if the the profile in Figure 4.1 is used. The easterlies change where planetary waves break and might therefore influence the wave-mean flow interaction in the Northern hemisphere. However, it is unclear why this effect would be more pronounced for wavenumber one topography. Beside the initial response, experiments with and without easterlies show a remarkably similar response to the switch-on forcing. In a second set of experiments, we extend one simulation with constant topography and one simulation with switch-on forcing to 10000. Figure 4.9 shows the zonal mean wind for both simulations, which were relaxed to the equilibrium height profile without solid body rotation. The simulations demonstrate that both experimental setups lead to the same quasi-equilibrium state, which includes SSW-like breakups and a relatively weak vortex. This result is remarkable as Liu and Scott (2015) describe a qualitative change in the vortex behavior after the initial breakup. The model response to wavenumber one forcing is to slowly degrade the polar vortex before SSW-like vacillations occur. This behavior is less
obvious with constant topography as the equilibrium vortex strength is never observed. The 250 day timescale in our simulations is too short to observe the SSW-like vortex breakups with wave one forcings. Thus, we are unable to create composites of SSWs with wave one forcings to further explore the role of wave-wave interactions in this model. However, the results can still be used to illustrate the fundamentally different interaction of the two wavenumbers with the same vortices. Many simulations in Figures 4.7 and 4.8 show a sharp decrease of the zonal wind without turning easterly at any point (e.g. all runs in Figure 4.7b). As a result, the differences between the runs are not adequately captured by the SSW definition of Charlton and Polvani (2007). We decide to adopt a different measure of the vortex response. Similar approaches are common in studies with simplified models (Esler and Matthewman 2011) and have also been used on reanalysis data (Martineau and Son 2015). We test two different measures. The difference between the 10 most westerly and 10 most easterly winds leads to intuitive results for most experiments (not shown). However, in some cases, the vortex is almost completely destroyed, so that there are very few days resembling the undisturbed vortex. As a result, the wind response is small even if the vortex is extremely distorted. We find that the largest wave amplitude over the course of the 250 days is a more consistent measure of the planetary waves caused by wave two topography. The strong initial signal in the simulations with wave one topography are associated with large wave amplitudes (Figure 4.17). Thus, the largest amplitude is often associated with the unnatural behavior at the very beginning of the experiments. We decide to measure the response to wave one topography as the maximum wave one amplitude after day 30 and the wave two amplitude after day 100 of a given simulation. These thresholds are somewhat arbitrary, but we find that they most accurately characterize our results. Figure 4.10 shows the largest wave amplitude at 60°N for runs with wave two topography as a function of vortex strength. Columns represent results of simulations with the same value of $H_b$. The top row shows wave one amplitudes; the lower row wave two amplitudes of the same experiments. The previously described threshold behavior is even more apparent in Figure 4.10c. The wavenumber two amplitude increases abruptly for $H_b = 1.25 km$ for vortices with $h_{amp}$ larger than 3 km. The critical value of $h_{amp}$ increases with decreasing forcing. As a result all vortices produce large wave amplitudes in Figure 4.10f. For the smallest forcing, even the strongest vortex produces only small wave amplitudes. We enforce a perfect wave two topography and purely zonal equilibrium height profile. Consequently, we observe no wavenumber one amplitude at the beginning of runs with wave two forcing (e.g Figure 4.15a). As wave-wave interactions can only transfer energy following the rules outlined in Equations 3.6 and 3.10, it is not surprising that experiments with sub-critical forcing never develop large wave one amplitudes (see the weaker vortices in Figure 4.10b). However, for super-critical forcing amplitudes, the wavenumber one amplitudes are comparable to the wavenumber two amplitudes for the same simulations. The initial wave one amplitudes have to be the result of spontaneous instabilities. Similar behavior has been described by Flierl et al. (1987). Non-zero wave one amplitudes can then grow through wave-wave interactions such as sub-harmonic resonance. The timing of these processes will be discussed at the end of the chapter.
Figure 4.8: Zonal mean zonal wind at 60°N for 11 ensemble members with wave one topography. The forcing amplitude is $H_b = 2.0\, \text{km}$ ($H_b = 4.0\, \text{km}$) for the top (bottom) row. The vortex strength in the left (right) column is $h_{\text{amp}} = 4.5\, \text{km}$ ($h_{\text{amp}} = 6.5\, \text{km}$).
The quasi-resonant growth of wavenumbers one and two in response to wavenumber two topography can also be seen in Figure 4.11. Here, we show the maximum wave amplitudes as a function of the surface topography $H_b$. As in Figure 4.10, we find abrupt jumps in wave two amplitudes leading to sub- and super-critical regimes. The stronger jet (right column) leads to large wave amplitudes at smaller surface amplitudes. Simulations with large wave two amplitudes also produce significant wave one amplitudes. However, the maximum wave one amplitudes are less predictable and the spread among ensemble members is large. Additionally, the spread is also substantially larger for super-critical runs, which agrees with the observation that the behavior of vortex breakups is less consistent than the weak, sub-critical vortex response.

