Spring 1-1-2018

Lidar Exploration of a Vertical Picture of Gravity Waves in Antarctica from Near the Surface to the Thermosphere

Jian Zhao
University of Colorado at Boulder, jian.zhao@colorado.edu

LIDAR EXPLORATION OF A VERTICAL PICTURE OF GRAVITY WAVES IN ANTARCTICA FROM NEAR THE SURFACE TO THE THERMOSPHERE

by

JIAN ZHAO

B.S., University of Science and Technology of China, 2013

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Aerospace Engineering Sciences

2018

This thesis entitled:

Lidar Exploration of a Vertical Picture of Gravity Waves in Antarctica

from Near the Surface to the Thermosphere

written by Jian Zhao

has been approved for the Department of Aerospace Engineering Sciences

_____

Dr. Xinzhao Chu (Chair)

_____

Dr. Sharon L. Vadas

Date: _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above-mentioned discipline.

Abstract

Jian Zhao (Ph.D., Aerospace Engineering Sciences)

Lidar Exploration of a Vertical Picture of Gravity Waves in Antarctica from Near the Surface to the Thermosphere

Thesis directed by Professor Xinzhao Chu

As a key link connecting different atmosphere layers, gravity waves are the leading uncertainty and one of the most puzzling elements in modern general circulation and chemical climate models. In this dissertation, we aim to establish a vertical picture of gravity waves at McMurdo, Antarctica from near the surface to the thermosphere and to improve the understandings of gravity wave coupling via investigating wave characteristics, propagation, dissipation, and generation.

Utilizing lidar observations along with theory and model simulations, we investigate the causality of gravity wave events among different altitude regions. The sources of the stratospheric gravity waves are traced to the orographic gravity waves in the troposphere, along with possible *in-situ* wave sources in the stratosphere. The sources of the Mesosphere and Lower Thermosphere persistent gravity waves are traced back to the secondary gravity wave generation in the upper stratosphere and lower mesosphere. The origins of the ~1.5 hr fast gravity waves in the Thermosphere-Ionosphere Fe/Fe^+ (TIFe) layer are linked to both secondary gravity wave generation and possible sources in the troposphere. This vertical picture reflects the sophisticated wave coupling in Antarctica.

Along the research, we characterized the seasonal variations of gravity wave parameters such as vertical wavelengths, ground-relative periods, vertical phase speeds, and potential energy densities. A spectral proportion method was developed to accurately estimate wave energy from observations.

Overall, a complex gravity wave vertical picture from the lower to the upper atmosphere was established. This picture incorporates different gravity wave mechanisms over different altitudes and seasons. Not only can it provide physical bases for gravity wave parameterization in atmospheric models, but also can it provide a representative reference to the gravity wave research across the globe. This research paves its own way to human's ultimate understandings of atmospheric dynamics in terms of gravity wave coupling.

Dedication

Dedicated to my parents Mr. Zhenjia Zhao, Mrs. Xiling Yu, my sister Mrs. Hui Zhao, and my wife Dr. Yafang Guo for their endless and unconditional love.

Acknowledgements

First and foremost, I am deeply grateful to my advisor Professor Xinzhao Chu for her invaluable science guidance in my PhD research and the unique opportunity to work on the McMurdo lidar campaigns. This dissertation would not have been possible without her continuous mentorship and encouragement. I am also in debt to Dr. Sharon L. Vadas for her remarkable insights and suggestions along the way. I sincerely acknowledge my committee members Dr. Jeffrey Forbes, Dr. Hanli Liu, Dr. Stanley C. Solomon, Dr. John Plane, and Dr. Timothy J. Fuller-Rowell for their advice and feedback throughout the PhD research.

I would like to thank Dr. V. Lynn Harvey, Dr. R. Michael Jones, and Dr. Andreas Dörnbrack for the insightful discussions and the data access to Modern Era Retrospective-analysis for Research and Applications (MERRA) and European Centre for Medium-range Weather Forecasts (ECMWF). I would like to express my gratitude to our group members Dr. Wentao Huang, Dr. Xian Lu, Dr. Zhangjun Wang, Dr. Zhibin Yu, Dr. John Smith, Dr. Weichun Fong, Dr. Cao Chen, Mr. Brendan R. Roberts, Mr. Ian Barry, Mr. Muzhou Lu, Mr. Zhengyu Hua, and Mr. Dongming Chang for their contributions to the projects, which eventually made this dissertation possible. I gratefully acknowledge Dr. Vladimir Papitashvili, Dr. Julie Palais, Mr. Richard Dean, and the staff of the United States Antarctic Program, McMurdo Station, Antarctica New Zealand, and Scott Base for their superb support of the McMurdo lidar campaigns.

The research is supported by NSF grants OPP-0839091, OPP-1246405, and OPP-1443726.

Table of Contents

CONTENTS

TABLES

Table

FIGURES

Figure

CHAPTER I

INTRODUCTION

1.1. Scientific Motivations

"Two possibilities exist: either we are alone in the Universe or we are not. Both are equally terrifying." —Arthur C. Clarke. One of the most basic scientific questions that we are supposed to answer, whether we are on our own in the grand Universe, has intrigued generations of outstanding scientists and engineers to explore the unknown. Among numerous galaxies in the Universe, we live on the third planet in the solar system. As a small rocky planet, the Earth is filled with sophisticated and coupled systems that shape it just enough to harbor life. It is not too close to the Sun like Venus, where combined with the greenhouse effect, the temperature can reach 400° Celsius, nor is it too far away from the Sun like Mars, where the temperature could be as blazing cold as minus 100° Celsius. Needless to say, there are also many other factors that make planets such as Venus and Mars not habitable to life (e.g., atmosphere composition, pressure, and magnetic fields). However, our blue marble, the Earth, has everything just appropriate to support life.

The Earth's Space-Atmosphere Interaction Region (SAIR), which extends from the stratosphere to the thermosphere, is a complicated and important region which is essential for sustaining life on Earth through absorbing extreme solar radiation, ablating meteoric materials, regulating gaseous escape, dissipating energetic particles, and deflecting magnetic fields from space, while balancing influences from the planet itself in the forms of wave energy and momentum originating from the lower atmosphere [OASIS, 2014; Chu et al., 2016]. Through research on the fundamental and universal physical, chemical, thermodynamical, electrodynamical, and wave dynamical processes taking place within the Earth's SAIR, we will be able

to infer how it shapes the atmospheres of Earth-like planets, thereby helping in the search for exoplanets. However, the physics, dynamics, and chemistry in the SAIR are rather sophisticated and our understanding of them is still in progress. Undoubtedly, comprehensively understanding the SAIR requires continuous efforts of observational discoveries and theoretical understandings from multiple generations of scientists. This dissertation focuses on a specific but crucial aspect in the Earth's SAIR — the wave dynamics.

1.1.1. The Thermal Structure, Circulation, and Chemistry in the SAIR

The Sun powers the Earth primarily via solar radiation across a wide spectral range from Infrared to Gamma-rays. The atmosphere hence gains energy and forms vertical thermal structures, which eventually lead to the formation of general circulation from local to global scales. Certainly, the atmosphere is not a closed system since it receives influences from the Earth's land, sea, and space, and gives feedback to them. The dynamics within the Earth's atmosphere is complicated and is highly dependent on time and geolocation.

Generally, the atmosphere is divided into the troposphere, stratosphere, mesosphere, and thermosphere according to the thermal structures. The white line on the left in Figure 1.1 shows a representative temperature profile with respect to altitude. The temperature in the troposphere decreases with altitude and almost all the weather systems happen within it. In the stratosphere, the temperature increases with altitude due to the solar radiation being absorbed by ozone. Consequently, the stratosphere is more vertically stable with rare bulk motions in altitude. The temperature in the mesosphere again decreases with altitude and the mesopause region in the summer is often the coldest place in the Earth's atmosphere. In the thermosphere, the temperature rises with altitude rapidly due to the low atmospheric density and the absorption of strong solar radiation by atomic O and O_2.

The SAIR, which extends from the stratosphere to the thermosphere, hosts a variety of dynamical and chemical processes that are vital to the Earth's habitable environment. Generally speaking, there are three major factors that determine the thermal structure, circulation, and variability in the Earth's atmosphere, namely, radiative forcing, chemical forcing, and dynamical forcing [e.g., Roble, 1995; Garcia, 1989; Friedman and Chu, 2007]. To a great extent, scientists are able to predict the state of the atmosphere with a combination of radiative forcing and chemical forcing, assuming that the atmosphere reaches photochemical equilibrium. However, there are certain phenomena that cannot be described within such a limited framework such as the zonal wind reversal in the upper stratosphere and the mesosphere, the deceleration of the polar night jet, and the coldest mesopause in the summer, etc. [Smith, 2012]. Dynamical forcings such as the propagation and dissipation of waves must be considered to explain numerous observational facts.

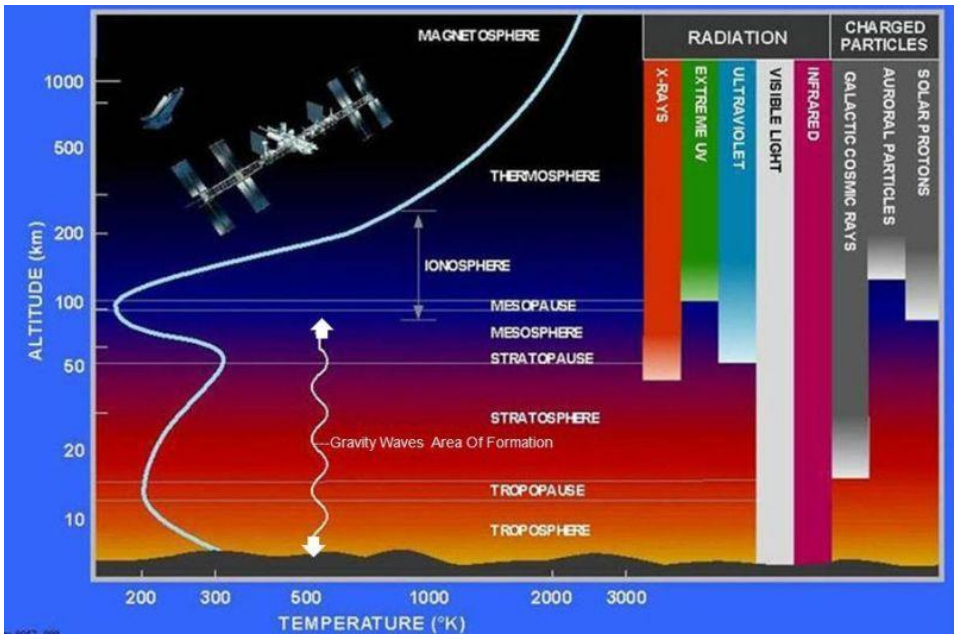

Figure 1.1. The vertical division of the atmosphere according to the thermal structure. [Credit: Charles Welch, http://www.theozonehole.com/atmosphere.htm, The Ozone Hole].

Like the waves in the oceans, there are also waves in the atmosphere. Mainly according to their periods, spatial scales, and possible origins, waves in the atmosphere are categorized into planetary waves, tides (tidal waves), gravity waves, and acoustic waves, etc. as Figure 1.2 illustrates. Planetary waves induced primarily by the conservation of potential vorticity are influenced by the Coriolis force and the pressure gradient in the atmosphere. They can cause strong wind, temperature, and constituent variabilities with scales of the planet and periods of days [Madden, 2007; Forbes, 2013]. Tides mostly result from the absorption of solar radiation by water vapor, ozone, etc. They have periods of 24-h or its harmonics. Gravity waves can be exited in the atmosphere with a broad spectrum. Their frequencies fall into the interval between the inertial frequency and the buoyancy frequency [Fritts and Alexander, 2003]. The generation, propagation, and dissipation of gravity waves change the background mean circulation of the atmosphere and complicate the general picture of the dynamics of the atmosphere. In particular, waves and the background mean circulation have mutual interactions with each other. Background mean flows can influence the generation, propagation, and dissipation of waves, while waves can change or induce background mean flows. Planetary waves, tides, and gravity waves carry momentum and energy away from their source regions. Via the deposition of momentum and energy, they alter the background mean flow and temperature structures during their breaking and dissipation.

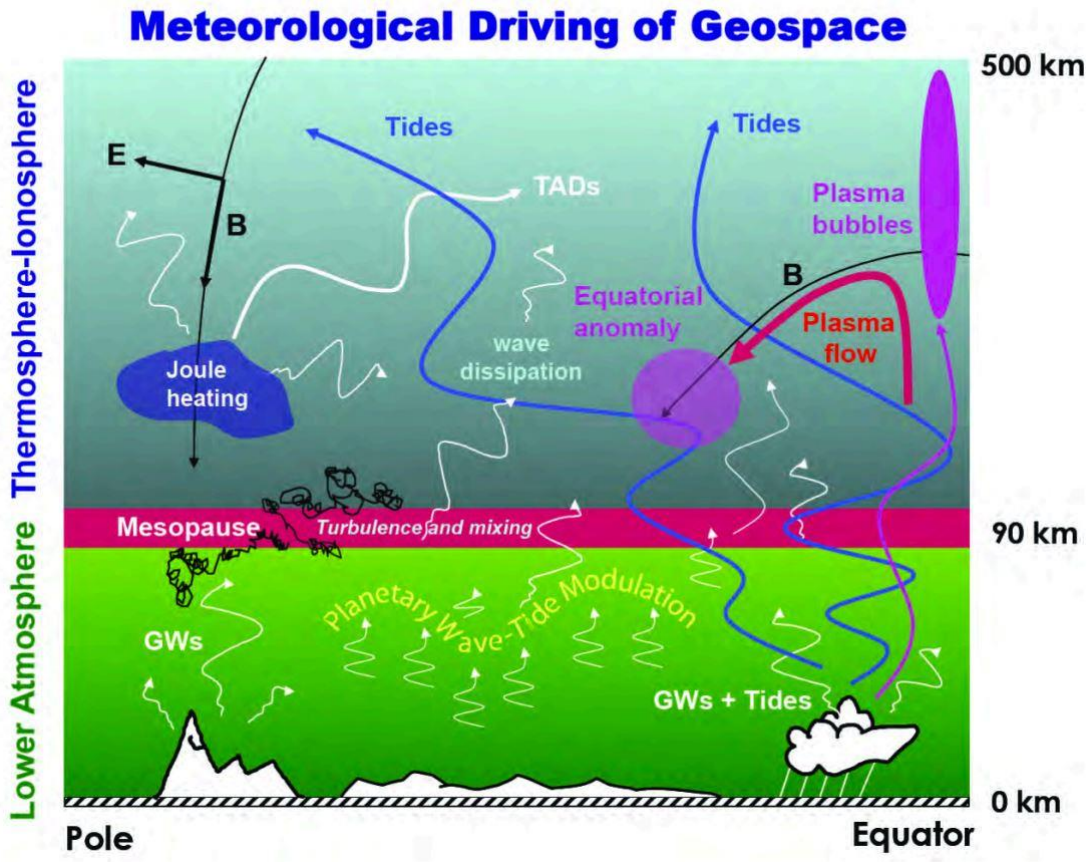

Figure 1.2. The generation, propagation, and dissipation of waves and other dynamics in the SAIR. [Credit: National Research Council (2013), Solar and Space Physics].

1.1.2. Significance of Gravity Waves in the SAIR

Among the atmospheric waves, large-scale waves such as planetary waves and tides are reasonably understood by the community as they can be physically resolved by General Circulation Models (GCMs) and Chemistry Climate Models (CCMs). Gravity waves are one of the least understood components due to their small scales, broad spectra, and the difficulties in observing them with our current instruments.

Gravity waves result from the vertical bulk displacement of air parcels in a stably stratified atmosphere. The restoring force is the resultant force of buoyancy and gravity. A general picture of gravity wave propagation was depicted by *C. Hines* in the 1960s for waves generated at lower altitudes of the atmosphere with a range

of spectra. The effects of amplification, reflection, intermodulation, and dissipation act to change the wave spectra continuously with increasing altitude, and result in different dominant modes at different altitudes [Hines, 1960; 1964; 1974]. There are a variety of gravity wave sources such as wind flow over topography, deep convection, wind shears, jet streams, geostrophic adjustment, local body forces created by primary wave breaking, and wave-wave interactions [e.g., Zhu and Holton, 1987; Alexander and Pfister, 1995; Vadas et al., 2003; Fritts et al., 2006; Alexander and Holton, 2004; Sato and Yoshiki, 2008; Vadas et al., 2009; Liu et al., 2014; Mzé et al., 2014; Eckermann et al., 2016; Smith et al., 2016]. Primary gravity waves generated in the lower atmosphere propagate and transport momentum and energy upward in the atmosphere, thereby affecting the atmospheric thermal structure, circulation, and variability [e.g., Eckermann and Preusse, 1999; Watanabe et al., 2006; Alexander and Teitelbaum, 2007; Baumgaertner and McDonald, 2007; Vincent et al., 2007; Plougonven et al., 2008; Plougonven and Zhang, 2014]. The amplitudes of some of these gravity waves are small upon generation at the source region, but they grow exponentially with altitude due to the exponential decrease of the atmospheric density with altitude. Due to this exponential growth, even though primary gravity waves might have small amplitudes in the lower atmosphere where they are generated, they grow into relatively large-amplitude waves in the middle and upper atmosphere, and play a significant role in redistributing momentum and energy there [Wilson et al., 1991]. These primary gravity waves can break/saturate and dissipate energy into the atmosphere when encountering a critical level due to the background wind or experiencing convective/shear (dynamic) instabilities due to the exponential growth of their amplitudes [e.g., Lindzen, 1981; Holton, 1983; Whiteway et al., 1997; Pfenninger et al., 1999; Dornbrack et al., 1995]. It should be noted that a gravity wave will never actually reach its theoretical critical level where the background wind is equal to the wave's phase speed and have zero vertical wavelength in a real

atmosphere [Becker, 2012]. As a matter of fact, it will encounter instabilities and break, thereby dissipating momentum and energy into the background atmosphere below the critical level a few kilometers. Gravity waves can also be saturated once their amplitudes reach high values, which causes the wave fields to be strongly unstable [Lindzen, 1981]. Eddy diffusion in the lower and middle atmosphere (below the turbopause) and molecular diffusion and thermal conductivity in the upper atmosphere (above the turbopause) work together to inhibit the continuous exponential growth of gravity wave amplitudes. The processes of gravity wave breaking/saturation and the dissipation of energy into the mean flow likely involve the generation of turbulence or even smaller scale gravity waves, i.e., the cascade of energy from large scales to small scales [Lindzen, 1981; Becker, 2012]. Secondary gravity waves can also be excited during the time period when primary gravity waves deposit their momentum and energy into the background atmosphere and create local body forces (i.e., temporal and spatial localized accelerations of the mean flow) [Vadas et al., 2003; 2018]. These secondary gravity waves have much larger horizontal scales than those of the corresponding primary gravity waves. A portion of the exited secondary gravity waves will further transport momentum and energy to even higher altitudes because their small amplitudes and large spatial scales enable them to propagate large distances vertically before breaking. The other portion will propagate downward and transport energy and momentum downward in the atmosphere. However, their amplitudes will decrease rapidly due to the exponential increase of the atmospheric density as they propagate downward. Other mechanisms such as gravity wave reflection and ducting also play important roles in the momentum and energy budget within the atmosphere [e.g., Hines, 1960, 1964; Lindzen, 1981; Fritts and Alexander, 2003; Preusse et al., 2008]. In-situ generation of gravity waves in the stratosphere due to the adjustment of the geostrophically unbalanced polar night jet

is also discussed extensively in the literature [e.g., Sato et al., 1999; Zink and Vincent, 2001a, 2001b; Sato and Yoshiki, 2008].