Figure 4.12 and 4.13 show the equivalent analyses to Figures 4.10 and 4.11 for the experiments with wavenumber one topography. In Figure 4.12, the wavenumber one amplitude increases almost linearly with vortex strength. For the smaller forcing (left column), the response of vortices with $h_{amp}$ between 3 km-5 km is somewhat smaller than would be expected from purely linear behavior (see Figure 4.12a). The spread among ensemble members increases with the strength of the vortex, but there is no indication of resonant behavior or a threshold that divides the simulations into different regimes. As we discussed in Chapter 2, wave-wave interactions can transfer energy from wavenumber one into wavenumber two. Thus it is not surprising that even weak jets and small topographies result in non-zero
Figure 4.10: Maximum geopotential wave amplitudes for wavenumber one (top) and two (bottom) at 60°N as a function of vortex strength for experiments with wave two topography. a) H_b = 0.5 km  b) H_b = 1.25 km  c) H_b = 2.5 km
Figure 4.11: Maximum geopotential wave amplitudes for wavenumber one (top) and two (bottom) at 60°N as a function of the amplitude of topography for experiments with wave two topography. a) c) $H_{amp} = 2.0 \text{ km}$ b) e) $H_{amp} = 5.5 \text{ km}$
Figure 4.12: Maximum geopotential wave amplitudes for wavenumber one (top) and two (bottom) at 60°N as a function of vortex strength for experiments with wave one topography. The maxima were calculated for the period after day 30 (100) for wavenumber one (two). a) $H_b = 2.0\, km$ b) $H_b = 4.5\, km$
wave two amplitudes. The wavenumber two amplitudes increase linearly with $h_{\text{amp}}$.

The growth of wavenumber one with increasing topography is less linear than with vortex strength. The amplitudes appear to respond more strongly to increased topography for smaller surface amplitudes (Figures 4.13a and 4.13b). The wave two amplitudes even decrease for $H_b$ larger than 3 km. The largest mountains lead to extreme distortion of the polar vortex and a significant loss of vortex material. We believe that the reduced amplitudes are the result of the weak vortices in these runs.

The previous figures only show slices through the explored parameter space. In order to summarize all experiments, we calculate ensemble averages of the maximum amplitudes at 60°N. Table 4.1 shows all parameter combinations, which were used to explore the parameter space. While, we use equally spaced intervals of $h_{\text{amp}}$, more simulations are performed with wavenumber two topography amplitudes close to the observed threshold. Simulations with strong vortices and large wavenumber one topography are numerically unstable in the current model setup. Note the non-linear color bar, which is used to span the large range of values. The top (bottom) row represents runs with wave one (two) topography. The most prominent feature of Figure 4.14 is the similarity between the four plots. In all cases, the maximum wave amplitude is mostly determined by the height of the surface topography for small values of $H_b$. For large surface topography, the wave amplitudes depend mostly on the strength of the vortex. We believe that each equilibrium state might be associated with a maximum wave amplitude and experiments with large forcing could approach this maximum.

There are however important differences between the two sets of experiments. As expected from the previous discussion, we find a steep increase in wave two amplitudes in response to wave two topography. In general, weaker vortices experience quasi-resonant behavior at larger forcing amplitudes. It might be counter-intuitive that stronger forcings are required to collapse a weaker vortex. However, these results agree with the results of Esler and Scott (2005), Matthewman and Esler (2011) and Liu and Scott (2015), who find that stronger vortices split at lower forcing amplitudes. Additionally, we observe in Chapter 2 that the strongest heat flux events are usually preceded by anomalously strong vortices. Furthermore, we find that the threshold becomes less pronounced for weaker vortices. The most prominent feature of wavenumber one amplitudes for simulations with wavenumber two topography is the large fraction of the parameter space with zero wave one amplitudes. As previously discussed, spontaneous instabilities are responsible for the wavenumber one amplitudes at larger forcings or vortex strengths. The rapid wave growth appears at the same parameter combinations for wave one and two amplitudes. However, the wave one threshold is more pronounced for weak vortices. For super-critical forcing, wave one amplitudes can be comparable to the wave two amplitudes in the same simulation.