In the SAIR region, gravity waves are known to play an essential role in influencing atmospheric circulation, thermal structures, composition, and variability through transporting momentum and energy among stratified layers from the troposphere to the thermosphere [e.g., Hines, 1960, 1964, 1974; Lindzen, 1981; Holton, 1982, 1983; Hitchman et al., 1989; Fritts and Alexander, 2003; Haynes, 2005]. Gravity waves are very important to the atmosphere because of the following reasons.

1) Gravity waves have high dynamical variability over time and space (i.e., in the frequency and wavenumber domain). The large variability of gravity waves can influence the dynamics and chemistry in the atmosphere to different extents, which can be difficult to quantify and predict. There are established theories that explain the spectral shapes and slopes of various gravity wave induced atmospheric parameters (e.g., temperature, winds), which include Linear Instability Theory, Diffusive Filtering Theory, Saturated Cascade Theory, Diffusive Damping Theory, and Doppler Spreading Theory [Gardner, 1996; Mitchell and Howells, 1998; Chen et al., 2016b]. Gardner [1996] compared different theories and gave predictions of spectral shapes and slopes of various atmospheric parameters in terms of the frequency and wave numbers. Although generally there are so called "universal" spectra for different atmospheric parameters, the specific detailed spectral shapes and slopes still highly depend on the geolocations, altitudes, and seasons according to a variety of observations. For example, in terms of frequency spectra of the vertical wind, Chen et al. [2016b] found -0.4±0.1 for Boulder from 85 to 105 km; Tsuda et al. [1990] found -5/3 for Japan from 60 to 90 km; Sato et al. [2017] found -1 for Syowa station from 81 to 93 km; Zhang et al. [2017] found -0.43 and -0.31 in winter and summer stratosphere, -0.62 and -0.60 in winter and summer troposphere at Miramar NAS;

Gardner et al. [1995] found near 0 for the Na layer altitudes at Maui, Hawaii. The results obviously depend on the altitude, the data analysis procedures, seasons, and geolocation etc. Hence, it is important to note that the high dynamical variability of gravity waves over time and space play a significant role in atmospheric general circulations.

2) Gravity waves are capable of decelerating the polar night jet and inducing zonal wind reversals in the upper stratosphere and mesosphere via the deposition of energy and momentum [Watanabe et al., 2006; Alexander et al., 2009]. When gravity waves saturate or break and deposit momentum and energy into the background mean flow, the forcing exerted on the background mean flow is called gravity wave drag. The Eliassen-Palm theorem states that gravity wave drag accelerates/decelerates the mean flow towards the phase velocity of the corresponding gravity wave [Eliassen and Palm, 1961]. This effect along with critical level filtering partially explains the zonal mean wind structure (i.e., the zonal wind reversal) and the deceleration of the polar night jet in the upper atmosphere. Generally, as is shown in Figure 1.3, in the stratosphere and the mesosphere, an eastward jet stream (polar vortex or polar night jet) exists in the winter hemisphere and a westward jet stream exists in the summer hemisphere. In the winter hemisphere, due to the critical level filtering effect of the background eastward jet stream in the stratosphere, most of the gravity waves that reach the mesosphere are westward propagating. When such westward propagating gravity waves break in the winter mesosphere, the induced gravity wave drag decelerates the eastward jet. Similar effects take place in the summer hemisphere and the jet streams in the mesosphere in both hemispheres are weakened by gravity wave drag. Hence, we observe the deceleration of the polar night jet and the zonal wind reversals in the upper atmosphere.

3) Gravity wave drag can also cause the polar winter stratosphere and summer mesopause to deviate from radiative equilibrium and induce a residual mean meridional circulation from the summer pole to the winter pole [Lindzen, 1981; Holton, 1982; Garcia and Solomon, 1985; Garcia and Boville, 1994; Alexander *et al.*, 2009; Becker, 2012]. Needless to say, the residual mean meridional circulation plays a vital role in the momentum and energy balance in the Earth's middle and upper atmosphere due to its global scale.
4) Polar gravity waves are also expected to have a strong impact on the composition and chemistry of the atmosphere because the temperature perturbations induced by these waves can alter the chemical reaction rates and, therefore, the concentrations of important atmospheric constituents such as ozone [Lee et al., 2014], metal species [Chu et al., 2011a, b; Chen et al., 2016a], and the formation of polar stratospheric clouds (PSCs) and polar mesospheric clouds (PMCs) [Steele et al., 1983; McDonald et al., 2009; Chu et al., 2009, 2011a; Chen et al., 2016a]. PSCs appear in the polar winter stratosphere. They are responsible for the Antarctic ozone hole due to their significant roles in catalytic ozone destruction. PMCs are observed in the polar summer mesopause region where the temperature is often the coldest on the Earth. PSCs and PMCs are both highly sensitive to variations of the atmospheric water vapor content and temperature. Thus, they are potential indicators for monitoring the global climate changes.
5) When gravity waves propagate into the ionosphere, the induced perturbations can create oscillations in the electron density in that region. This stunning phenomenon is referred as traveling ionospheric disturbances (TID) [e.g., Klostermeyer 1972; Hocke and Schlegel, 1996; Liu and Vadas, 2013]. When such events take place, the propagation of radio waves can be disrupted, and radio communications and GPS navigations can be degraded to various degrees [Hocke and Schlegel, 1996].

6) Gravity waves are often parameterized in General Circulation Models (GCMs) and Chemistry Climate Models (CCMs) due to their sub-grid scales in these models. However, poor gravity wave parameterizations in numerical weather prediction and climate research models still present a major issue which lead to discrepancies between model results and observations [e.g., Kim et al., 2003; McLandress, 2006; Alexander et al., 2010; Richter et al., 2010; Garcia et al., 2014; Mzé et al., 2014]. For example, most GCMs and CCMs have the cold-pole problem, i.e., the simulated winter polar temperatures in the stratosphere are much colder than observed [e.g., Hamilton et al., 1999; Mclandress et al., 2012]. Numerical simulations have indicated missing gravity wave drag in the southern stratosphere as a possible cause [Mclandress et al., 2012; Tan, 2012; Alexander and Grimsdell, 2013]. The full parameterization of gravity waves in GCMs and CCMs is still not perfect owing to limited observations, especially in polar regions due to the harsh environment. In terms of high-resolution GCMs such as the Kühlungsborn Mechanistic Circulation Model (KMCM), which resolves gravity waves explicitly, observations can be used to validate the model and greatly improve model implementations [Watanabe et al., 2006; Becker, 2012, 2017; Becker and Vadas, 2018]. Thus, it is crucial to characterize gravity waves from a variety of observations, especially in the polar regions. Such observations will help improve our understanding of gravity wave characteristics and place parameterizations on a physical basis, which will help improve the diagnostic and predictive capabilities of the current numerical models.

Therefore, gravity waves are key elements in establishing atmospheric large-scale circulation, influencing atmospheric thermal structures, and coupling different atmospheric vertical layers.

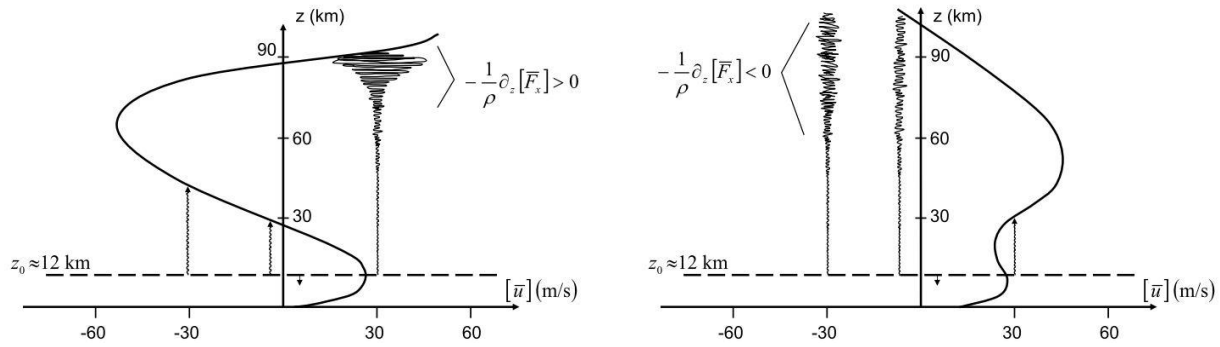

Figure 1.3. The zonal mean wind structure in summer (left panel) and winter (right panel) and the gravity wave induced deceleration of the polar night jet and zonal wind reversals. [Becker, 2012].

1.2. Observational Studies of Gravity Waves

Due to the importance of gravity waves in the SAIR, research in characterizing gravity wave features, determining the wave source and propagation, evaluating the impacts of gravity waves, etc. has been ongoing for many decades; yet large uncertainties and discrepancies still exist. The quantification of gravity wave parameters (such as wavelengths, periods, phase speeds, group velocities, potential and kinetic energy densities, and momentum fluxes) remains a challenge due to the limited number of ground-based observational sites and the limited spatial and temporal resolution of satellite observations [Alexander et al., 2011]. Scientists have developed many types of instruments to detect gravity waves in the atmosphere. Because of the inherent restrictions from different observational instruments, a wide spectra of gravity waves have to be studied with a combination of various observational techniques. In this section, we briefly summarize the observations of gravity waves (mainly in the Antarctica) in terms of various detection techniques.

1) Balloon Radiosonde

Balloon radiosonde refers to the sounding instruments carried by balloons into the atmosphere. It is usually battery powered and is capable of recording atmospheric temperature, horizontal winds, pressure, humidity, and more. The data is

transmitted back to the ground via radio waves. Because of the burst altitude of the balloons, radiosondes focus on gravity wave detection below 25 km. Radiosonde balloons are usually launched twice a day, leaving the observations with poor temporal resolution. Hence, the analysis of radiosonde data mainly focuses on the dimension of altitude. Polynomial fittings performed on the observed vertical profiles are usually regarded as the background, and the remaining perturbations are considered the gravity wave components.

From radiosonde observations, gravity wave activity in the lower stratosphere are investigated nearly globally. In the Northern Hemisphere, gravity wave activity is characterized as most intense in the winter polar regions, whereas at mid-latitudes, such activity reaches a maximum in either the winter or spring [Kitamura and Horita, 1989, Yoshiki and Sato, 2000]. In the southern atmosphere, gravity wave potential energy density becomes the highest in the winter at mid-latitudes and during equinox in the polar regions [Allen and Vincent, 1995; Yoshiki and Sato, 2000; Zink and Vincent, 2001; Moffat-Griffin et al., 2011, 2013]. The hodographs derived from radiosonde-observed horizontal winds usually indicate more upward propagation of gravity waves, although Murphy et al. [2014] found 50% of gravity waves propagating downward from 13 to 30 km, which might be explained by secondary gravity wave generation. Interestingly, in the Antarctic winter and spring, downward propagating gravity waves increase with respect to other seasons, indicating a potential wave source in the stratosphere related to the formation of the polar vortex [Yoshiki and Sato, 2000; Moffat-Griffin et al., 2011, 2013]. In Antarctica, two historical super-pressure balloon campaigns, VORCORE and CONCOREDIASI, were designed to detect lower stratospheric gravity waves around or below 20 km [Hertzog et al., 2007; Rabier et al., 2012]. Momentum fluxes were derived from these balloon radiosonde measurements with lower (1 h or greater) and higher (12 min) resolutions for VORCORE and CONCOREDIASI, respectively [Hertzog et al., 2008; Walterscheid et

al., 2016]. The average peak flux over the Antarctica Peninsula is about –375 mPa during the CONCOREDIASI campaign, which is ~10 times the campaign-averaged zonal flux for the VORCORE campaign [Walterscheid et al., 2016].

2) Airglow Imager

Airglow is the light emitted by trace gases in the mesopause region due to the complicated chemistry. Several airglow trace gases utilized to observe the atmosphere are hydroxyl (OH, ~ 87 km), molecular oxygen (O_2, ~94 km), and atomic oxygen (O(I), ~96 km and ~250 km). These layers all have several to tens kilometers of extension in the vertical. The gravity wave induced temperature perturbations can affect the altitude and brightness of such layers. All-sky airglow cameras can observe the optical emissions of the airglow layers in the mesopause region. Horizontal wavelengths and periods of gravity waves of various scales can be studied with such datasets. The first observation of the airglow in the sky was pioneered in 1973 via the infrared photography of OH [Peterson and Kiefabber, 1973]. Such observations were soon utilized to analyze the wave structures in the mesosphere in 1977 at the French Alps. Internal gravity waves were proposed for causing the band-like waves in the photograph [Moreels and Herse, 1977]. Following these pioneering works, airglow emissions from other trace gases (wavelengths) were utilized to study the dynamics in the mesopause region [Taylor et al, 1987]. Later on, with the employment of sensitive, solid state CCD cameras, an incredible improvement in the data quality was achieved and detailed gravity wave dynamics were characterized.

The Antarctica Gravity Wave Instrument Network (ANGWIN) was initiated in 2011 to unite the international airglow imagers across the Antarctic continent. The network was formed to address the source, propagation, and dissipation of a broad spectrum of mesospheric gravity waves, and their effects on the general circulation for this continental scale. Horizontal gravity wave parameters such as horizontal wavelength, ground-relative period, ground-relative phase speed, wave propagation

direction, occurrence, and duration can be quantified from such observations. Matsuda et al. [2014] obtained the power spectra of airglow intensity variation in the horizontal phase velocity domain via a new spectral analysis method utilizing a three-dimensional Fourier transform method. The averaged phase velocity distributions at Syowa, Halley, Davis, and McMurdo are used to characterize the preferential propagation direction of the gravity waves [Matsuda et al., 2017]. Gravity waves tend to propagate west in the mesosphere above Syowa, McMurdo, and Halley. This is possibly related to the critical level filtering by the background wind field due to the formation of the polar vortex. Similar explanations are also given to explain the preferential propagation direction of gravity waves at Mountain Research Station near Nederland, Colorado [Taylor et al., 1993]. However, it is found that the gravity waves above Davis propagate into every direction uniformly. This finding indicates that such gravity waves may be generated above the stratosphere, where the critical level filtering is negligible. Secondary gravity wave generation could potentially account for the source. In terms of the study searching for the gravity wave sources, tropospheric sources such as thunderstorms and *in-situ* sources near the mesopause are discussed in combination with observations [Taylor et al, 1991; 1995].

3) Radar

Radars can be employed to study the gravity wave dynamics in the atmosphere via the observations of wind velocities based on the Doppler shift theory. Gravity wave basic properties, kinetic energy, momentum flux, propagation direction, etc. can be investigated via such datasets. Various radars (~0.3 to 3000 MHz) are utilized to monitor gravity waves in the atmosphere from 10 to 30 km and from 60 to 90 km through Bragg scattering from weak permittivity fluctuations and through reflections from sharp changes in permittivity [Rastogi, 1981]. The major causes that influence the permittivity in the atmosphere with height is the variation of humidity in the troposphere, the temperature in the stratosphere, the number density of free

electrons in the mesosphere. UHF radar (typical frequency of 300 to 3000 MHz, wavelength < 1m) such as Incoherent Scatter Radars have been used for observing the ionosphere for decades [e.g., Santos et al., 2011; Martinis et al., 2013]. VHF radar (typical frequency of 30 to 300 MHz, wavelength of ~6 m) can simultaneously observe the mesosphere, stratosphere, troposphere through receiving echoes by scattering or reflections from permittivity fluctuations. Hence, some of them are called MST radar like the Jicamarca radar. HF and MF radar (typical frequency of 3 to 30 MHz for HF and 0.3 to 3 MHz for MF, wavelength of ~100 m) can observe the mesosphere region due to the existence of free electrons. The region of 35 to 55 km is invisible to radio waves due to the decline in permittivity fluctuations. Due to the lack of free electrons during the night-time, some mesospheric radar observations are limited to day time only. In the upper mesosphere from 80 to 95 km, ionized trails from incoming meteors provides reflection for VHF radio waves. Meteor radars were thus invented to monitor the dynamics of the upper mesosphere from this phenomenon.

Nakamura et al. [1996] performed a study utilizing radar data from 3 observational sites at mid-latitudes in both hemispheres. Semiannual variation of the gravity wave kinetic energy with maxima during solstice and minima during equinox were found in the MLT. This statement is generally true in both hemispheres [Vincent and Fritts, 1987; Manson and Meek, 1993]. Momentum flux can also be derived from radar measured wind observations. Research from Tsuda et al. [1990b], Manson and Meek [1993], Nakamura et al. [1993] etc. show that the zonal momentum flux of gravity waves has an apparent annual cycle and is eastward in the summer and westward in the winter, and is anti-correlated with the background wind direction in the MLT. The meridional momentum flux does not have a yearly cycle, but has similar amplitudes compared to the zonal momentum flux. Beldon and Mitchell [2009] revealed the characteristics of polar gravity waves with periods shorter than 3 hr at altitudes of 80 to 100 km utilizing meteor radar data at Rothera

(68° S) in the Antarctic and Esrange (68° N) in the Arctic. Gravity wave activity increases with height in all seasons except below 90 km in the summer months. The previously mentioned semi-annual seasonal cycle in gravity wave activity was also observed; wave activity is strongest during the solstices and is weakest during the equinoxes. Gravity wave activity in the lower troposphere are observed with seasonal cycles at Davis, Antarctica via VHF radar wind measurements [Alexander et al., 2013]. The horizontal and vertical wind variances indicate maximum gravity wave activity in the winter and minimum gravity wave activity in the summer; this is related to the orographic gravity waves generated by the synoptic winds and an ice ridgeline. A similar gravity wave trend is also observed in the stratosphere, which is related to the formation of the polar vortex [Tsuda et al., 1994]. At Syowa station, the PANSY radar is monitoring the gravity waves from the troposphere to the mesosphere [Mihalikova et al., 2016]. Inertial gravity waves are observed with smaller intrinsic frequencies in the summer and larger intrinsic frequencies in the winter. Downward-propagating gravity waves are observed with higher rates in the winter, which is consistent with the above-mentioned radiosonde observations and is related to *in-situ* wave sources in the middle atmosphere.

4) Rockets

Typically, rocket sounding can record vertical profiles of the atmospheric temperature and horizontal winds along the flight, which enable gravity wave analysis. Long-term rocket measurements can be used to characterize the seasonal and latitudinal variations of gravity waves in the stratosphere. It is generally found that gravity wave activity exhibits a strong annual cycle with a maximum during the winter and a minimum during the summer at high latitudes, and exhibits a semiannual cycle with maxima at the equinox at mid latitudes from 20 to 65 km [Hirota, 1984; Eckermann et al., 1995]. At low latitudes, there are no obvious seasonal variations of gravity wave activity [Eckermann et al., 1995]. The upward propagation

of gravity waves is indicated via hodographs from the rocket observed horizontal winds, which agrees with the results from radiosondes and radars [Hirota and Niki, 1985; Hamilton, 1991]. At mid-latitudes, observed gravity waves have significant westward propagation direction during the winter and eastward propagation direction during the summer, which possibly results from the prevailing background wind [Hamilton, 1991].

5) Satellite-borne Instruments

Undoubtedly, satellites can provide global coverage of gravity wave measurements with relatively frequent time intervals, but it is difficult to infer gravity wave information at very high latitudes near the poles due to the analysis technique and satellite viewing geometry. Satellites also have difficulty resolving small temporal and spatial scale gravity waves. Generally speaking, limb-viewing satellites yield better vertical resolution and nadir-viewing satellites yield better horizontal resolution. Satellites like Limb Infrared Monitor of the Stratosphere (LIMS) instrument, Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA), Microwave Limb Sounder (MLS), and the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) have been utilized to monitor gravity waves in the stratosphere and mesosphere for decades.

Observations from the Atmospheric Infrared Sounder (AIRS) onboard the National Aeronautics and Space Administration Aqua satellite and the Infrared Atmospheric Sounding Interferometer (IASI) onboard the European MetOp satellites have been combined to study stratospheric gravity waves by Hoffmann *et al.* [2014]. AIRS and IASI are capable of observing stratospheric gravity waves with long vertical and short horizontal wavelengths. AIRS and IASI observe these waves most frequently at mid- and high latitudes during the winter, where they are related to orographic sources and the polar night jet, and at low and mid-latitudes during the summer, where convective sources play an important role. Satellite AIRS

observations provide extensive datasets for global studies of stratospheric gravity waves. Alexander and Barnett [2007] discussed how satellite observations such as AIRS could be used to constrain parameterizations of gravity waves in global models. A statistical study based on AIRS data for the region near the Patagonian Andes of South America and northernmost Antarctic Peninsula during September 2003 reveals that horizontal wavelengths range from ~30–500 km with a peak of ~100 km, while wave propagation angles from the background horizontal wind range from ~90–270° with a peak of ~190° [Alexander and Barnett, 2007]. Hoffmann and Alexander [2009] developed a data retrieval scheme for AIRS that made it feasible to study small scale gravity waves with vertical wavelengths as short as 20 km. Gong et al. [2012] studied the climatology of the gravity-wave-induced temperature variances at pressure levels from 2 to 100 hPa (~16–41 km) with AIRS measurements, and demonstrated that AIRS is capable of detecting gravity waves with shorter vertical wavelengths (~12 km). Efforts were also spent on forming a global view of stratospheric gravity wave hotspots and discriminating gravity waves between orographic and non-orographic sources [Hoffmann et al., 2013; Hoffmann et al., 2016]. Utilizing (CHAMP/GPS) radio occultation data, significant seasonal variations of gravity wave potential energy density with winter maxima were found and attributed to the seasonal variation of critical level filtering and the Doppler shift effect by the mean winds (strong background wind causes Doppler shifting of gravity waves to longer vertical wavelengths, where they can reach larger amplitudes before reaching saturation) over Antarctica Peninsula around 15 to 35 km [Baumgaertner and McDonald, 2007]. Wave activity enhancements were also found around the edge of the polar vortex. Tsuda et al. [2000] determined monthly mean values of gravity wave potential energy density at 15 to 20 km altitude around Japan and found a seasonal variation with an enhancement during the winter months using GPS occultation data. However, the difficulty in interpreting satellite observations of gravity wave activity

and their coarse vertical resolutions suggest that examining the region between 30 and 200 km require detailed measurements by ground-based instruments.

6) Lidar

Lidar (LIght Detection And Ranging) is a powerful remote-sensing instrument which is capable of providing measurements (temperatures, winds, constituent densities, etc.) with unprecedented temporal and spatial resolutions in the atmosphere, and has enabled various cutting-edge findings. Chanin and Hauchecorne [1981] studied gravity waves using atmospheric density and temperature data through Rayleigh lidar observation for the first time. Apart from observing gravity waves, other phenomena in the atmosphere such as tides, planetary waves, heat flux, momentum flux, energy flux, atmospheric instabilities, thermal structures, thermospheric metal layers, and polar mesospheric clouds etc. have been studied extensively from various lidar campaigns [e.g., Gardner et al., 1995, 2002; Lübken and von Zahn, 1991; Senft et al., 1994; Fong et al., 2014, 2015; Chen et al., 2013, 2016a, 2016b; Lu et al., 2013, 2015a, 2015b; Chu et al., 2002, 2003, 2011a, 2011b]. Resonance fluorescence lidar [Chu and Papen, 2005] fits in the atmospheric research community because of its capability to retrieve range-resolved atmospheric parameters over a long altitude range (from ~ 15 km to as high as 170 km) [Chu et al., 2011b, 2016]. The detection of the stratosphere and lower mesosphere is enabled by Rayleigh scattering of the atmospheric molecules while the detection of the Mesosphere and Lower Thermosphere (MLT) is made possible by the resonance fluorescence from a layer of metal atoms in that region. This layer of neutral atomic metals (Na, Fe, K, Ca, etc.) are ultimately created from the sputtering and ablation of the incoming meteors to the Earth. Within and above this altitude range, these neutral metal atoms will be ionized due to situations such as solar radiation or energetic electrons. Whereas below this altitude range, the neutral metal atoms will generally be oxidized to form more complicated compounds.

Most observations of gravity waves over Antarctica made from various instruments such as radar, balloon radiosonde, airglow imager, and GPS radio occultation focused on the region below 30 km or above 80 km [e.g., Collins et al., 1994, 1996; Collins and Gardner, 1995; Moffat-Griffin et al., 2011, 2013; Espy et al., 2004, 2006; Hibbins et al., 2007; Tsuda et al., 2000; Baumgaertner and McDonald, 2007; Yoshiki and Sato, 2000]. Lidar observations are suitable to fill this observational gap and further extend the altitude coverage. In polar regions, researchers have characterized the gravity wave signatures at the South Pole (90°S), McMurdo (78°S), Davis (69°S), Syowa (69°S), and Rothera (67.5°S) from various lidar campaigns [Collins et al., 1994, 1996; Collins and Gardner, 1995; Yamashita et al., 2009; Alexander et al., 2011; Chu et al., 2011, 2018; Kaifler et al., 2015; Zhao et al., 2017]. Case studies and statistical studies have been performed. Yamashita et al. [2009] characterized stratospheric gravity wave activities (in terms of basic gravity wave parameters such as vertical wavelengths, periods, vertical phase speeds, etc.) and compared the potential energy densities (E_{pm}) between Rothera (67.5°S, 68.0°W) and the South Pole (90°S). E_{pm} showed very different seasonal variations at these two sites. At Rothera, E_{pm} variations are characterized by maximum values in the winter and minimum values in the summer, while E_{pm} remains roughly constant throughout the year at the South Pole. Yamashita et al. [2009] explained the results at Rothera via selective critical-level filtering of orographic waves and the polar night jet being a possible new wave source during the winter. Following this study, gravity wave activity in the upper stratosphere and lower mesosphere was investigated using temperature data retrieved from Rayleigh lidars at Davis [Alexander et al., 2011; Kaifler et al., 2015]. A similar seasonal behavior of E_{pm} as observed at Rothera was reported for Davis. Rayleigh/Raman lidar observations of gravity wave activity from 15 to 70 km altitude over Syowa station is discussed in detail by Kogure et al. [2017] excluding summertime. Above 30 km altitude, gravity wave potential energy density

was larger in the winter than in the spring and fall. Gravity wave potential energy densities were found to have high correlations with tropospheric wind, stratospheric wind, critical layer filtering, and the position of the polar vortex using measurements from Eureka Weather Station (80° N, 86° W) [Whiteway et al., 1997; Duck et al., 2001]. However, there was a big latitudinal gap of observations between the South Pole and Antarctic Circle. It was unknown how the stratospheric gravity waves behave in this region.

The University of Colorado Boulder lidar group deployed the Fe Boltzmann Temperature Lidar to McMurdo Station, Antarctica (77.85° S, 166.67° E) in December 2010 to fill in this observational gap. A long-lasting observational campaign has been going on around the clock since then, weather permitted. Hence, McMurdo lidar observations near 80°S are very important. This lidar has the capability of monitoring gravity waves in the stratosphere, mesosphere, and lower thermosphere [e.g., Chu et al., 2011a,b; Chen et al., 2013, 2016; Lu et al., 2015], providing extremely valuable long-term datasets. Many remarkable scientific findings have resulted from this unique observational dataset. To name a few, the discovery of the MLT persistent gravity waves (Figure 1.4) [Chen et al., 2013, 2016] and the discovery of the thermospheric Fe layers that exhibit clear signatures of 1.5-hr fast gravity waves with downward phase progression (Figure 1.5) [Chu et al., 2011a, 2011b]. The MLT persistent gravity waves refers to the perpetually-existing, dominant, large-amplitude (20 to 30 K amplitude in temperature) gravity waves with periods of 3 to 10 hr and vertical wavelengths of 20 to 30 km as shown in the MLT region in Figure 1.4. The word "persistent" describes the fact that these waves can exist continuously for a prolonged period. Applying wavelet analysis methods, we identify single wave events that last from 10 hr to 50 hr. However, if we do not separate the waves into individual wave events with individual periods/wavelengths, but regard these waves as a group of waves with periods ranging from 3 to 10 hr, then

this wave group is perpetual. That is, these waves appear in every lidar observation, and occur frequently enough to be regarded as endless and uninterrupted [Chen, PhD dissertation, 2016]. It is surprising that these waves are so dominant and so strong in the McMurdo MLT because of presumed filtering by the polar night jet, which suggests a significant gap in our understanding of the wave sources in the SAIR.

The thermospheric Fe layers (Figure 1.5, Panel (a)) were reported for the first time in lidar observations of neutral Fe layers having 1.5 hr fast gravity waves that reached 155 km at McMurdo (77.8°S, 166.7°E), Antarctica [Chu et al. 2011b]. It was the first time that gravity wave signatures (with periods of ~ 1.57 hr) were observed in the thermospheric metal layers having downward phase progression at such high altitudes. The discovery of such neutral metal layers reaching nearly 200 km in the thermosphere have significant scientific importance because these discoveries challenge the current understanding of the upper atmospheric composition, chemistry, dynamics, electrodynamics and energetics via the space-atmosphere interactions [Chu and Yu, 2017]. These neutral metal layers provide tracers for direct measurement of the neutral properties in this little understood but crucially important space-atmosphere interaction region in the altitude range of 100–200 km. As is shown in panel (b) and (c) of Figure 1.5, these metal layers make direct and range-resolved measurements of the neutral temperatures in the thermosphere at altitudes from 100 to 200 km possible, as has been demonstrated up to 170 km at McMurdo using the Fe Boltzmann lidar technique [Chu et al., 2011b, 2013, 2016]. With modern Doppler lidar techniques, it is anticipated that neutral winds in the thermosphere can also be directly measured using the resonance fluorescence excited from such metal layers. In fact, some preliminary results of neutral winds have been demonstrated up to ~140 km using the Na Doppler lidar technique at Cerro Pachón [Liu et al., 2016]. The University of Colorado Boulder lidar group also deployed a Na Doppler lidar at McMurdo station in the same laboratory beside the Fe Boltzmann

lidar in December 2017. The campaign has generated preliminary results of the temperature, vertical wind, and Na layer observations in the McMurdo MLT. As the campaign continues, it is likely that more exciting findings will be discovered. It is reported that such thermospheric metal layer is not restricted to polar regions, but extends to mid- and low-latitude sites as well [Friedman et al., 2013; Gao et al., 2015]. Properly utilizing these tracers will enable us to open up a new window into the structures and dynamics of the SAIR via ground-based instruments. The 1.5 h gravity waves with obvious downward phase progression in the TIFe layers are crucial to TIFe formation. According to model simulations on the mechanism study, it is the vertical shears induced by gravity wave winds that enable the convergence of Fe^+ such that the high-density Fe^+ layer can be neutralized to form neutral Fe via direct recombination with electrons [Chu and Yu, 2017]. Basically, it is found that the 1.5 h fast gravity wave determines the overall shape of the TIFe layer. Hence, investigating the characteristics and sources of these waves will help us better understand the dynamics in the thermosphere and improve the model simulations of TIFe events.

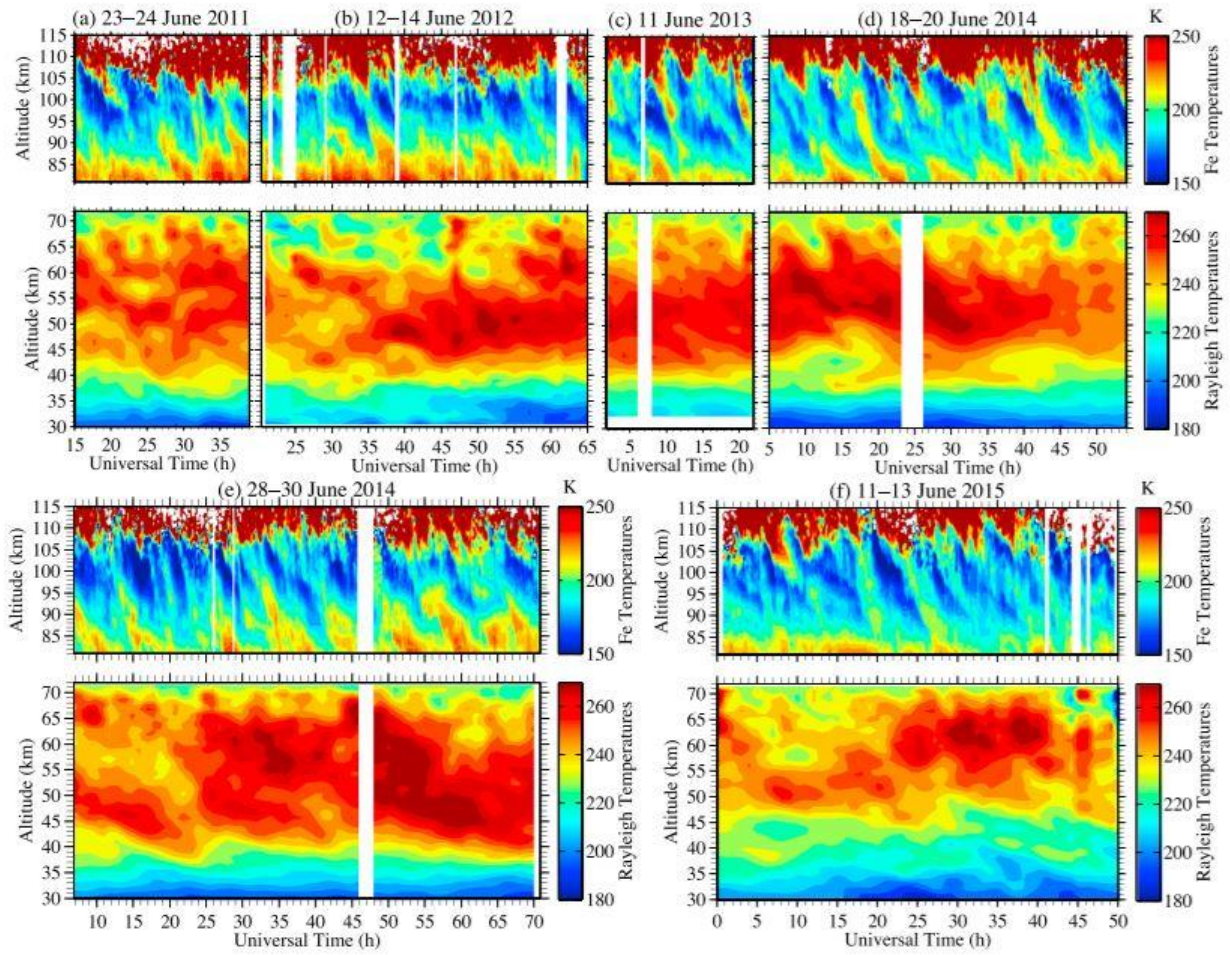

Figure 1.4. The lidar-discovered persistent gravity wave in the McMurdo MLT (80 to 115 km) temperature observations on various days. The temperature observations in the stratosphere (30 to 72 km) are also provided [Chen et al., 2016].