Considering that experiments with wave one topography reach a new steady state, while wave two topography leads to SSW-like variability, the differences between the maximum wave amplitudes in the two sets of experiments are remarkably similar. This result illustrates
Figure 4.13: Maximum geopotential wave amplitudes for wavenumber one (top) and two (bottom) at 60°N as a function of the amplitude of topography for experiments with wave one topography. The maxima were calculated for the period after day 30 (100) for wavenumber one (two). a) $H_{amp} = 3.0 \text{ km}$ b) $H_{amp} = 6.0 \text{ km}$
Figure 4.14: Summary of all experiments listed in Table 4.1. Ensemble averages of maximum amplitudes of wavenumbers one (left) and two (right) at 60°N as a function of $h_{amp}$ and $H_b$. Experiments with wave one (two) topography are in the top (bottom) row. Wave one (two) amplitudes for wave one forcing exclude days 1-29 (1-99). Note the non-linear color bar.
Figure 4.15: Geopotential wave amplitudes of wavenumber one (top) and two (bottom) at 60°N as a function of time for experiments with wave two topography and \( h_{amp} = 4.5 \text{km} \). a) \( H_b = 1.0 \text{km} \) b) \( H_b = 1.25 \text{km} \)
Figure 4.16: Geopotential wave amplitudes of wavenumber one (top) and two (bottom) at 60°N as a function of time for experiments with wave two topography and $h_{amp} = 6.5 \text{ km}$. a) $H_b = 1.0 \text{ km}$ b) $H_b = 1.25 \text{ km}$
that the largest amplitude is only one measure to understand the presented experiments. Additional analysis is needed to discover the difference between the SSW-like vortex breakups and the slow degradation of the polar vortex. The largest differences between the two sets of experiments is the gradual vortex response in simulations with wave one topography. There is no sign of resonant growth in the top row of Figure 4.14. Furthermore, the largest wave one amplitudes are significantly larger than the wave two maxima of the same simulations. As mentioned above, we decided to exclude the wave amplitudes associated with the initial response to wave one topography when we calculate the maximum wave amplitudes. This decision mostly affects runs with small topography, where the largest amplitude is often detected within the first few days of the simulation. As a result, the wave one amplitudes for runs with wave one topography appear to depend solely on the strength of the vortex (not shown).

Our final analysis investigates the wave amplitudes in geopotential as a function of time. Figures 4.15, 4.16 and 4.17 show the wave amplitudes of the first two harmonics of the simulations presented in Figures 4.7 and 4.8. Figure 4.15 shows the wavenumber one and two
amplitudes of all ensemble members for the simulations with $h_{\text{amp}} = 4.5 \text{ km}$ and $H_b = 1 \text{ km}$ or $H_b = 1.25 \text{ km}$. The two sets of experiments show declining wave two amplitudes following the build-up of wave two topography. However, the qualitative difference between the zonal wind time series in Figures 4.7a and 4.7c is also reflected in the wave amplitudes. The 11 ensemble members with smaller topography develop almost identically until day 650. From this point on, the wave two amplitudes increase and the ensemble members start to diverge. While the polar vortex remains intact, the zonal wind time series in Figure 4.7a begin to diverge at the same time. Finally, this transition is characterized by the development of wavenumber one amplitudes. The wave one amplitudes are comparable to wavenumber two around day 680. However, they decline from here while the wavenumber two amplitudes continue to grow. This result shows that both wavenumbers can be important for the development of the polar vortex below the critical forcing amplitudes. Figure 4.15d shows the abrupt increase in wave two amplitudes for the runs with super-critical forcing around day 580. Following the sudden increase, the amplitudes remain large and variable for about 80 days before they decline to very small values and finally begin to increase more gradually. This life cycle goes along with the collapse, recovery and second collapse of the polar vortex in Figure 4.7c. Interestingly, the growth of wavenumber one occurs around the same day as in the sub-critical experiments, despite very different states of the polar vortex at that time. In the sub-critical case, wavenumber one begins to grow as the ensemble members diverge. The super-critical wave one amplitudes grow after the collapse of the vortex during a time of vortex recovery. The wave one amplitudes remain relatively small and only become comparable to the wave two amplitudes for a short period of time around day 700.

Figure 4.16 shows the wave amplitudes for the simulations in Figures 4.7b and 4.7d, which were forced with the same topography, but have stronger jets than the experiments discussed above. As a result, both ensembles experience rapid growth of wavenumber two. We find that the timing of wave two growth and the collapse of the polar vortex can be used to characterize the response of the polar vortex in our experiments. The larger the forcing and the stronger the vortex, the earlier the vortex collapses and the earlier the waves grow. On the other hand, the formation of significant wave one amplitudes over the second half of the experiments does not align with the rapid growth of wave two or the collapse of the polar vortex. The extremely large amplitudes in Figure 4.16 lead to a large spread among different ensemble members. Thus, the predictability of the vortex behavior decreases quickly at the end of runs with large forcings (Figure 4.16d). Liu and Scott (2015) argue that vortex breakups fundamentally change the dynamics of the system and therefore represent a limit for a meaningful interpretation of their experiments.

Finally, the wave amplitudes for three different wave one forcings and $h_{\text{amp}} = 4.5 \text{ km}$ are presented in Figure 4.17. The strong and short-lived response of the polar vortex at the beginning of the simulations in Figure 4.8 is reflected in a spike in wavenumber one amplitudes. Neither the analysis shown here, nor the subjective interpretation of the vortex behavior, indicate that the initial response alters the base state of the experiments. Following the short displacement, the vortex is reestablished at the pole where it remains during
the runs with small forcings. During strongly forced runs, the vortex is substantially weakened and then displaced off the pole. The amplitude of the initial peak increases with the strength of the vortex. We interpret the strong initial response as an artifact of the model setup. Thus, we calculate the maximum wave one amplitude after day 30. In simulations with very strong jet velocities, wave-wave interactions lead to a substantial wavenumber two peak as well. As the wavenumber two amplitudes peak in the second half of the experiments, we determine the maximum wave two amplitudes after day 100. Comparing Figures 4.17a and 4.17b shows that the maximum amplitude, following the initial spike, depends on the forcing amplitude. As previously described, the wave one experiments show no SSW-like vortex collapse, but reach a quasi steady-state with a weaker vortex (similar to the results of Gerber and Polvani (2009)). This development results in relatively constant wave amplitudes. Larger forcing amplitudes lead to larger wave one and two amplitudes during the simulation. However, the strongest forcings (Figures 4.17c and 4.17f) completely deteriorate the vortex, which results in smaller wave amplitudes at the end of the simulations. At this stage, the vortex has almost disappeared. The wavenumber two amplitudes only spike during the first days of the simulations with very strong vortices (not shown). Following the first few days of the simulations, the wave two amplitudes increase with the height of the topography until \( H_b \) reaches about 3 km. The wave two amplitude stabilizes and is less variable in time for larger forcings.

Discussion and Conclusions

In this chapter, we use a shallow-water model to isolate the effect of wavenumber one and two on idealized polar vortices. We modify the model code to include surface topography, which causes the formation of planetary waves. Our model evaluation shows that the chosen equilibrium height profile allows us to create polar vortices of variable and predictable strength. Additionally, long runs with constant topography result in realistic looking SSWs. The frequency of vortex breakups increases with the height of the topography until the vortex is unable to recover from breakups. Our results are comparable to previous studies with simplified models of the stratosphere (Gerber and Polvani 2009) for reasonable values of \( h_{amp} \) and \( H_b \). To answer the first question from the introduction, we classify the SSWs as splits and displacements. In this set of experiments, we detect splits and displacements for forcings of either wavenumber. While the ratio of splits and displacements is not close to the real world, it still represents an improvement over previous publications. The larger geopotential gradients associated with larger wavenumbers result in stronger vortex responses to wave two topography compared to wave one mountains of the same height. In summary, we show that the shallow-water equation model can be a useful tool to explore the interaction between planetary waves and the stratospheric polar vortex.

We address the second and third question from Section 4.1 by performing a large number of experiments that explore the interaction of wavenumber one and two topography of various heights with vortices ranging from 30 m/s to 130 m/s. The experiments with wave
two forcing are similar to the results of Liu and Scott (2015) and Scott (2016). However, we expand the previous work in multiple ways. First, we use an equilibrium height profile that produces a more realistic zonal mean wind with weaker easterlies at low latitudes. Furthermore, we address the large variability of polar vortex breakups by creating ensembles for every parameter combination. Our analysis includes the response of wavenumber one to wavenumber two forcing. Finally, we perform an entirely new set of experiments, which are forced by wavenumber one topography.