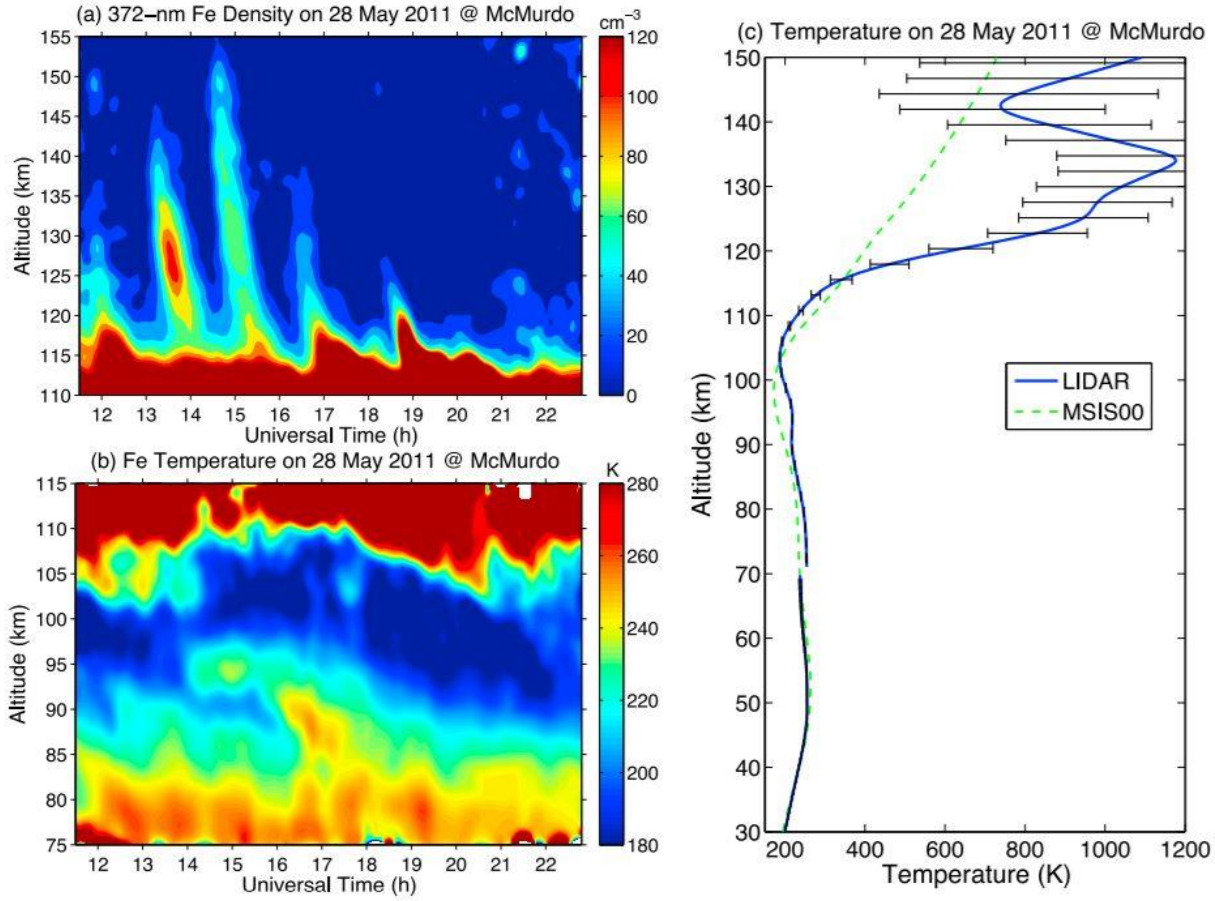

Figure 1.5. The lidar-discovered thermospheric Fe layers (a) and the retrieved temperature observations (b) in the McMurdo MLT. (c) illustrates a continuous temperature observation from the stratosphere all the way up to 150 km. The data is retrieved from Fe Boltzmann lidar observation on 28 May 2011 [Chu et al., 2011b].

1.3. Theories of Gravity Waves

1.3.1. Primary Gravity Wave Theory

The linear theory is usually invoked to describe gravity waves. In such a theory, we treat the wave induced atmospheric perturbations as small departures from a stably stratified background state. The parameters describing the background state usually vary only in the vertical. Here we briefly review this fundamental gravity wave theory. The following contents within this section are adapted from the classic

gravity wave papers Fritts and Alexander [2003, 2012], Vadas and Fritts [2005], and Vadas [2013]. In the Cartesian coordinates following the motion of particular fluid parcels [Holton, 2004], the compressible fluid equations that describe the conservation of momentum, mass, and energy are as follows:

$$\frac{du}{dt} - fv + \frac{1}{\rho}\frac{\partial p}{\partial x} = X \tag{1.1}$$

$$\frac{dv}{dt} + fu + \frac{1}{\rho}\frac{\partial p}{\partial y} = Y \tag{1.2}$$

$$\frac{dw}{dt} + \frac{1}{\rho}\frac{\partial p}{\partial z} + g = 0 \tag{1.3}$$

$$\frac{1}{\rho}\frac{d\rho}{dt} + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{1.4}$$

$$\frac{dT}{dt} + (\gamma - 1)T\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right) = Q \tag{1.5}$$

Here p is pressure, ρ is density, $\rho = \rho_0 \exp(-(z - z_0)/H)$ with $\rho_0 = \rho(z_0)$ being the density at a reference height z_0 and H being the scale height, d/dt is the total, or advective derivative, u, v, w are the fluid velocities in zonal, meridional, and vertical directions, $f = 2\Omega \sin\phi$ is the Coriolis parameter with Ω being the Earth rotational angular speed and ϕ being the latitude, $\gamma = \frac{c_p}{c_v}$ is the ratio of specific heats at constant pressure (c_p) and constant volume (c_v), X, Y, and Q are the unspecified forcings, heatings, and coolings.

Next, we linearize the above equations with respect to a stable background state. The unforced (X, Y, and Q equal to 0) form of the equations linearized about a horizontally uniform hydrostatic basic state with background wind ($\bar{u}, \bar{v}, 0$), temperature $\bar{T}$, pressure $\bar{p}$, and density $\bar{\rho}$ varying only in z are

$$\frac{Du'}{Dt} + w'\frac{\partial \bar{u}}{\partial z} - fv' + \frac{\partial}{\partial x}\left(\frac{p'}{\bar{\rho}}\right) = 0 \tag{1.7}$$

$$\frac{Dv'}{Dt} + w'\frac{\partial \bar{v}}{\partial z} + fu' + \frac{\partial}{\partial y}\left(\frac{p'}{\bar{\rho}}\right) = 0 \tag{1.8}$$

$$\frac{Dw'}{Dt} + \frac{\partial}{\partial z}\left(\frac{p'}{\bar{\rho}}\right) - \frac{1}{H}\left(\frac{p'}{\bar{\rho}}\right) + g\left(\frac{\rho'}{\bar{\rho}}\right) = 0 \tag{1.9}$$

$$\frac{DT'}{Dt} + (\gamma - 1)\bar{T}\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z}\right) = 0 \tag{1.11}$$

$$\frac{D}{Dt}\left(\frac{\rho'}{\bar{\rho}}\right) + \frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z} - \frac{w'}{H} = 0 \tag{1.12}$$

$$\frac{p'}{\bar{p}} = \frac{\rho'}{\bar{\rho}} + \frac{T'}{\bar{T}}, \tag{1.13}$$

where the derivative D/Dt is the linearized form of the time derivative,

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x} + \bar{v}\frac{\partial}{\partial y} \tag{1.14}$$

Quantities with a prime represent the wave induced perturbations with respect to the background state. Equation (1.11) is from equation (15) in Vadas and Fritts [2005] obtained by setting the kinematic viscosity to 0. Equation (1.13) is from Vadas [2013] (following equation (23) in that work). We neglect the background shear terms by assuming $(\bar{u}, \bar{v})$, H, and N vary only slowly over a vertical wavelength (the WKB approximation).

The gravity wave solutions are assumed to be,

$$\left(u', v', w', \frac{T'}{\bar{T}}, \frac{p'}{\bar{\rho}}, \frac{\rho'}{\bar{\rho}}\right) = \left(u_p, v_p, w_p, T_p, p_p, \rho_p\right) \bullet \exp\left[i\left(kx + ly + mz - \omega t\right) + \frac{z}{2H}\right] \tag{1.15}$$

This form describes a monochromatic wave perturbation with wave number components (k, l, m) and ground relative angular frequency ω. Substituting this assumed solution into equations (1.7) to (1.13) gives rise to six algebraic equations for

$\left(u_p, v_p, w_p, T_p, p_p, \rho_p\right)$. It is important to note that in Fritts and Alexander [2003], γ was set to infinity, which is unrealistic because γ is a finite number in reality (e.g., γ = 1.4 for diatomic molecules, representative of the stratosphere and mesosphere). This situation was corrected in Vadas [2013]. Therefore, here we use the results for finite γ .

$$-i\omega_I u_p - f v_p + ikp_p = 0 \tag{1.16}$$

$$-i\omega_I v_p + f u_p + ilp_p = 0 \tag{1.17}$$

$$-i\omega_I w_p + (im - \frac{1}{2H}) p_p = -g\rho_p \tag{1.18}$$

$$-i\omega_I p_p + [g(\gamma - 1) + iC_s^2(m + \frac{i}{2H})]w_p + iC_s^2(ku_p + lv_p) = 0 \tag{1.19}$$

$$-i\omega_I \rho_p + iku_p + ilv_p + (im - \frac{1}{2H})w_p = 0 \tag{1.20}$$

$$T_p = \gamma\, p_p / C_s^2 - \rho_p \tag{1.21}$$

Here $\omega_I = \omega - k\bar{u} - l\bar{v}$ is the intrinsic frequency (i.e., the frequency that would be observed in a frame of reference moving with the background wind $\left(\bar{u}, \bar{v}\right)$), $C_s = \sqrt{\gamma \frac{p}{\rho}}$ is the sound speed. With certain mathematical derivations, the above equations can be combined into a single equation for the vertical velocity amplitude. Setting the imaginary coefficients of this equation to equal zero yields:

$$\frac{g}{C_S^2} = \frac{1}{H} - \frac{N^2}{g} \tag{1.22}$$

$$\omega_I^2\left(k^2 + l^2 + m^2 + \frac{1}{4H^2} - \frac{\left(\omega_I^2 - f^2\right)}{C_s^2}\right) = N^2\left(k^2 + l^2\right) + f^2\left(m^2 + \frac{1}{4H^2}\right) \tag{1.23}$$

Here $N^2 = \frac{g}{\bar{T}}\left(\frac{d\bar{T}}{dz} + \frac{g}{c_p}\right)$ is the buoyancy frequency. Equation (1.23) supports both acoustic and gravity waves. By letting the sound speed $C_s \to \infty$, we obtain the gravity wave dispersion relation in terms of the intrinsic frequency ω_I ,

$$\omega_I^2 = \frac{N^2\left(k^2 + l^2\right) + f^2\left(m^2 + \frac{1}{4H^2}\right)}{k^2 + l^2 + m^2 + \frac{1}{4H^2}} \tag{1.24}$$

or alternatively for the vertical wave-number as

$$m^2 = \frac{\left(k^2 + l^2\right)\left(N^2 - \omega_I^2\right)}{\left(\omega_I^2 - f^2\right)} - \frac{1}{4H^2} \tag{1.25}$$

In this dispersion relation, the frequency of gravity waves is related to the wavenumbers (*k, l, m*) and to the background atmosphere properties such as the buoyancy frequency N and horizontal winds $\left(\bar{u}, \bar{v}\right)$. For a gravity wave that propagates vertically, the wavenumbers (*k, l, m*) are real, and the intrinsic frequency fall into the range $N > \omega_I > |f|$. While the gravity wave ground-relative frequency does not vary significantly unless $\frac{\partial \bar{u}}{\partial t} \neq 0$ or $\frac{\partial \bar{v}}{\partial t} \neq 0$, the intrinsic frequency and vertical wavenumber change accordingly as a wave propagates vertically through background wind shear and stability gradients. The gravity wave group velocities describing energy transport and wave packet propagation are,

$$\left(c_{gx}, c_{gy}, c_{gz}\right) = \left(\frac{\partial \omega}{\partial k}, \frac{\partial \omega}{\partial l}, \frac{\partial \omega}{\partial m}\right) = \left(\bar{u}, \bar{v}, 0\right) + \frac{\left[k\left(N^2 - \omega_I^2\right), l\left(N^2 - \omega_I^2\right), -m\left(\omega_I^2 - f^2\right)\right]}{\omega_I\left(k^2 + l^2 + m^2 + \frac{1}{4H^2}\right)} \tag{1.26}$$

It is important that we follow the sign convention during the derivation and application of the linear theory. Here ω_I is positive definite, wavenumber vector (*k, l*) denotes gravity wave horizontal propagation direction, vertical wavenumber *m* is

negative for upward group velocity and positive for downward. Hence, in this convention, for eastward ($k>0$) and upward ($m<0$) gravity wave energy propagation relative to the wind, the intrinsic phase speeds are eastward ($c_{Ix} = \omega_I / k$) and downward ($c_{Iz} = \omega_I / m$) relative to the wind.

From equation (1.16) to (1.21), we are able to derive polarization relations that relate the perturbation amplitudes and phases of different gravity wave parameters. Here are some examples,

$$u_p = \left(\frac{i\omega_I k - fl}{i\omega_I l + fk} \right) v_p \tag{1.27}$$

$$p_p = \left(\frac{\omega_I^2 - f^2}{\omega_I k + ifl} \right) u_p = \left(\frac{\omega_I^2 - f^2}{\omega_I l - ifk} \right) v_p \tag{1.28}$$

$$w_p = \frac{\left(-m - \frac{i}{2H} + \frac{i}{\gamma H} \right) \omega_I}{N^2 - \omega_I^2} p_p \tag{1.29}$$

$$T_p = \frac{-iN^2}{g\omega_I} w_p \tag{1.30}$$

Simplified but useful dispersion relations for high frequency ($\omega_I \gg f$), medium frequency ($N \gg \omega_I \gg f$), and low frequency ($\omega_I \sim f$) gravity waves with vertical wavelength $\lambda_z \ll 2\pi H$ are as follows,

$$\text{High Frequency: } \omega_I^2 = \frac{N^2(k^2 + l^2)}{k^2 + l^2 + m^2} = N^2 \cos^2 \alpha \tag{1.31}$$

$$\text{Medium Frequency: } \omega_I = N \left| \frac{k_h}{m} \right| \tag{1.32}$$

$$\text{Low Frequency: } \omega_I^2 = N^2 \frac{k_h^2}{m^2} + f^2 \tag{1.33}$$

Here α is the angle between lines of constant phase and the vertical, and $k_h = \sqrt{k^2 + l^2}$.

1.3.2. Secondary Gravity Wave Theory

We review the theory of secondary gravity wave generation based on the pioneering papers Vadas et al. [2003, 2018] and Vadas [2013]. The theory of secondary gravity wave generation was developed by Dr. Sharon Vadas via a series of work over 15 years. Because it is one of the most important theoretical bases for this dissertation, we recapitulate the mathematical derivations, adapting from the milestone paper Vadas et al. [2018].

We consider the secondary gravity waves excited by a localized (in space) and intermittent (in time) horizontal body force in a compressible fluid in the *f*-plane approximation. When primary gravity waves break, small-scale gravity waves and vortices are created from the large-amplitude, nonlinear interactions. The cascade of energy to smaller scales and eventually to turbulence then ensues. At this point, the momentum carried by the primary gravity waves is deposited into the background atmosphere on the spatial and temporal scales of the order of the primary gravity wave packet. During this momentum deposition, a local body force is created that accelerates the background flow horizontally in the direction that the primary gravity wave packet was propagating. Here we only focus on the linear effects that the resulting local body force has on scales comparable to or larger than the scales of the body force itself. In Cartesian coordinates, the fluid equations that govern the evolution of the flow are as follows,

$$\frac{D\vec{V}}{Dt}+\frac{1}{\rho}\nabla p-g\vec{e_z}+f\vec{e_z}\times\vec{V}=\vec{F}(x)\Psi(t) \tag{1.34}$$

$$\frac{D\rho}{Dt}+\rho\nabla\bullet\vec{V}=0 \tag{1.35}$$

$$\frac{DT}{Dt}+(\gamma-1)T\nabla\bullet\vec{V}=0 \tag{1.36}$$

Here $\vec{V}$ is the 3-dimensional velocity vector, and γ is the ratio of specific heats. The spatial portion of the body force is $\vec{F}(x) = (F_x(x), F_y(x), 0)$. The total zonal and meridional components of the body force are

$$F_{x,tot} = F_x(x)\Psi(t), F_{y,tot} = F_y(x)\Psi(t) \tag{1.37}$$

The temporal dependence of the total body force is given by the analytic function $\Psi(t)$.

Following linear theory, we decompose the flow variables as contributions from 1) the background flow (denoted with overlines), 2) the perturbations from the secondary gravity waves and from the induced mean response (denoted with primes) as following,

$$\begin{aligned} &u = \overline{U} + u', v = \overline{V} + v', w = w', \\ &\rho = \overline{\rho} + \rho', T = \overline{T} + T', p = \overline{p} + p', \end{aligned} \tag{1.38}$$

For simplicity, we regard $\overline{T}$, γ, $\overline{U}, \overline{V}$, and H (scale height) as constant. We linearize and solve these equations for the following smooth temporal function of the body force,

$$\begin{aligned} \Psi(t) &= \frac{1}{\chi}(1 - \cos \bar{a}t) \quad for\ 0 \le t \le \chi \\ &= 0 \qquad\qquad\quad for\ t \le 0\ and\ t \ge \chi \end{aligned} \tag{1.39}$$

Here $\Im$ has duration χ and frequency $\bar{a}$.

$$\bar{a} \equiv 2\pi n/\chi, n = 1, 2, 3... \tag{1.40}$$

Following the convention in Vadas [2013], we define

$$\begin{aligned} &\xi = e^{-z/2H}u', \sigma = e^{-z/2H}v', \eta = e^{-z/2H}w', \\ &\phi = e^{-z/2H}\,\rho'/\overline{\rho_0} = e^{-z/2H}\,\rho'/\overline{\rho}, \varphi = e^{-z/2H}\,p'/\overline{\rho_0} = e^{-z/2H}\,p'/\overline{\rho}, \\ &\varsigma = e^{-z/2H}\,T'/\overline{T_0}, F_{xs} = e^{-z/2H}F_x, F_{ys} = e^{-z/2H}F_y, \end{aligned} \tag{1.41}$$

The above parameters are expanded in a Fourier series such as,

$$\xi(x, y, z, t) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(kx+ly+mz)}\,\xi_{FFT}(k, l, m, t)\,dk\,dl\,dm \tag{1.42}$$