The experiments with wave two topography confirm the theory and simple models of Matthewman and Esler (2011) and the results in Liu and Scott (2015). In particular, we find rapid wave two growth and SSW-like vortex breakups for sufficient forcing amplitudes and vortex strength. Differences in the timing and intensity of the vortex collapse lead to a rapidly increasing spread among ensemble members that produce SSW-like events. The simple SSW definition of Charlton and Polvani (2007) is not useful to capture the large variability among ensemble members and the range of explored parameters. Our analysis shows that the maximum wave amplitude at 60°N is a useful measure to describe the behavior of the wave two experiments. Using this measure, we show quasi-resonant growth of planetary wave amplitudes at a threshold that depends on the strength of the vortex and the surface amplitude. Below this threshold the wave amplitudes depend on the surface amplitude, but are relatively constant over large ranges of of parameters. Above the critical forcing, the wave amplitudes are only a function of $h_{amp}$. Our results suggest that a maximum wave amplitude exists for each vortex and that forcings above a certain threshold usually reach saturation. The setup of our experiments allows us to investigate the timing of events more easily than simulations with slowly changing forcing parameters in Liu and Scott (2015). We find that the larger surface topography leads to earlier vortex breakups.

Furthermore, we observe the formation of wavenumber one amplitudes in runs with wavenumber two forcing. As these experiments have, by design, no wavenumber one amplitude, wave-wave interactions are not responsible for the initial growth of wavenumber one. Thus, we believe that instabilities in the wave-vortex interaction lead to the initial, small wavenumber one component in the geopotential field. Interestingly, the timing of wavenumber one growth is remarkably consistent among ensemble members. Wavenumber one grows rapidly during the first few days following the instability in most simulations, which might be due to wave-wave interactions. The growth of wavenumber one aligns with the divergence of the ensemble members at lower forcing amplitudes (see Figures 4.15a and 4.15c). In these simulations, the formation of wavenumber one can be interpreted as the onset of chaos. We observe this chaotic behavior at forcings below, but close to the threshold for resonant growth. In fact, our results suggest that in these idealized settings the growth of wavenumber one can be used to estimate the distance from resonant behavior. The wave-number one amplitudes of experiments with super-critical forcing are more difficult to interpret. The formation of wavenumber one is not clearly related to a particular phase of the vortex breakup or recovery. Furthermore, the ensemble spread in these experiments can be large before the onset of wavenumber one growth, which indicates that wavenumber one is
not necessary for chaotic behavior in the shallow-water model.

The polar vortex’s response to wavenumber one forcing fundamentally differs from the previously described results. Most importantly, imposing wave one topography leads to new quasi-steady states instead of sudden vortex breakups within the first 250 days. Comparing the experiments with wave one forcing in Sections 4.3.1 and 4.3.2, leads us to the conclusion that wave one forcing slowly degrades the strength of the vortex before SSW-like vacillation commence with a much weaker polar vortex. Thus, we observe no SSWs within the 250 day long simulations. While Matthewman and Esler (2011), Esler and Matthewman (2011) and other have described the onset of displacements as more gradual compared to splits, it is surprising that the shallow-water model responds fundamentally different mean states with either forcing.

The difference between the two sets of experiments is particularly striking considering the qualitatively similar relationship between the maximum wave amplitudes and the controlling parameters $h_{amp}$ and $H_b$. The two main differences in Figure 4.14 are the strong gradient in runs with wave two forcing (associated with the quasi-resonant wave growth) and the larger wave two amplitudes resulting from wave one topography compared to the wave one amplitudes in simulations with wave two topography. The latter result underlines the importance of wavenumber one formation in experiments that are initiated without any wavenumber one component. In summary, we conclude that the polar vortex responds fundamentally different to similar forcing amplitudes depending on the dominant wavenumber of the planetary wave spectrum.

The shallow-water model includes many simplifications of the real atmosphere, so one has to be careful applying the results to SSWs and stratospheric variability. However, our results show that the wave-mean flow interaction is very sensitive to the details of the experimental setup. Furthermore, our results show the emergence of chaotic behavior from small instabilities and infinitesimal wave amplitudes. While the real stratosphere is always interacting with a spectrum of planetary waves, the formation of wavenumber one in experiments with wave two topography demonstrate the type of exponential growth and positive feedback that limits weather forecasts of the real atmosphere to about a week (Tripathi et al. 2015). Finally, the maximum wave amplitudes are comparable for wavenumber one and two forcing despite the fundamentally different vortex behavior. We conclude that large wave amplitudes are not sufficient to cause SSWs. Future studies should address the question under which circumstances large waves lead to SSW-like behavior without permanently destroying the polar vortex.
Chapter 5

Conclusions and Outlook

The overarching conclusion from this thesis is that planetary wavenumber one and two affect the Northern polar stratosphere in fundamentally different ways. Previous studies have often identified planetary wave propagation and breaking without separating the contributions of different harmonics. Our results illustrate the importance to investigate the mechanisms behind vertical wave propagation and the growth of wave amplitudes in the stratosphere individually. This work has implications for the predictability of SSWs, theories on the onset of vortex breakups and the modeling of troposphere-stratosphere interactions.

In Chapter 2.2, we use MERRA reanalysis data to build a climatology of wave amplitudes and heat fluxes between October and April. Our results show the increasing importance of wavenumber one with increasing height and latitude. These effects had been predicted by Charney and Drazin (1961) and Hoskins and Karoly (1981), but we provide quantitative estimates, which serve as context for the results in Chapter 2.3 and previous studies. Our analysis in Chapter 2 helps to understand the impact of the choices that lead to the various event definitions in stratospheric dynamics. Furthermore, we emphasize the importance of long-term averages based on the large inter-annual variability of the relative importance of wavenumber one and two. The most surprising result in Chapter 2.2 are the fundamentally different correlations of wavenumber one and two amplitudes between the troposphere and the stratosphere. We find that wave amplitudes and heat fluxes are well correlated. However, the wave one amplitudes at 500 hPa are almost uncorrelated to the wave amplitudes in the stratosphere. However, the regression coefficient between the wavenumber one amplitudes in the troposphere and stratosphere are large. Both measures show small values around the tropopause. We interpret this result as a limit for the prediction of SSWs, which are dominated by wavenumber one. Numerous studies have tried to connect blockings (Martius et al. 2009; Colucci and Kelleher 2015), which are associated with large wave amplitudes in the troposphere, to stratospheric wave amplitudes and SSWs. Our results might explain why these efforts have not been very successful thus far. The effect of wavenumber one on the polar vortex becomes apparent in the correlations with the wave amplitudes at 100 hPa.
At this level, wavenumber one shows strong correlations throughout the stratosphere and significantly affects the strength of the polar vortex. The layer between 300 hPa-100 hPa appears to be crucial to the understanding of wave one propagation into the stratosphere, which agrees with the results of Polvani and Waugh (2004) and Birner and Albers (2017). Interestingly, wavenumber two amplitudes in the troposphere are well correlated to the amplitudes in the stratosphere, which suggests that wave two propagates more readily into the stratosphere. While the regression coefficients for wavenumber two are smaller than for wavenumber one, both statistics show no substantial effect of the troposphere on wave two propagation. This result is unexpected as the Charney-Drazin theorem predicts smaller waves to dissipate at lower levels. A possible explanation could be a difference in the climatological phase speed of waves one and two. However, the small amplitudes of wavenumber two have almost no effect on the zonal wind at 60°N.

Future research could focus on planetary wave propagation through the tropopause as multiple studies have identified this layer as a key to explain the planetary wave activity in the stratosphere. In particular, our research shows the effect of this layer is limited to wavenumber one. Furthermore, we suggest that attempts to predict displacements from the troposphere should not simply rely on blockings or the wave one amplitude in the mid-troposphere.

The composites of extreme heat flux events in Chapter 2.3 show that the largest wave one heat fluxes are the result of one long-lasting event. In contrast, we find two short pulses in the wave two composites. The effect of extreme heat fluxes on the polar vortex matches the vortex orientations in composites of SSWs Matthewman et al. (2009). We identify statistically significant geopotential anomalies weeks before the largest heat fluxes. The phases of waves one and two provide insight into the development of large heat flux anomalies. We observe that the phase-height profiles of individual events only align with each other and with the climatological profile through most of the atmosphere at the time of the largest heat fluxes. This result establishes that the theory of linear interference between the climatology and the anomaly (Watt-Meyer and Kushner 2015) holds for extreme events of either wavenumber. Thus, one may think of large heat fluxes as an exited mean state.

Future research could attempt to identify a chain of events, which cause waves to align with the climatology. In some sense, the prediction of large heat fluxes, and thus vertical wave propagation, can be reinterpreted as the prediction of the anomalous phase-height profile. Additionally, our work raises questions about the importance of particular levels to planetary wave propagation. Furthermore, future work could address why the correlations of wavenumber one highlight the role of the tropopause, while the extremes in Chapter 2.3 align over a large range of levels.