The subscript "*FFT*" denotes the Fourier transform of the variable. We then take the Laplace transform of the equations and solve them algebraically. After the force is finished, the mean and gravity wave solutions are,

$$\xi_{FFT,FH}(t) = \frac{i\bar{a}^{-2}l}{f}K + \frac{i\bar{a}^{-2}}{\chi(f^2-\omega_1^2)}\left[(kO+lfP)C(\omega_1)+(-k\omega_1 P+\frac{lfO}{\omega_1})S(\omega_1)\right] \quad (1.43)$$

$$\sigma_{FFT,FH}(t) = -\frac{i\bar{a}^{-2}k}{f}K + \frac{i\bar{a}^{-2}}{\chi(f^2-\omega_1^2)}\left[(lO-kfP)C(\omega_1)-(l\omega_1 P+\frac{kfO}{\omega_1})S(\omega_1)\right] \quad (1.44)$$

$$\eta_{FFT,FH}(t) = \frac{\bar{a}^{-2}(i\gamma H m_s - 1)}{\chi\gamma H(N_B^2-\omega_1^2)}\left[OC(\omega_1)-P\omega_1 S(\omega_1)\right] \quad (1.45)$$

$$\phi_{FFT,FH}(t) = \frac{im_s g(\gamma-1)\bar{a}^{-2}}{c_s^2 N_B^2}K + \frac{\bar{a}^{-2}}{\chi c_s^2}\left(\frac{im_s g(\gamma-1)-\omega_1^2}{N_B^2-\omega_1^2}\right)\left[\frac{O}{\omega_1}S(\omega_1)+PC(\omega_1)\right] \quad (1.46)$$

$$\varphi_{FFT,FH}(t) = \bar{a}^{-2}K + \frac{\bar{a}^{-2}}{\chi}\left[\frac{O}{\omega_1}S(\omega_1)+PC(\omega_1)\right], \quad (1.47)$$

where $m_s = m - i/2H$, N_B is the buoyancy frequency, and

$$S(\omega) = \sin\omega t - \sin\omega(t-\chi) \quad (1.48)$$

$$C(\omega) = \cos\omega t - \cos\omega(t-\chi) \quad (1.49)$$

$$A_F = kF_{xs,FFT} + lF_{ys,FFT} \quad (1.50)$$

$$B_F = kF_{ys,FFT} - lF_{xs,FFT} \quad (1.51)$$

$$K = \frac{f}{\bar{a}^{-2}s_1^2 s_2^2}(ic_s^2 B_F N_B^2) \quad (1.52)$$

$$O = \frac{1}{s_1^2(s_2^2-s_1^2)(\bar{a}^{-2}+s_1^2)}(-ic_s^2 fB_F(N_B^2+s_1^2)) \quad (1.53)$$

$$P = \frac{1}{s_1^2(s_2^2-s_1^2)(\bar{a}^{-2}+s_1^2)}(-ic_s^2 A_F(N_B^2+s_1^2)) \quad (1.54)$$

The gravity wave intrinsic frequency is $\omega_{GW} = \omega_1 = -is_1$, which satisfies the dispersion relation,

$$\omega_I - [f^2 + c_s^2(k^2 + 1/4H^2)]\omega_I^2 + c_s^2[k_H^2 N_B^2 + f^2(m^2 + 1/4H^2)] = 0, \quad (1.55)$$

where

$$s_1^2 = -\omega_1^2 = -\frac{a}{2}[1 - \sqrt{1 - 4b/a^2}] \quad (1.56)$$

$$s_2^2 = -\omega_2^2 = -\frac{a}{2}[1 + \sqrt{1 - 4b/a^2}] \quad (1.57)$$

$$a = -(s_1^2 + s_2^2) = [f^2 + c_s^2(k^2 + 1/4H^2)] \quad (1.58)$$

$$b = s_1^2 s_2^2 = c_s^2[k_H^2 N_B^2 + f^2(m^2 + 1/4H^2)]. \quad (1.59)$$

This set of gravity wave solutions include the effects from compressibility. The square brackets in equations (1.43) to (1.47) contain the gravity wave terms, because these terms are proportional to $1/\chi$; no secondary gravity waves are excited if $\chi \to \infty$. The terms that do not depend on χ are the mean response terms, and describe the flow components of the counter-rotating cells. The gravity wave amplitudes are linearly proportional to the body force amplitudes since they are proportional to F_{xs} and F_{ys}.

1.4. McMurdo Fe Boltzmann Lidar Campaign

1.4.1. Campaign and Instrument Summary

The datasets involved in this dissertation are from the McMurdo lidar campaign. McMurdo Station (77.8° S, 166.7° E) is located on Ross Island, Antarctica (see Figure 1.6). The University of Colorado Boulder lidar group deployed the Fe Boltzmann lidar to McMurdo Station to fill a critical data gap in the atmospheric observations between the South Pole and the Antarctic Circle [Chu et al., 2011a, 2011b]. The lidar was installed in the Antarctica New Zealand facility at Arrival

Heights near McMurdo and it is a collaboration between the United States Antarctic Program (USAP) and the Antarctica New Zealand (AntNZ). Ever since the first lidar signal was obtained in December 2010, lidar researchers from the University of Colorado Boulder have been collecting data around the clock whenever weather permitted. Thanks to our dedicated winter-over lidar students (Zhibin Yu, Brendan R. Roberts, Weichun Fong, Cao Chen, Jian Zhao, Ian F. Barry, Zhengyu Hua, and Dongming Chang), the project has accomplished its 8th successful year. The Fe Boltzmann lidar consists of two laser channels operated at 372 nm and 374 nm, respectively. Using Rayleigh and Boltzmann techniques, we retrieve the temperature, Fe density, and atmospheric density from ~30 to ~ 170 km above ground [Chu et al., 2002, 2011a, 2011b]. This lidar has the capability to make measurements 24/7 owing to the high-power Pulsed Alexandrite Lasers (PALs) employed in the lidar transmitters, and the narrowband interference filters and Fabry-Perot etalons as daytime filters employed in the lidar receivers. A detailed introduction of this lidar in terms of its principles, capabilities, developments, and error analyses can be found in Chu et al. [2002].

Figure 1.6. Research Stations in Antarctica. [Credit: Map Collection Around the World, https://encykorea.com/davis-weather-station-map/antarctica-time-new-davis-weather-station-map/]

The Fe Boltzmann lidar was designed and constructed at the University of Illinois more than a decade ago [Gelbwachs, 1994; Chu et al., 2002]. The system is illustrated in Figure 1.7. The transmitter consists of lasers of two wavelengths (372 nm and 374 nm), which correspond to two resonance absorption lines originating from the Fe atom's ground state and a low excited state. The two laser channels are enabled by two injection-seeded, flashlamp-pumped, frequency-doubled, Pulsed Alexandrite Lasers (PAL) developed by Light Age, Inc. The seed lasers are DL-Pro (748 nm) and DL-100 (744 nm) from Toptica, which are monitored by a Bristol wavelength meter and a scanning Fabry-Perot Interferometer (FPI) to ensure the

lasers are well locked on the peak of Fe absorption lines and operating in single mode. The whole seed laser assembly were set up in a climate chamber to minimize temperature, vibration, and humidity variations. The seed lasers are then coupled into single mode polarization maintain (SM-PM) fibers to be directed to PAL laser cavities for injection seeding. After the frequency doubling in the PALs, the pulsed lasers are then directed and filtered by two dichroic filters and expanded in beam size to reduce the divergence angles. The lasers are then directed into the sky via steering mirrors. A fast photo diode with an oscilloscope, a pulsed laser spectrum analyzer, a beam profiler camera, and a power meter are used to monitor the laser performance in the time domain, spectral domain, spatial domain, and energy/power domain.

The laser beams interact with atmospheric molecules and Fe atoms, and generate Rayleigh backscattered photons and resonance fluorescence photons, respectively, which eventually are collected by the lidar receivers. Each of the receiver channels consists of a 40-cm Schmidt-Cassegrain telescope with custom-made UV enhanced dielectric coatings. The telescopes are aligned to collect photons in the zenith direction. The return signals collected by the telescope are reflected by a 45° reflection mirror and focused through a field-stop pinhole. A mechanical chopper was placed right after the pinhole to blank the signals from the lower atmosphere to avoid photomultiplier tube (PMT) saturation. The collected signals are then collimated by lenses and pass through a daytime etalon and interference filter to suppress the background light of other wavelengths. A PMT then detects photons in photon counting quantum mode to convert light signals into electronic signals. A discriminator is employed to identify the real signals while rejecting noise-induced counts to maximize the signal-to-noise ratios (SNRs). The output signals are then accumulated and recorded by two multi-channel scalers (SR430) and stored into a computer. From the recorded photon counts, this lidar is capable of detecting atmospheric density and temperature from 30 to 75 km (50 km in summertime when

solar background is prominent) using the Rayleigh integration technique, and Fe density and temperature from 85 to 115 km (~200 km when there are thermospheric Fe layer events) using the Boltzmann technique.

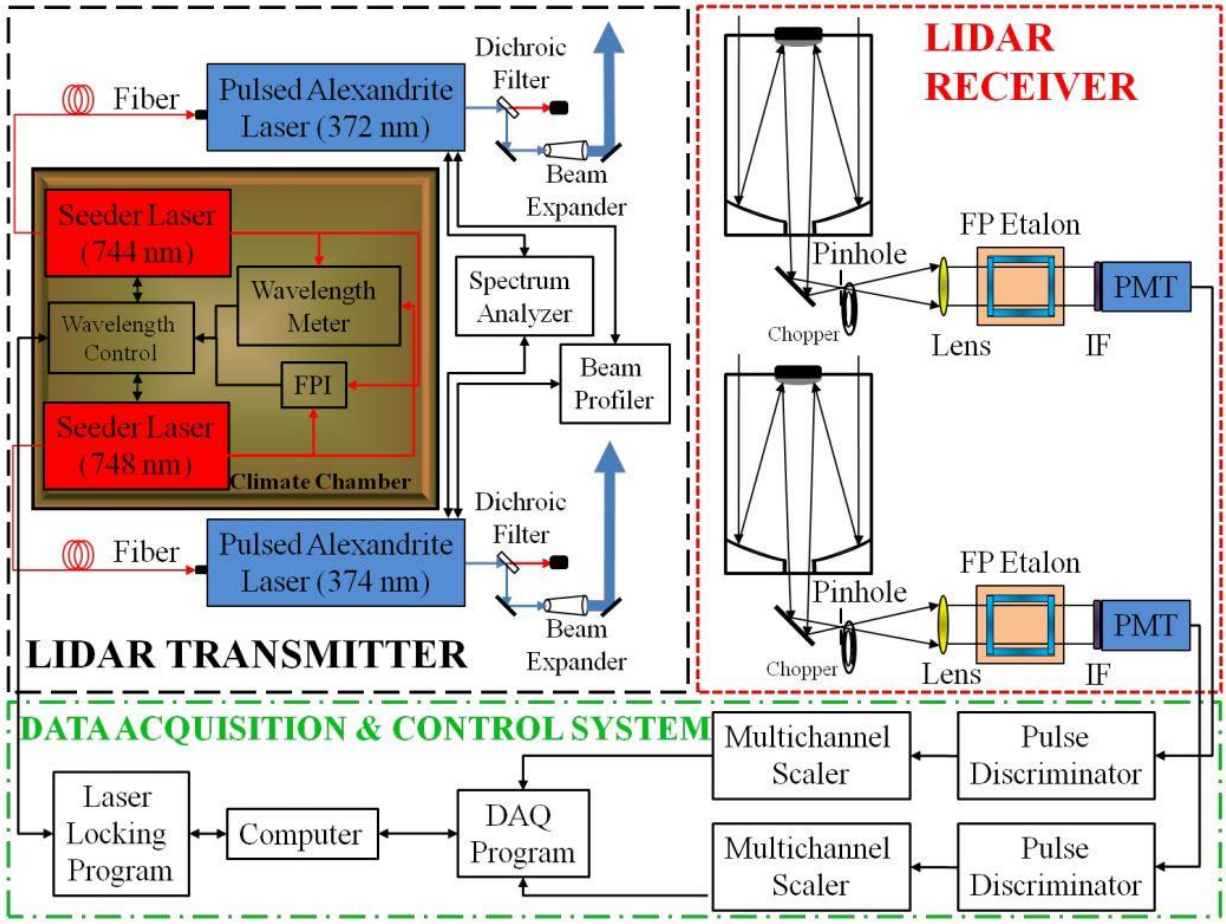

Figure 1.7. Schematic layout of the Fe Boltzmann lidar [Wang et al., 2012].

Before the lidar's deployment to McMurdo station, it was refurbished and upgraded at the University of Colorado Boulder in all three components (Transmitter, Receiver, and Data Acquisition System). The power of the transmitter was improved via replacing worn-out optical parts and realigning optical paths. The receiver was equipped with more powerful day-time observational capabilities via the calibration and upgrade of Fabry-Perot etalons and interference filters. The DAQ and laser control was modernized via LabVIEW interface. After the refurbishment, this lidar

has 24/7 observation capabilities which enabled the gravity wave observations in various temporal and spatial scales.

1.4.2. My Roles in McMurdo Lidar Campaigns

My involvement with the legendary McMurdo lidar campaigns started in December 2013 when I traveled with Dr. Chu to Antarctica for the first time in my life. I was amazed by the beauty of Antarctica and the unique opportunities of conducting the cutting-edge research. During this 2-month trip in Antarctic summer, I focused on learning the lidar technologies and operation from Dr. Chu and a senior PhD student (now Dr. Cao Chen), and helped Drs. Chu and Chen to complete a lidar marathon run — 174 hours of continuous lidar observations by the three of us.

In October 2014 I returned to McMurdo Station and completed a 13-month assignment as the sole winter-over lidar student for winter 2015. Dr. Chen left McMurdo in November 2014. Before that, he handed me over the detailed information on the lidar status. He further trained me and refreshed my memory about the system, which prepared me for the mission in the coming whole year. Dr. Chu and Ian Barry came to the station at the end of Dec 2014. The three of us refurbished the system during Dr. Chu's one-month stay to ensure a smooth campaign the following winter. I helped Dr. Chu to train Ian Barry during his two-month stay to prepare him for his one-year mission after mine. After Ian left in Feb 2015, I officially became the sole winter-over lidar student for the year 2015. During my year, I performed regular lidar maintenance, resolved multiple lidar system technical failures, drew assistance from USAP and AntNZ, arranged the procurement and shipment of lidar consumable parts, and ensured a smooth lidar campaign. One particular setback that happened in my year and is worthwhile mentioning is that the High Voltage Power Supply (HVPS) in the Pulsed Alexandrite Laser (PAL) failed during March 2015, which leads to the paralysis of the 372 nm probing channel. Following the instructions from Dr. Chu and Light Age, Inc, I successfully switched on a spare HVPS, and shipped the broken

HVPS out of the station on the last flight before the station closed. Electrical and plumbing failures also happened every now and then, however, they were relatively easier to fix. A number of record-breaking datasets were also made during that winter, including a 66 hr continuous observation. This was the longest dataset collected by a single winter-over operator at that time. Over my 13-month continuous stay in McMurdo, a total of ~1400 hr of data were collected. These observations have enabled the study of atmospheric planetary wave, tides, gravity waves, and more. Besides managing to operate and maintain the Fe Boltzmann lidar for the best possible data collection, I spent a significant amount of time analyzing the Rayleigh lidar data and reading literature, which eventually formed the basis of this dissertation. My PhD research direction was largely decided upon when I performed various studies on the characterization of gravity waves utilizing the Rayleigh temperature observations. Ian Barry returned to McMurdo at the end of Sep 2015 to take over the lidar campaign and fulfill his one-year mission. I stayed for another month to train him before departure. For me to hand over the campaign smoothly, I showed him the inventory of the lidar spare parts, refreshed his memory on the lidar operation and maintenance, and discussed the diagnostic process if the lidar malfunctions, etc. I left McMurdo by the end of Oct 2015.

After two years away from the ice, I took the 3rd trip to McMurdo in November 2017 to assist Dr. Chu on a new mission — installing a new STAR Na Doppler lidar for simultaneous Fe and Na lidar observations and helping train new winter-over students on the Fe Boltzmann lidar. The shipment, transportation, assembly, diagnostics, and optimization of the lidar required tremendous amount of attentive planning and arrangements. During my two-month stay, I was mainly involved with the mechanical set-up of the lidar, which required a tremendous amount of help from the station to achieve its successful installation. This Na Doppler lidar is equipped with the capability to record the temperature, wind, Na density, and atmospheric

density in the atmosphere, making it a powerful tool to explore the middle and upper atmosphere.

Overall, I had a total of 17 months of "romance" with Antarctica during my PhD research. For the legendary Fe Boltzmann lidar, from December 2013 to February 2014, from October 2014 to October 2015, and from November 2017 to January 2018, I participated in this lidar's observational campaigns, maintenance and training. For the new Na Doppler lidar, I assisted in its initial installation, and witnessed its "birth" from scratch in Antarctica. Being involved in the McMurdo lidar campaigns has greatly broadened my horizon on scientific research and excited my interests in the space-atmosphere sciences, which eventually made this dissertation possible. I am certain that the knowledge, skills, abilities, and experiences I gained from the experience will help me launch a career and assist me well in my future career.

1.5. Gravity Wave Research in This Dissertation

The original motivation for this research was to search the sources of the MLT persistent gravity waves and of the ~1.5 hr fast gravity waves in the Thermosphere-Ionosphere Fe/Fe^+ (TIFe) layer at McMurdo, Antarctica. Such investigation was very important because the persistent gravity waves dominate the temperature variations in the MLT, while the ~1.5 hr fast gravity waves are vital to the formation of the TIFe layers. Along the way of this research, we have updated the ultimate goal of this dissertation to exploring a vertical picture of gravity waves in Antarctica from near the surface to the thermosphere. Before going into depth to reveal detailed information in each chapter, we first explain the overall meaning flow of this dissertation to help the reader understand its contents. To identify if the recently discovered MLT persistent gravity waves come directly from the dominant stratospheric gravity waves, we perform a statistical study on the stratospheric gravity waves in Chapter 2. Based on the large discrepancy of the gravity wave

horizontal wavelengths in the stratosphere and MLT, we rule out the possibility that the stratospheric dominant gravity waves are the direct source of the MLT persistent gravity waves. We discuss the possible sources of the stratospheric gravity waves via the characterization of the gravity wave potential energy density in Chapter 3. In Chapter 4, we utilize the theory of secondary gravity wave generation to link the stratospheric gravity waves with the MLT persistent gravity waves. We find that this mechanism is the possible source of the MLT persistent gravity waves. The spectral proportion method, used to calculate the gravity wave potential energy density, is introduced in Chapter 5. In Chapter 6, we attempt to characterize the ~1.5 hr fast gravity waves from the stratosphere to the thermosphere to investigate the possible sources of the discovered fast gravity waves in TIFe layers. Secondary gravity wave generation is again found to be a possible source. In these chapters, we aim to establish a relatively clearer picture of the gravity waves above McMurdo via the Fe Boltzmann lidar observations. Let us expand on these points slightly below.

Although several theories have been raised to explain their source since the MLT persistent gravity waves were discovered, there was no determination of their exact origin. To determine if the MLT persistent gravity waves come directly from the stratospheric dominant gravity waves (as was postulated by Chen et al. [2013]), 5 years (2011 to 2015) of temperature observations from the Fe Boltzmann lidar are analyzed in order to characterize the dominant gravity wave properties such as the vertical wavelengths, periods, vertical phase speeds, frequency and vertical wavenumber spectra in the stratosphere (30 to 50 km) above McMurdo in Chapter 2. The vertical wavelengths, periods, and vertical phase speeds are found to obey log-normal distributions. Seasonal variations with winter maxima and summer minima are found for gravity wave vertical wavelengths and periods. Interestingly, the vertical phase speeds remain almost constant over the year. Gravity wave vertical

wavelengths increase from ~5.5 km in the summer to ~8.5 km in the winter, while the periods increase from 4.5 hr in the summer to 6 hr in the winter. Linear correlations are found between the monthly mean stratospheric zonal background wind (from ECMWF) and the monthly mean gravity wave vertical wavelengths and periods. We explain these linear correlations using the gravity wave dispersion relation and the Doppler shift effect.