We present potential and kinetic eddy energy budgets in Chapter 3. This tool enables us to quantify the effect of wave-wave interactions within the stratosphere. The work is motivated by large wavenumber one amplitudes preceding the growth of wavenumber two in Chapter 2.3 and in early observations of SSWs Labitzke (1977). We find that the energy transfer from wavenumber one toward wavenumber two played a key role in the vortex split.
in January 2013. Consequently, wave-wave interactions within the stratosphere can be an alternative explanation of split events. Furthermore, we identify a number of SSWs (splits and displacements) where wave-wave interactions explain most of the energy tendency of wavenumber two. However, no statistically significant results are found for the composite of all splits in the MERRA dataset. Our analysis shows that not all splits are associated with the growth of wavenumber one. Additionally, the timing of planetary wave growth is not necessarily captured by the common definitions of SSWs, which leads to small signals in the composite average. We suggest that wave-wave interactions are responsible for the growth of wavenumber two in the absence of resonance.

The employed energy budgets may be applied to numerous questions. While the interpretation of energy budgets can be misleading (Plumb 1983), the budgets could still be used to gain insight into a variety of stratospheric processes by quantifying the relative importance of stratospheric processes. In particular, the dominance of wavenumbers one and two simplifies the explicit representation of all wave-wave interactions. In principle, the methodology could also be applied in the troposphere. However, a more comprehensive representation of non-conservative processes might be needed at lower levels. Additionally, wave-wave interactions would probably be affected by larger wavenumbers. A future project could attempt to distinguish the growth of wavenumber two due to resonance from wave-wave interaction. Our results suggest that a large number of events is necessary to increase the sign-to-noise ratio. Thus, the analysis of comprehensive GCM simulations appears most promising.

Finally, the results in Chapter 4 highlight the sensitivity of the polar vortex response to the dominant wavenumber of the forcing. In particular, experiments with wave one topography produce SSW-like vortex breakups only after the vortex is gradually weakened. Thus, we observe no SSWs with wave one forcing within the initial 250 days of the simulations. In contrast, wavenumber two forcing leads to abrupt vortex collapses if the forcing exceeds a critical threshold. In agreement with Liu and Scott (2015), the critical forcing amplitude decreases with increasing vortex strength. The maximum wave amplitudes prove to be a more appropriate characterization of the experiments than traditional SSW definitions as they capture the qualitatively different vortex responses to wave two forcings over a large range of parameters. However, an analysis of the time-dependent zonal winds and wave amplitudes is necessary to distinguish vortex breakups from gradual decays. We expand previous studies by performing ensembles for each combination of wavenumber, height of topography and vortex strength. Additionally, we analyze the amplitude of wavenumber one in runs with wave two forcing. Assessing the ensemble spread in zonal winds and wave one amplitudes, we find that the formation of wavenumber one amplitudes coincides with the development of chaotic behavior in simulations with sub-critical forcing.

We unlock a number of new applications of the FMS shallow-water model by implementing surface topography. Future work could further explore how easterly winds at low latitudes affect the interaction of wavenumber one and the polar vortex. Since the simulations are numerically inexpensive, it is feasible to repeat a subset of the presented simulations to
further explore the onset of SSWs in simulations with wavenumber one topography and their connection to wave-wave interactions.

Lastly, our results have important implications for studies with simplified models. These studies often prescribe forcings of one particular wavenumber. We show here that the results of these studies have to be interpreted cautiously. The propagation of waves one and two can not simply be predicted by the Charney-Drazin theorem for stationary waves and wave-wave interactions might not be negligible. In addition, simplified models might be sensitive to details in the model setup. As a result, the correct magnitude of planetary wave forcing is no guarantee for realistic results. Our research suggests that the relative importance of wavenumbers one and two, a realistic tropopause and the zonal mean winds throughout the Northern hemisphere, might have large effects. Our results indicate the importance to choose model setups carefully and to distinguish, as much as possible, between physical results and model artifacts.