Based on the assumption that the gravity wave monthly mean horizontal phase speeds are roughly constant over a year, we inferred the horizontal wavelengths, intrinsic periods, horizontal group velocities, vertical group velocities, wave propagation azimuth, and elevation angles. Importantly, we found that the horizontal wavelengths of the dominant gravity waves in the stratosphere (~400 km) are much shorter than that of the persistent gravity waves (~1000 to 3000 km in the MLT); therefore, we rule out the possibility that the MLT persistent gravity waves come directly from the dominant stratospheric gravity waves.

At this point, not only did we not know the source of the MLT persistent gravity waves, but also could we not answer the question of the possible sources for the analyzed stratospheric gravity waves. Thus, in Chapter 3, we derive the gravity wave potential energy density per unit mass (E_{pm}), potential energy density per unit volume (E_{pv}), and various related atmospheric parameters in order to investigate the origin of the stratospheric gravity waves. Here the gravity wave E_{pm} is a measure of the wave amplitudes whereas the vertical profile E_{pv} can be used as an indicator of wave dissipation [e.g., Wilson et al., 1991; Whiteway and Carswell, 1995]. In the process, we developed the spectral proportion method to calculate E_{pm} in order to overcome the issue of low signal-to-noise ratios in the summer observations.

The stratospheric gravity wave activity reaches a maximum in the winter and a minimum in the summer. Via correlating E_{pm} with the background wind and

information from the polar vortex, the seasonal variations of E_{pm} are explained as a combined effect from orographic gravity waves from the lower atmosphere, critical level filtering from the background wind, *in-situ* wave source from unbalanced flow induced by the formation of the polar vortex, and the Doppler shift effect from strong stratospheric winds. Critical levels can dissipate the orographic gravity waves when they propagate upward and the horizontal wind speed approaches the gravity wave horizontal phase speed. The Doppler shift effect refers to the situation where strong stratospheric winds induced by the formation of the polar vortex can shift gravity waves to longer vertical wavelengths whereby they need to have larger amplitudes before they can break and saturate.

At this point, we had a generally clearer picture of the possible wave sources for the stratospheric gravity waves. As for the source of the mysterious MLT persistent gravity waves, since the dominant gravity waves in the stratosphere were not the direct source, we continued to the search for their origin. In Chapter 4, we investigated the secondary gravity wave generation theory proposed by Vadas [2003, 2018] as a possible source. This theory predicts that when primary gravity waves propagate upward and break in the atmosphere, the deposition of gravity wave momentum can induce a temporally and spatially localized acceleration of the background mean flow (local body force). The generated intermittent local body force excites secondary gravity waves which have larger horizontal scales than the corresponding primary gravity waves. In this dissertation, we analyze two Boltzmann lidar cases and find that they fit a list of eight reasonable criteria; we therefore conclude that these observed gravity waves are secondary gravity waves. Since the generated secondary gravity waves have horizontal scales much larger than that of the corresponding primary gravity waves, we find that the secondary gravity wave

generation is a possible source of the persistent gravity waves discovered in McMurdo MLT having horizontal wavelengths of thousands of kilometers.

We introduce the mathematical derivations and forward modeling validations of the spectral proportion method in Chapter 5. The motivation to develop this method is that we ran into difficulties when calculating E_{pm}. Due to the low signal-to-noise ratios in the summer, measurements with relatively large temperature uncertainties have to be excluded when utilizing the traditional method. The reason is that the subtraction of noise variance from the total temperature perturbation variance employed in this common practice can lead to negative E_{pm}. Hence, this method might lead to a waste of viable observational data when the desired signals are barely obscured by the noise. In order to overcome such a constraint, alternatively, we proposed the spectral proportion method where we try to quantify how much energy is in the form of the demanded wave signals among all the energy detected via spectra analysis and Monte Carlo simulations. The spectral proportion method has the advantage that with the derived wave energy being only a proportion of the total energy, the unrealistic results of negative wave potential energy density is avoided. With such a method, gravity wave information can be extracted from observations with relatively large observational uncertainties.

In Chapter 6, we attempt to characterize the ~1.5 hr gravity waves from the stratosphere to the thermosphere to discuss the possible source for the ~1.5 hr fast gravity wave in the Thermosphere-Ionosphere Fe/Fe^+ (TIFe) layer. The ~1.5 hr gravity waves are crucial to the formation of the TIFe layers, and a better understanding of them will enhance our knowledge on the dynamics in the thermosphere and optimize the TIFe model implementation. In the stratosphere, the gravity wave atmospheric density and temperature perturbations have a ~180° phase difference with each other from the gravity wave polarization relations. Through a

case study on 28 May 2011, the possibility of wave sources from the lower atmosphere (below 30 km) and from the secondary gravity wave generation mechanism in the stratosphere for these ~1.5 hr gravity waves are investigated. In this case, we find that either lower atmosphere wave sources or the secondary gravity wave generation mechanism may account for the consistent existence of these ~1.5 hr fast gravity waves from the stratosphere all the way up to the thermosphere ~ 155 km. Future work may reveal a single unified source for the ~1.5 hr fast gravity waves in the TIFe layers.

Overall, this dissertation focuses on forming a vertical picture of gravity wave source, propagation, dissipation, and regeneration of gravity waves from near the surface to the thermosphere above McMurdo, Antarctica. Needless to say, organizing a comprehensive picture describing all these mechanisms is beyond the scope of this work. We hope to make contributions to the scientific understandings about gravity waves and yield a clearer understanding of the role that gravity waves play in the atmospheric dynamics of the Earth's atmosphere.

1.6. Research Objectives

Through this dissertation, we hope to generate a clearer vertical picture of gravity waves above McMurdo, Antarctica. The potential links between gravity waves in different atmospheric vertical layers are also investigated by characterizing the stratospheric and the MLT gravity wave signatures. The research objectives of this thesis are:

1) Characterize the dominant gravity wave properties in the stratosphere above McMurdo, Antarctica utilizing the Fe Boltzmann lidar temperature data. The gravity wave parameters we solve for are period, vertical wavelength, vertical phase speed, frequency spectra, vertical wavenumber spectra, and potential energy density over 5 years.

2) Infer the monthly mean horizontal wavelength, intrinsic period, horizontal group velocity, vertical group velocity, propagation azimuth and elevation angles of the stratospheric dominant gravity waves.
3) Investigate the wave source for the gravity waves in the McMurdo stratosphere.
4) Investigate the wave source for the persistent gravity waves in the McMurdo MLT.
5) Provide an observational basis for the secondary gravity wave generation theory, which is postulated to be the possible source for MLT persistent gravity waves.
6) Study the ~1.5 hr gravity waves from the stratosphere to the MLT and speculate possible sources.

1.7. Arrangement of the Dissertation

The organization of the dissertation is:

1) Chapter 1 is the introduction.
2) Chapter 2 characterizes the stratospheric gravity waves in terms of periods, vertical wavelengths, vertical phase speeds, frequency spectra, vertical wavenumber spectra. It is based on a published paper [Zhao et al., 2017].
3) Chapter 3 investigates the stratospheric gravity wave potential energy density and possible wave sources. It is based on a published paper [Chu et al., 2018].
4) Chapter 4 introduces the secondary gravity wave generation theory to explain the possible source for the MLT persistent gravity waves. Two case studies from lidar observations are presented to provide the observational basis. It is based on a published paper [Vadas et al., 2018].
5) Chapter 5 focuses on the development of the spectral proportion method for the calculation of the wave energy density. Mathematical formalism, forward modeling, and validation utilizing lidar observations are provided. It is based on a manuscript to be submitted [Zhao and Chu, 2018].

6) Chapter 6 describes the ~1.5 hr gravity waves from the stratosphere to the thermosphere over McMurdo. Secondary gravity wave generation for such ~1.5 hr fast gravity waves is also discussed.
7) Chapter 7 summarizes the conclusions and future research possibilities.
8) Appendix I presents the mathematical formalism of the spectral proportion method.

CHAPTER II

CHARACTERIZATION OF THE STRATOSPHERIC GRAVITY WAVES: 1. PHYSICAL PROPERTIES AND SOURCE INDICATION

2.1. Introduction

The discovery of the persistent gravity waves (with periods of 3–10 h and vertical wavelengths of 20–30 km) in the mesosphere and lower thermosphere (MLT) by the Fe Boltzmann lidar observations above McMurdo (77.83°S, 166.67°E), Antarctica has posed intriguing questions and challenged the understanding of gravity wave sources [Chen et al., 2016; Chen and Chu, 2017]. No known source could provide perpetual generation of gravity waves with such significant amplitudes (±30 K around altitudes of 100 km), so upon the discovery various theories were invoked to attempt explanations. For example, Godin and Zabotin [2016] proposed resonant vibrations of the Ross Ice Shelf as a source of persistent gravity waves. In a case study, Chen et al. [2013] ray-traced a 7.7-h inertia-gravity wave (IGW) from the MLT above McMurdo to an unbalanced flow in the upper stratosphere around 45 km on the other side of the Antarctic continent. This result motivated us to pursue a statistical study on the dominant gravity waves in the stratosphere using the same Fe lidar data in the 30–65 km altitude range.

In this chapter, we analyze five years (2011 through 2015) of Fe Boltzmann lidar temperature data from McMurdo (77.8° S, 166.7° E), Antarctica, and characterize the lognormal distributions of gravity wave vertical wavelength, period, and vertical phase speed in the stratosphere (30–50 km). An interesting result from the 5-year lidar data is the seasonal variations of vertical wavelength (λ_z) and ground-based period (τ) of the dominant gravity waves in the stratosphere at 30–50 km above McMurdo. Monthly-mean λ_z and τ vary from their summer minima (~5.5

km and 4.5 h) to winter maxima (~8.5 km and 6 h), while monthly-mean vertical phase speed (c_z) remains nearly constant (~0.4 m/s) throughout the year. λ_z and τ exhibit statistically significant linear correlations with the mean background stratospheric winds given by the European Center for Medium-Range Weather Forecasts (ECMWF) model. These linear correlations are explained via the combination of the gravity wave dispersion relation and the Doppler shift effect.

Based upon the observational result of nearly constant c_z, we assume that the gravity wave monthly-mean horizontal phase speeds are nearly constant over the year, and then inferred the gravity wave horizontal wavelengths, intrinsic periods, group velocities, and propagation azimuth and elevation angles. Because the inferred horizontal wavelengths of 350–460 km in the stratosphere are much shorter than those of the persistent gravity waves in the MLT (1000–3000 km) [Chen et al., 2013; Chen and Chu, 2017], we conclude that the dominant gravity waves in the stratosphere at McMurdo are not the direct source of the large-scale persistent gravity waves in the MLT. This conclusion is also supported by the fact that the persistent gravity waves in the MLT have much longer vertical wavelengths than that for the waves in the stratosphere. As noted by Becker and Vadas [2018], wintertime gravity waves in the stratosphere are predominantly orographic or propagating westward. Hence, the vertical wavelengths of these waves decrease with altitude above the polar-night jet maximum (in the mesosphere), rather than increase. This work is published in Zhao et al. [2017].

2.2. Lidar Observational Campaign at Arrival Heights

The University of Colorado lidar group deployed an Fe Boltzmann temperature lidar [Chu et al., 2002; Wang et al., 2012] to Arrival Heights observatory near McMurdo in late 2010 [Chu et al., 2011a]. Ever since, the observational campaign has been ongoing for over six years, recording multiple parameters of the atmosphere

from ~15 km to nearly 200 km, whenever weather permits [Chu et al., 2011a, 2011b, 2016]. Benefited from this long-lasting campaign, invaluable datasets were accumulated for unraveling mysteries in the Antarctic atmosphere. Lidar temperature data from the Rayleigh scattering region (~30–70 km) are used in this study. This analysis covers a 5-year time period from 1 January 2011 to 31 December 2015.

Over the five years, around 5000 h of data were accumulated owing to the hard work from our dedicated winter-over lidar students. Only 4084 h of data are chosen in this study as documented in Table 2.1 where the individual monthly observational lengths in hours are listed over the 60 months of studies. The total observational hours for each individual month vary from 14 to 160 h. Yearly and monthly totals are also shown in this Table. The data screening that results in such data statistics is largely based on the following criteria:

(1) Only datasets with observational length equal to or longer than 6 hours are chosen.

(2) Data with large temporal gaps (4 hours or above) are excluded in this study.

(3) Data with low signal-to-noise ratios (SNRs) are excluded. Low SNRs usually occur during thin cloud coverage or high solar background in summer.

The raw photon counts were collected in the temporal and spatial resolutions of 1 min and 48 m, respectively. During the data retrieval process that converts the raw photon counts to real temperature, different resolutions and altitude ranges are used for the purposes of investigating the different gravity waves parameters. Achieving sufficient SNRs while keeping resolutions as high as possible is the major driver behind these selections. In Sections 2.4. and 2.5., temperature data were retrieved with a temporal integration window of 2 h and an altitudinal window of 0.96 km to retain range information in order to investigate the vertical wavelength and period, etc., when summer data with low SNRs were included. The display resolutions were set to 1 h and 0.96 km. In Section 2.6., resolutions of 0.5 h by 0.96

km were used for the frequency spectrum studies with June data (high quality, nearly zero solar background in Antarctic winter) in order to compare with the work of Chen et al. [2016], while the wavenumber spectrum studies through the entire year adhered to the resolutions of 2 h by 0.96 km. Similarly, the selection of different altitude ranges was also based on the needed SNR for specific topics. For the studies of vertical wavelength, period, phase speed, and vertical wavenumber spectrum covering the entire year, we chose the altitude range of 30–50 km in order to achieve sufficiently high SNRs in summer. Altitude range from 30 to 65 km was chosen for the studies of frequency spectra using the winter data in June only. As a matter of fact, during the wintertime, owing to the high SNRs, this lidar has the capability of probing atmospheric temperatures to as high as 70 km using Rayleigh integration technique [Chu et al., 2011b; Fong, et al., 2014; Chen et al., 2016].

Table 2.1. Statistics on Observational Lengths (in hours) for Individual Month from 2011 to 2015

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2011	82	85	36	40	59	72	54	67	14	38	148	160	855
2012	108	95	19	26	53	72	21	26	49	98	96	49	712
2013	90	19	49	99	45	71	83	63	47	53	91	145	855
2014	137	41	54	30	126	104	79	34	31	82	50	130	898
2015	61	51	57	52	66	121	104	42	67	48	14	81	764
Total	478	291	215	247	349	440	341	232	208	319	399	565	4084

2.3. Analysis Methods to Derive Gravity-Wave-Induced Perturbations

Five years of temperature data are used to analyze the gravity wave activities in the stratosphere and lower mesosphere over McMurdo. Among the significant amount of data accumulated from 2011 through 2015, there are long-lasting datasets with various durations (up to ~65 h). For this study we choose only datasets lasting longer than 6 hours as elaborated above. Except for the frequency and vertical

wavenumber spectra in Section 2.6., for the studies in Sections 2.4. and 2.5., all the datasets longer than 12 h are divided into 12-h observational segments without overlaps in order to ensure reasonable statistics on gravity wave parameters while including gravity wave spectra as much as possible. As the inertial period at McMurdo is 12.24 h and gravity waves with periods of 3–10 h are persistently observed in the MLT region [Chen et al., 2016], 12-h segments are the best choice of window size in our case. In this process of division, if the remaining segment at the end has observational length less than 6 h, then this observational segment is abandoned. Hence, all the observational segments used in this study have data length equal to or longer than 6 h but equal to or shorter than 12 h. Owing to this division process, the total observational hours actually used in Sections 2.4. and 2.5. have been reduced to 3784 h as enumerated in Table 2.2. Note that the frequency spectra (f-spectra) are derived with zero padding to the longest data set (~65 h), while the vertical wavenumber spectra (m-spectra) are derived from individual altitude profiles at a temporal display resolution of 1 h. As a result, the screened data are used without division of segments in the f- and m-spectra.

Table 2.2. Statistics on Seasonal Distributions of Vertical Wavelength, Period, and Vertical Phase Speed

		Summer	Winter	Spring + Fall	Mean/Total
λ_z (km)	Downward	5.97±0.17	8.07±0.18	7.14±0.25	7.06
Mean±Std. Error		(2.86)	(2.88)	(3.18)	
(Std. Deviation)	Upward	5.85±0.24	8.35±0.29	7.12±0.32	7.11
		(3.26)	(2.94)	(3.10)	
Median	Downward	5.26	7.68	6.60	6.51
	Upward	4.84	7.93	6.09	6.29
Skewness, Kurtosis	Downward	0.98, -0.09	0.48, -0.52	0.95, 0.39	
	Upward	1.87, 4.87	0.99, 1.13	1.15, 1.41	
τ (h)	Downward	4.86±0.09	5.72±0.11	5.44±0.14	5.34
Mean±Std. Error		(1.53)	(1.73)	(1.72)	
(Std. Deviation)	Upward	4.50±0.10	5.78±0.18	5.21±0.19	5.16
		(1.32)	(1.83)	(1.79)	
Median	Downward	4.57	5.57	5.33	5.16
	Upward	4.13	5.57	5.02	4.91
Skewness, Kurtosis	Downward	0.43, -0.68	0.58, 0.14	0.48, 0.03	
	Upward	0.82, 0.27	0.27, -0.53	0.47, -0.48	
C_z (m/s)	Downward	0.37±0.01	0.43±0.01	0.39±0.02	0.40
Mean±Std. Error		(0.21)	(0.21)	(0.21)	
(Std. Deviation)	Upward	0.39±0.02	0.44±0.02	0.42±0.02	0.42
		(0.24)	(0.20)	(0.22)	
Median	Downward	0.32	0.39	0.34	0.35
	Upward	0.32	0.39	0.36	0.36
Skewness, Kurtosis	Downward	1.44, 2.32	1.35, 2.65	1.59, 2.92	

	Upward	1.97, 5.37	1.09, 0.98	1.26, 1.55	
Observation hours		1633	1261	890	3784

[1] λ_z, τ and C_z represent vertical wavelength, ground-relative period, and vertical phase speed.

[2] Shown as mean ± standard error (Std. Error), standard deviation (Std. Deviation), median, skewness, and kurtosis.

[3] The Downward and Upward represent downward and upward phase progressions of gravity waves, respectively.

[4] Regarding the Mean/Total column, Mean is for row λ_z, T, and C_z, Total is for row Observation hours.

The following procedure is implemented to estimate the gravity wave perturbations. As the inertia-gravity waves (IGWs) with periods of 3–10 h are persistent and dominant in the McMurdo MLT region [Chen et al., 2013, 2016; Chen and Chu, 2017] while the planetary waves (PWs) are clearly seen in the stratosphere [Lu et al., 2013; Chen et al., 2016], this procedure aims to keep the IGW spectra as full as possible but significantly filter out PWs and tides.

(1) Temperature perturbations $\Delta T(z,t)$ are calculated through subtracting the temporal mean $T_0(z)$ of the 12-h segment at each altitude first, and then subtracting the altitudinal mean at each time grid for every observational segment. Such altitudinal mean subtraction is to remove the nearly vertical stripes found in some segments of the Rayleigh temperature data, equivalently removing waves with long vertical wavelengths. Dividing the absolute temperature perturbations by the temporal mean gives the raw relative temperature perturbations before filtering.

(2) At each individual altitude, a high-pass filter in time domain is applied to the temperature perturbations obtained above to derive temporally filtered

temperature perturbations. The high-pass filter is designed to remove waves with periods longer than 11 h, and it is implemented as the following. First, a one-dimensional fast Fourier transform (1DFFT) is applied to a time series to attain its frequency spectrum. Then a sixth order Butterworth filter window is multiplied with the obtained frequency spectrum to remove frequency components outside the desired frequency range. Finally, an inverse 1DFFT is applied to this multiplied frequency spectrum to attain a new time series with only the selected frequency components.

(3) At each individual time grid, by applying a high-pass filter in the altitudinal domain to the temporally filtered temperature perturbations, the final temperature perturbations $T'(z,t)$ are obtained. This filter and its implementation are identical to the high-pass filter used in Step (2) except it is designed to filter out waves with vertical wavelengths longer than 30 km.

(4) The temporal mean $T_0(z)$ in Step (1) is considered as the temperature background, and the filtered relative temperature perturbations of every observational segment are calculated as

$$T'_{Rel}(z,t) = T'(z,t) / T_0(z) \qquad (2.1)$$

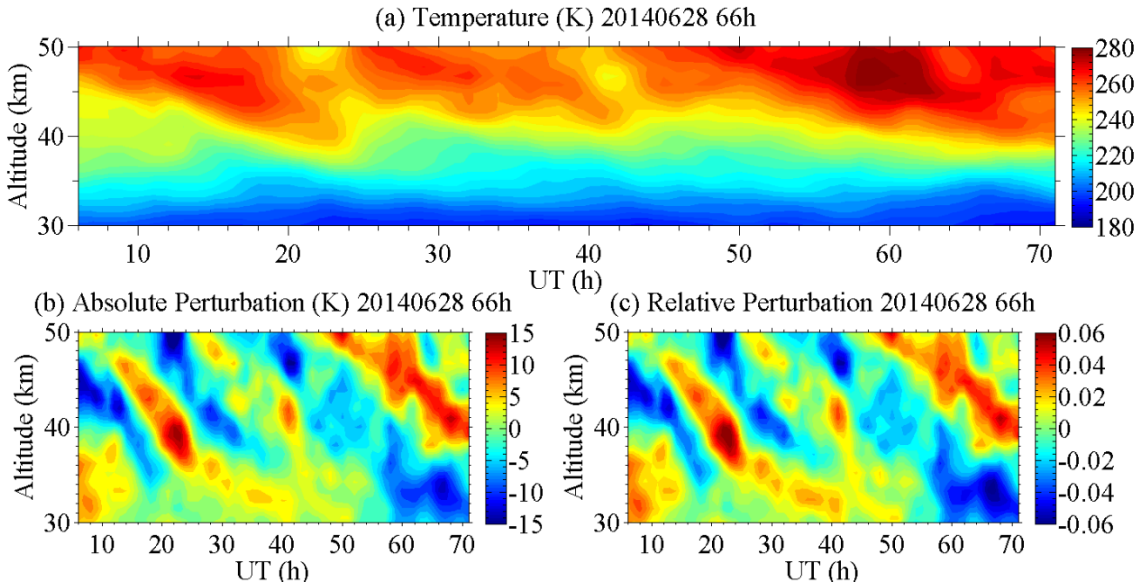

Figure 2.1. An illustration of data processing procedures before the high-pass filtering based on the observation on 28–30 June 2014. (a) Raw lidar temperatures (K) versus UT (h) and altitude (km). (b) Absolute (in unit of K) and (c) relative temperature perturbations are perturbations without high-pass filtering.

Figure 2.1 is an illustration of data processing procedures using the dataset of 28–30 June 2014. Clear downward phase progression can be identified in both the absolute (Figure 2.1b) and raw relative (Figure 2.1c) temperature perturbations without high-pass filtering, which usually indicates upward energy propagation. Note that due to the lack of background wind information, the inferred upward or downward energy propagation here may have ambiguities. Strong wave signatures can be easily identified in these plots. As planetary wave components with periods over 1 d are dominant in the stratosphere [Lu et al., 2013] and visible in the raw relative perturbations (Figure 2.1c), it is necessary to filter out waves with long periods and long vertical wavelengths as elaborated above in Steps (2) and (3) in order to derive pure gravity wave perturbations and minimize the contamination from planetary waves. An example result is shown in Figure 2.2a where gravity wave signatures show up clearly. The filtered relative temperature perturbations derived

above represent the perturbations induced by gravity waves within certain ranges of frequency and vertical wavenumber spectra. Further analyses are performed on these perturbations in order to characterize the McMurdo gravity waves on various aspects. These data analysis methods are introduced before the results are presented in the following Sections 2.4.–2.5.

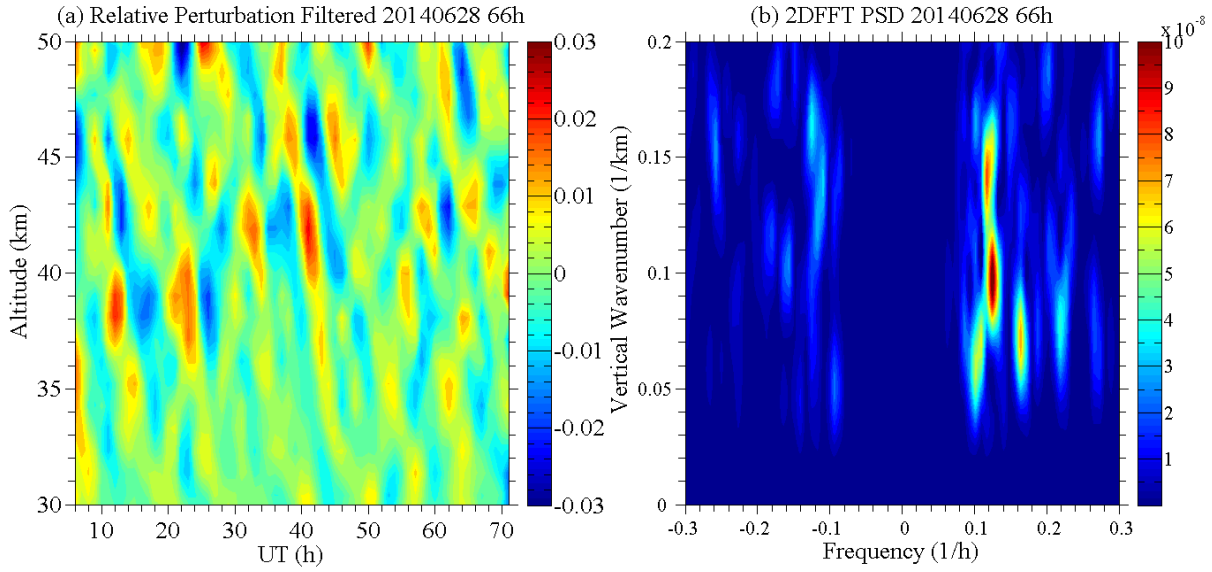

Figure 2.2. (a) High-pass filtered relative perturbations versus UT (h) and altitude (km). (b) Power spectral density versus frequency (1/h) and wavenumber (1/km). The unit of the PSD is arbitrary. Both (a) and (b) are based on the observation on 28–30 June 2014.

2.4. Characterization of Vertical Wavelength, Period, and Phase Speed

Characterizing gravity wave parameters, such as wavelength, period, phase speed, f- and m-spectra, and phase progression direction, etc. for McMurdo has the potential to shed light on the understanding of polar wave dynamics, e.g., in which specific way gravity waves are transporting momentum and energy, and how gravity wave properties change with the time of the year. To derive the basic wave parameters, a two-dimensional fast Fourier transform (2DFFT) is first applied to the relative perturbations of segments to obtain the power spectral density (PSD). These

PSD plots are then scaled by their corresponding temporal length and altitudinal range so that their magnitudes do not depend on the observational window lengths. The scaling is done through dividing the original PSD value on every grid point by the total temporal length and the total altitudinal range. We further convert these scaled PSD plots to area-preserving form through multiplying PSDs by their corresponding wavenumbers and frequencies following the approaches in Yamashita et al. [2009] and Tsuda et al. [2000]. The spectral noise floor of the area-preserving PSD, induced by the lidar measurement errors, is estimated using the Monte Carlo method described below. First, 1000 Gaussian white-noise simulation data are constructed at each grid point of the measurement (time vs. altitude) with a standard deviation equal to the measurement error at that grid point. Then, we run each of the above constructed 1000 sets of 2-D simulation data through the same filtering and 2DFFT processes to obtain their corresponding area-preserving PSD. Finally, the spectral noise floor is estimated by taking the mean of these 1000 simulated noise spectra. The spectral noise floor obtained above is then subtracted from the measured 2DFFT power spectrum. The number of 1000 is chosen with the consideration of computational expenses. In fact, the noise floor barely varies once this number exceeds 500. Figure 2.2 shows an example of the filtered relative perturbations and its area-preserving PSD plot on 28–30 June 2014 with the spectral noise floor subtracted already. Note that as an example of illustrating the procedures to derive relative perturbations and PSD, we did not divide this observation into 12-h segments in Figures 2.1 and 2.2. However, this observation is actually divided into 12-h segments for the study below as is done for all the other observations.

2DFFT spectral analysis is capable of distinguishing the gravity wave vertical phase progression directions (upward or downward) [e.g., Yamashita et al., 2009; Lu et al., 2015b; Lu et al., 2016]. Our 2D FFT formalism leads to the fact that the spectral peaks with positive frequencies correspond to downward phase progression waves,

while the peaks with negative frequencies represent upward phase progression, with the vertical wavenumbers being positive. Using Figure 2.2b as an example, there are several downward phase progression waves with the strongest peak locating at a positive frequency of ~0.13 h^{-1} and a vertical wavenumber of ~0.10 km^{-1}. In comparison, the upward phase progression waves with negative frequencies have much lower power spectral densities in this particular example. In this study, the first 3 dominant waves (among both upward and downward phase progression wave fields) above the spectral noise floor are picked in every one of these area-preserving PSD plots in order to identify the signatures of dominant gravity waves. The choice of three waves is to align in accordance with the statistical studies performed by Yamashita et al. [2009] for the South Pole and Rothera. Table 2.3 summarizes the total number of waves identified from all the qualified 12-h segments through five years from 2011 to 2015, which amount to 1062 waves. The corresponding numbers of ground-relative downward and upward phase progression waves in the seasons of summer, winter, and "spring + fall" are also given in Table 2.3. The seasons are defined as summer from November through February, winter from May through August, fall for March and April, and spring for September and October. The fractions of downward-phase-progression waves out of the total waves in the individual seasons, range from ~59% in summer to ~70% in winter, with an overall average of ~64%. The trend of increasing fraction of downward phase progression waves from summer to winter is consistent with the observations at Rothera by Yamashita et al. [2009]. Note that the ground-relative (not intrinsic) frequency is inferred from the lidar data only.

Table 2.3. Gravity Wave Propagation Direction in the 30–50 km at McMurdo

	Total Number of Waves	Downward Phase Progression	Upward Phase Progression	Downward Phase to Total Wave Ratio
Summer	462	273	189	59.1%
Winter	348	245	103	70.4%
Spring + Fall	252	160	92	63.5%
Total	1062	678	384	63.8%

The statistical studies are conducted on the spectra of the 1062 dominant gravity waves identified above, including both downward and upward phase progression waves. Histograms in Figure 2.3 illustrate the seasonal distributions of vertical wavelengths λ_z, ground-relative periods τ, and vertical phase speeds c_z. For the convenience of presentation, parameters (λ_z, τ, and c_z) of downward phase progression waves are plotted as positive values in Figure 2.3 while the negative values are for upward phase progression waves. The vertical wavelengths of dominant gravity waves distribute from a few to less than 20 km, while the dominant wave periods (ground-relative) range from ~3 to ~10 h. The vertical phase speeds vary from ~0.1 to ~1 m/s. The means, standard errors, and standard deviations of all individual distributions are summarized in Table 2.2, in which all numbers are presented in positive values of their magnitudes. The medians are also provided in Table 2.2. It is obvious from Table 2.2 that the vertical wavelengths and periods increase in winter as shown by the mean and the median values, i.e., λ_z grows from ~5.5 km in summer to ~8 km in winter while τ increases from ~4.5 h in summer to ~5.7 h in winter. Such seasonal variations can be seen for both the downward and upward phase progression waves. There are slight increases in the vertical phase speeds in winter, but not as obvious as λ_z and τ. The seasonal variations of λ_z and τ will be further investigated in Section 2.5. Comparing our results to those by

Yamashita et al. [2009], both the vertical wavelengths (~7.1 km) and periods (~5.3 h) at McMurdo are generally longer than those at the South Pole (~4.1 km and ~1.7 h) and Rothera (~4.4 km and ~1.8 h) while the vertical phase speeds (~0.4 m/s) at McMurdo are slower than those at the South Pole and Rothera (~0.7 m/s). Such differences are likely due to the fact that different spectra of gravity waves are selected in these two studies. Gravity waves with periods of ~1–6 h and vertical wavelengths of 2–30 km are included in Yamashita et al. [2009], whereas here we focus on waves with periods of ~2–11 h and vertical wavelengths of ~2–30 km. The differences could also be due to different excitation mechanisms that generate gravity waves with different spectral characteristics.

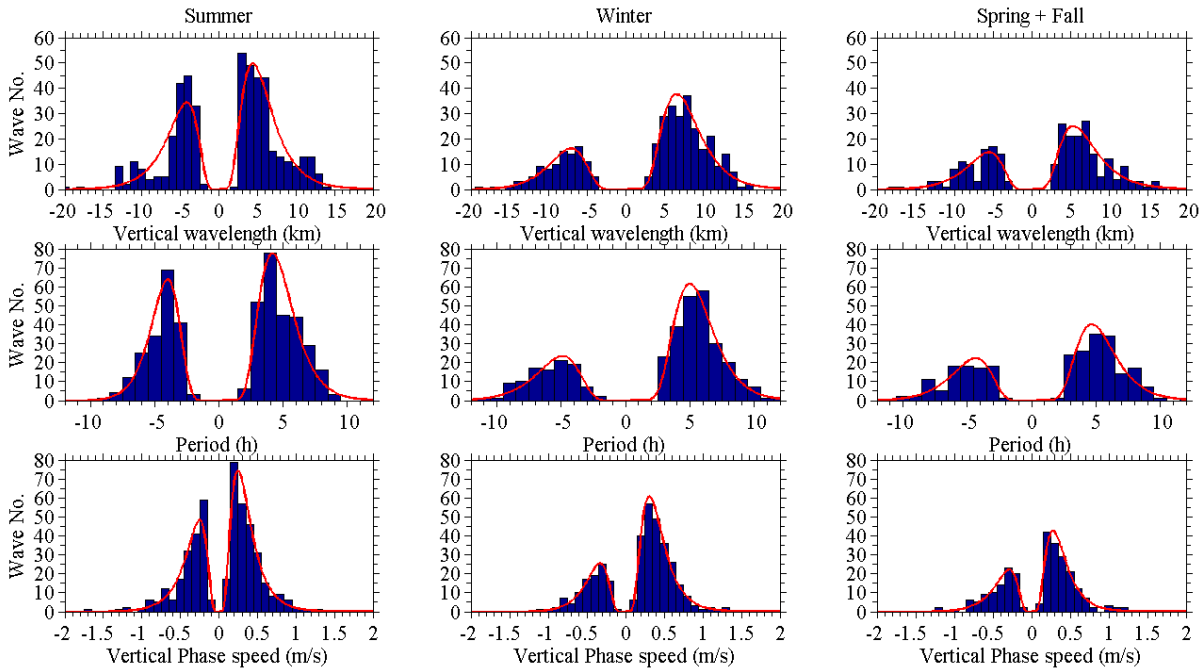

Figure 2.3. Seasonal distributions of (top) vertical wavelength (km), (middle) period (h), and (bottom) vertical phase speed (m/s). For each panel, y-axis shows number of waves. Positive values denote waves with downward phase progression whereas negative values denote waves with upward phase progression.

The individual distributions in Figure 2.3 obviously deviate from normal distributions. To quantify such features, we calculate the skewness and kurtosis for each distribution via equations (2.2) and (2.3). Skewness is the third moment—a measure of asymmetry of a probability distribution around its mean, and Kurtosis is the fourth moment—a measure of the peakedness or flatness relative to a normal distribution [Press et al., 1986; Chu et al., 2006].

$$Skewness(x_1,...,x_i,...,x_N) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\bar{x}}{std}\right)^3 \tag{2.2}$$

$$Kurtosis(x_1,...,x_i,...,x_N) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\bar{x}}{std}\right)^4 - 3 \tag{2.3}$$

where $\bar{x}$ is the mean of values $x_1,...,x_i,...,x_N$, std is the standard deviation of the distribution, and N is the number of data points. The standard errors of the skewness and kurtosis for a normal distribution are approximately $\sqrt{6/N}$ and $\sqrt{24/N}$, respectively. Taking the winter downward phase progression c_z as an example of the skewness and kurtosis in Table 2.2, with 68% confidence level, we have skewness 1.35 ± 0.16, and kurtosis 2.65 ± 0.31; with 95% confidence level, we have skewness 1.35 ± 0.32, and kurtosis 2.65 ± 0.62. The statistics in Table 2.2 reveal that the distributions in Figure 2.3 clearly deviate from normal distributions.

Given that previous observations have shown lognormal distributions of gravity wave momentum flux and potential energy density [Alexander et al., 2008; Hertzog et al., 2012; Baumgaertner and McDonald, 2007; Murphy et al., 2014], we test if the histograms in Figure 2.3 are also log-normally distributed. The lognormal distribution defined by equation (2.4)

$$h(x) = \frac{A}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2} \right] \tag{2.4}$$

is fitted as the red lines in Figure 2.3. The fitting parameters μ, σ and A are summarized in Table 2.4. The correlation coefficients for these lognormal fittings are very high, ranging from 90% to 100% at 95% confidence level. It is therefore unequivocal that the observed λ_z, τ, and c_z are log-normally distributed at McMurdo. Under the definition of lognormal distribution in equation (2.4), the most probable value (MPV) is given by $x_{MPV} = e^{\mu}$, which is summarized in Table 2.4. The trends of vertical wavelengths and ground-relative periods increasing from the summer minima to the winter maxima also clearly show up in the most probable values. Comparing Table 2.4 with Table 2.2, we find that the MPV, median, and mean values have the following relation: $x_{MPV} < x_{median} < x_{mean}$. This fact is consistent with the lognormal distributions that are highly skewed toward lower values as we have characterized from the lidar observations.