In summary, we believe that this thesis has expanded our understanding of the propagation and growth of planetary waves one and two. In particular, our results highlight the complexity of stratospheric dynamics and challenge conventional wisdom in multiple ways. First, we find that the average impact of planetary waves one and two is not predicted by a simple Charney-Drazin argument. Second, the analyses of the MERRA climatology and extreme heat fluxes appear to emphasize the upper troposphere and a large range of levels respectively. Finally, we find evidence for the importance of wave-wave interactions in reanalysis data and shallow-water simulations. The latter also emphasizes the chaotic nature of the atmosphere and the importance of processes far away from the polar vortex. Our results advance our knowledge of the mean state of the atmosphere as well as extreme events.
Appendix A

Additional Figures to Section 2.3

In this appendix, we show the geopotential anomalies associated with the extreme heat flux composites in Chapter 2.3. The figures below cover the entire length of the composite, while the chapter only contains the anomalies at the time of the largest heat fluxes to improve readability.
Figure A.1: Geopotential anomalies for the 5-day average around day -27 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.2: Geopotential anomalies for the 5-day average around day -20 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.3: Geopotential anomalies for the 5-day average around day -13 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.4: Geopotential anomalies for the 5-day average around day -6 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.5: Geopotential anomalies for the 5-day average around day 0 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.6: Geopotential anomalies for the 5-day average around day 6 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.7: Geopotential anomalies for the 5-day average around day 13 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.8: Geopotential anomalies for the 5-day average around day 20 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.9: Geopotential anomalies for the 5-day average around day 27 of the composite based on the wave-number one component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.10: Geopotential anomalies for the 5-day average around day -27 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.11: Geopotential anomalies for the 5-day average around day -20 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.12: Geopotential anomalies for the 5-day average around day -13 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.13: Geopotential anomalies for the 5-day average around day -6 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.14: Geopotential anomalies for the 5-day average around day 0 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.15: Geopotential anomalies for the 5-day average around day 6 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.16: Geopotential anomalies for the 5-day average around day 13 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.17: Geopotential anomalies for the 5-day average around day 20 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Figure A.18: Geopotential anomalies for the 5-day average around day 20 of the composite based on the wave-number two component of the meridional heat flux at six different levels. White shading indicates statistical significance at the 95% level. Note that the color bar is different for each level.
Appendix B

Comparison of SSWs in different reanalysis products
Table B.1: List of central dates and types ((S) for splits and (D) for displacements) of SSWs. The left two columns show the results in Charlton and Polvani (2007), which analyzed data before 2002. The central dates in the right column are calculated with the same algorithm and MERRA reanalysis data. The MERRA events are categorized using the algorithm in Miller et al. (2013).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NCEP</td>
<td>ERA-40</td>
</tr>
<tr>
<td>1</td>
<td>22/02/1979 (S)</td>
<td>22/02/1979 (S)</td>
</tr>
<tr>
<td>2</td>
<td>29/02/1980 (D)</td>
<td>29/02/1980 (D)</td>
</tr>
<tr>
<td>3</td>
<td>04/03/1981 (D)</td>
<td>04/03/1981 (D)</td>
</tr>
<tr>
<td>4</td>
<td>04/12/1981 (D)</td>
<td>04/12/1981 (D)</td>
</tr>
<tr>
<td>5</td>
<td>24/02/1984 (D)</td>
<td>24/02/1984 (D)</td>
</tr>
<tr>
<td>6</td>
<td>01/01/1985 (S)</td>
<td>01/01/1985 (S)</td>
</tr>
<tr>
<td>7</td>
<td>23/01/1987 (D)</td>
<td>23/01/1987 (D)</td>
</tr>
<tr>
<td>8</td>
<td>08/12/1987 (S)</td>
<td>07/12/1987 (S)</td>
</tr>
<tr>
<td>9</td>
<td>14/03/1988 (D)</td>
<td>14/03/1988 (S)</td>
</tr>
<tr>
<td>10</td>
<td>22/02/1989 (S)</td>
<td>21/02/1989 (S)</td>
</tr>
<tr>
<td>11</td>
<td>15/12/1998 (D)</td>
<td>15/12/1998 (D)</td>
</tr>
<tr>
<td>12</td>
<td>25/02/1999 (S)</td>
<td>26/02/1999 (S)</td>
</tr>
<tr>
<td>13</td>
<td>20/03/2000 (S)</td>
<td>20/03/2000 (S)</td>
</tr>
<tr>
<td>14</td>
<td>11/02/2001 (S)</td>
<td>11/02/2001 (S)</td>
</tr>
<tr>
<td>15</td>
<td>02/01/2002 (D)</td>
<td>30/12/2001 (D)</td>
</tr>
<tr>
<td></td>
<td>17/02/2002 (D)</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>18/01/2003 (S)</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>05/01/2004 (D)</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>21/01/2006 (S)</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>24/02/2007 (D)</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>22/02/2008 (D)</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>24/01/2009 (S)</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>09/02/2010 (S)</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>24/03/2010 (D)</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>06/01/2013 (S)</td>
</tr>
</tbody>
</table>
Chapter 6

References


134