Table 2.4. Parameters of the Lognormal Fittings to Histograms in Figure 2.3

			μ	σ	A	Correlation	MPV
λ_z (km)	Summer	Down	1.40±0.05	0.43±0.04	58	94%	4.1
		Up	1.39±0.07	0.32±0.05	38	96%	4.0
	Winter	Down	1.88±0.04	0.41±0.04	36	97%	6.5
		Up	1.91±0.06	0.36±0.04	14	98%	6.8
	Spring Fall	Down	1.65±0.07	0.45±0.04	28	94%	5.2
		Up	1.67±0.08	0.42±0.05	16	90%	5.3
τ (h)	Summer	Down	1.39±0.04	0.37±0.03	65	98%	4.0
		Up	1.36±0.04	0.29±0.03	46	99%	3.9
	Winter	Down	1.62±0.04	0.33±0.02	47	98%	5.0
		Up	1.62±0.06	0.37±0.04	20	98%	5.0
	Spring	Down	1.53±0.05	0.38±0.03	34	96%	4.6
	Fall	Up	1.44±0.08	0.41±0.05	21	94%	4.2
C_z (m/s)	Summer	Down	-1.46±0.06	0.54±0.04	101	99%	0.23
		Up	-1.46±0.07	0.49±0.05	68	98%	0.23
	Winter	Down	-1.18±0.06	0.50±0.04	72	100%	0.31
		Up	-1.17±0.08	0.49±0.05	30	99%	0.31
	Spring	Down	-1.36±0.08	0.51±0.05	55	99%	0.26
	Fall	Up	-1.32±0.10	0.53±0.07	30	98%	0.27

[1] The lognormal fitting function is $h(x) = \frac{A}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right]$.

[2] λ_z, τ and C_z represent the vertical wavelength, ground-relative period, and vertical phase speed, respectively.

[3] A, μ, and σ represent the amplitude, mean, and standard deviation of the variable's natural logarithm, respectively.

[4] Correlation denotes the correlation coefficient of the lognormal fitting.

[5] MVP denotes the most probable value of the lognormal distribution.

[6] Down and Up represent downward and upward phase progressions of gravity waves.

[7] The lognormal fittings and correlations are done at 95% confidence level.

2.5. Vertical Wavelengths and Periods versus Background Stratospheric Winds

To further investigate the seasonal variations of λ_z and τ, the monthly means and standard errors of wave parameters averaged over the 5-year analysis

period are plotted in Figure 2.4. The monthly means are calculated by taking the means of certain wave parameters from all the individual gravity waves in a particular month through all five years. Blue (red) lines in these plots correspond to downward (upward) phase progression waves. Figure 2.4a shows clearly that the vertical wavelength increases when approaching winter and decreases as returning to summer. Vertical wavelength reaches its peak of 8.8±0.69 km in June for upward phase progression waves and of 8.6±0.51 km in August for downward phase progression waves. The dominant periods follow a similar annual cycle with the winter peak values of nearly 6 h while the summer values in December and January are of ~4.5 h, as shown in Figure 2.4b. As the case for dominant vertical phase speeds, they roughly remain constant (~0.4 m/s) throughout the year. Because sufficient numbers of gravity waves were recorded for each month over the five years of lidar observations, the errors associated with the derived wave parameters are small enough that the observed seasonal variations of λ_z and τ are statistically significant.

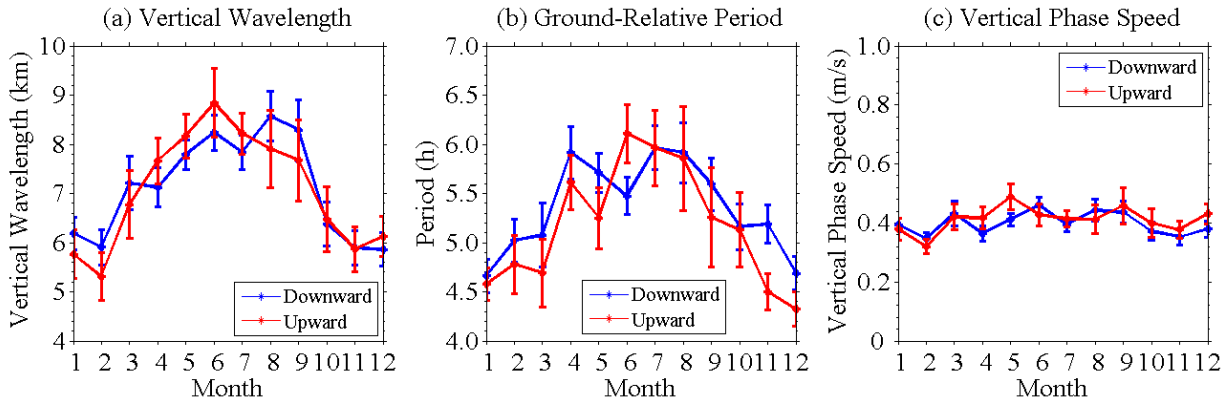

Figure 2.4. Monthly mean vertical wavelengths (a), periods (b), and vertical phase speeds (c) for gravity waves with downward (red) and upward (blue) phase progression. The error bars represent the standard errors, which are equal to the standard deviations divided by the square root of the numbers of independent samples.

Next we investigate whether the seasonal variations of λ_z and τ are related to the mean background winds in the stratosphere [Whiteway et al., 1997; Alexander et al., 2011]. For this purpose, the European Centre for Medium-Range Weather Forecasts (ECMWF) model [Dee et al., 2011] is invoked to provide wind information. The ECMWF model outputs four wind profiles per day near McMurdo station, so there are ~120 wind profiles per month. Figure 2.5 is an illustration of monthly mean zonal and meridional wind profiles from ECMWF at McMurdo averaged over five years from 2011 through 2015. We group λ_z observations with their corresponding total horizontal wind speeds (daily averages of the observational segments) at 40 km from ECMWF. Figures 2.6a and 2.6b shows the probability density functions and cumulative distribution functions of λ_z grouped by wind speed larger than 40 m/s and smaller than 20 m/s. The results can be summarized as that, with higher stratospheric winds, waves tend to have longer vertical wavelengths.

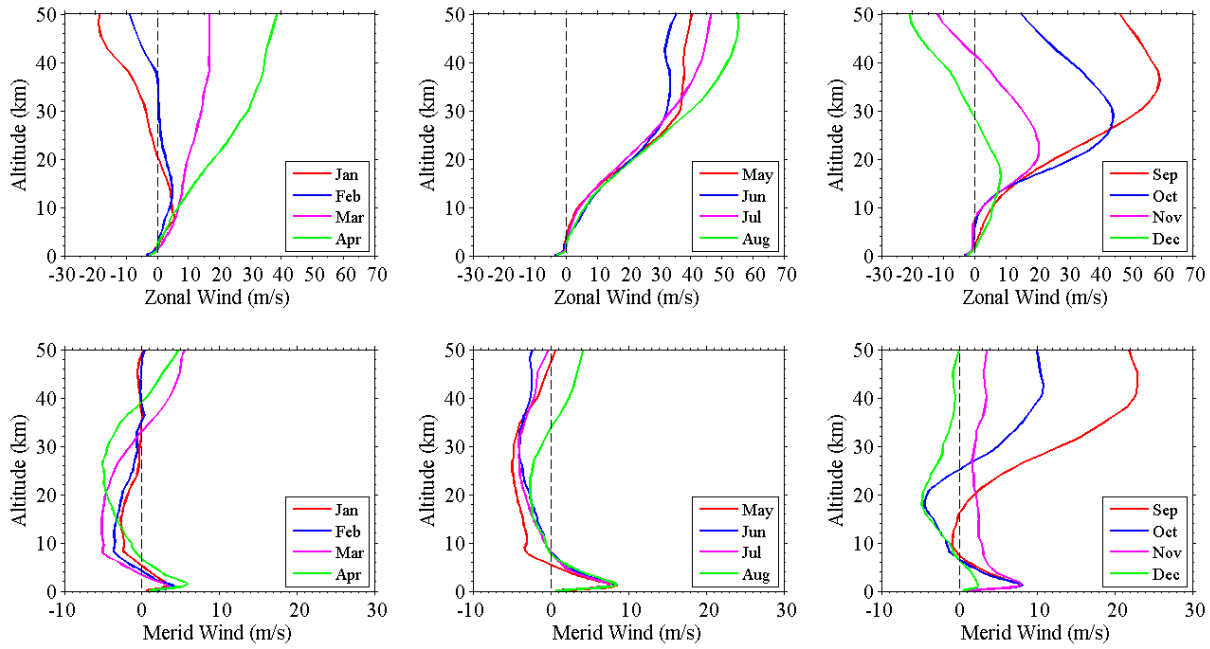

Figure 2.5. Monthly mean zonal (top) and meridional (bottom) wind profiles averaged over five years from 2011 through 2015 using ECMWF data at McMurdo, Antarctica.

We further look into the correlation coefficients between the monthly mean vertical wavelengths (shown in Figure 2.4) and the monthly mean zonal and meridional wind velocities in the stratosphere, respectively. The monthly mean wind velocities are obtained by averaging the zonal and meridional winds, respectively, from 30 to 50 km in the monthly mean profiles of Figure 2.5. The vertical wavelengths λ_z are plotted against the mean zonal winds in Figures 2.6c and 2.6d, and against the mean meridional winds in Figures 2.6e and 2.6f, where the green lines represent the linear regressions with both x- and y-axis errors considered using the York Curve fitting method [York et al., 2004]. Fitting parameters are summarized in Table 2.5. For downward phase progression waves, the vertical wavelength and the mean zonal wind velocity correlate at 0.88 at 95% confidence level. The linear fitting curve has a slope of 39.8 ± 5.1 m/(m/s) and an intercept of 6.22 ± 0.15 km at zero zonal wind. For upward phase progression waves, they correlate at 0.83 at 95% confidence level. The linear fitting curve has a slope of 45.6 ± 7.0 m/(m/s) and an intercept of 6.11 ± 0.20

km at zero zonal wind. In the case of λ_z versus the mean meridional winds, the linear regressions lead to the λ_z intercepts of 6.14 ± 1.28 and 6.03 ± 1.37 km for downward and upward phase progression waves, respectively. However, the correlation coefficients are low, less than 35% (see Table 2.5), which are likely caused by the meridional winds (except September data points) being much smaller than the corresponding zonal winds, while the variances associated with the meridional winds are large over 5 years at McMurdo. Such large wind variances make it hard to show any definite relationship with the small meridional wind velocities. Nevertheless, the zero-wind points in the mean zonal and meridional winds correspond to intercepted λ_z of ~6.1–6.2 km, comparable with each other. Such a result will be used in Section 2.7 to estimate intrinsic properties.

Table 2.5. Parameters Inferred from Figures 2.6 and 2.7

	Zonal Wind		Meridional Wind	
	Downward	Upward	Downward	Upward
S_{l_z} (m/(m/s))	39.8±5.1	45.6±7.0	-1178±613	-1419±643
Intercept (km)	6.22±0.15	6.11±0.20	6.14±1.28	6.03±1.37
Correlation for λ_z	88.3%	83.2%	-25.8%[1]	-33.6%[1]
S_t (s/(m/s))	67.7±10.1	84.6±13.0	-5280±3728	-2746±1380
Intercept (h)	4.94±0.08	4.67±0.09	4.57±1.28	4.51±1.16
Correlation for τ	87.8%	80.7%	-24.5%[1]	-30.8%[1]
c_h (m/s)	21.2	20.8		
$\cos(\theta)$	-0.14	-0.16		

[1] Linear fittings for meridional winds exclude the September data points.

[2] In the calculation, N = 2.14×10^{-2} rad/s, H = 7 km

Similar analyses are done to the ground-relative period τ versus the mean zonal and meridional wind velocities, respectively. Shown in Figures 2.7a and 2.7b are the probability density functions and cumulative distribution functions of periods grouped by total wind speed larger than 40 m/s and smaller than 20 m/s. Similar to the results for λ_z, waves tend to have longer periods under higher winds in the stratosphere. The 12 monthly mean periods are plotted against the mean zonal winds in Figures 2.7c and 2.7d, and against the mean meridional winds in Figures 2.7e and 2.7f. A similar linear regression (with both the x- and y-axis errors considered using the York Curve fitting method) is applied and shown as the green straight lines in Figures 2.7c–2.7f. Fitting parameters are summarized in Table 2.5. For downward phase progression waves, the period and the mean zonal wind velocity correlate at 0.88 at 95% confidence level. The linear fitting curve has a slope of 67.7 ± 10.1 s/(m/s) and an intercept of 4.94 ± 0.08 h at zero zonal wind. For upward phase progression waves, they correlate at 0.81 at 95% confidence level. The linear fitting curve has a

slope of 84.6 ± 13.0 s/(m/s) and an intercept of 4.67 ± 0.09 h at zero zonal wind. Similar to the λ_z case, the meridional wind plots (Figures 2.7e and 2.7f) give low linear correlation coefficients, likely due to the small wind velocities but large variances. The zero-wind points in the mean zonal and meridional winds correspond to intercepted τ of ~4.5–4.9 h.

The 12 months of January through December are marked as the numbers from 1 to 12 in Figures 2.6c–2.6f as well as in Figures 2.7c–2.7f. It is clear from these figures that the six data points from April through September are "clustered" in the right-upper corner of zonal wind plots, the four data points from summer months (December through February) are "clustered" toward the left-lower corner, and March and October are near the middle during transitions. Such summer to winter transition and vice versa are also seen in the ECMWF monthly mean wind profiles in Figure 2.5. Both the vertical wavelengths and ground-relative periods of both downward and upward phase progression waves show statistically significant linear correlations with the mean zonal winds. Such linear correlations will be further investigated in Section 2.7. Overall such linear correlations are stronger for the downward phase progression waves than for the upward phase progression waves, likely because more data points in the downward phase progression waves give better statistics than the upward phase progression case.

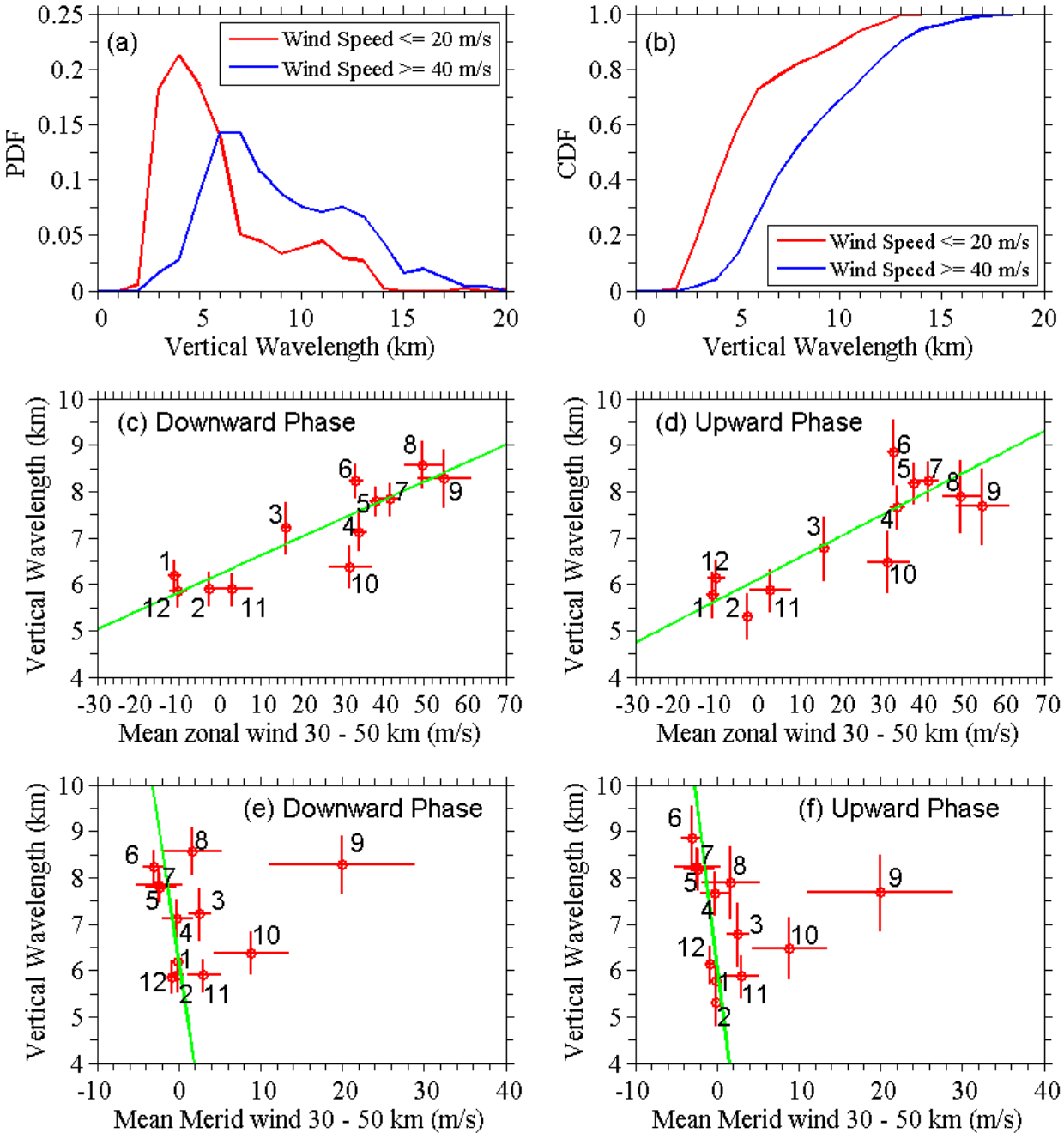

Figure 2.6. (a) Probability density functions, and (b) cumulative distribution functions for vertical wavelength distribution under different wind conditions at 40 km. (c) and (d) are the correlations of the monthly mean vertical wavelengths with the monthly mean zonal wind velocities in the stratosphere (30–50 km) for downward and upward phase progression waves, respectively. (e) and (f) are the correlations of vertical wavelengths with the monthly mean meridional winds. The green lines show

the linear regressions with both x- and y-axis errors considered. The numbers in panels (c-f) denote the month of each individual data point. Note that the September data points are excluded in the linear regressions in (e) and (f).

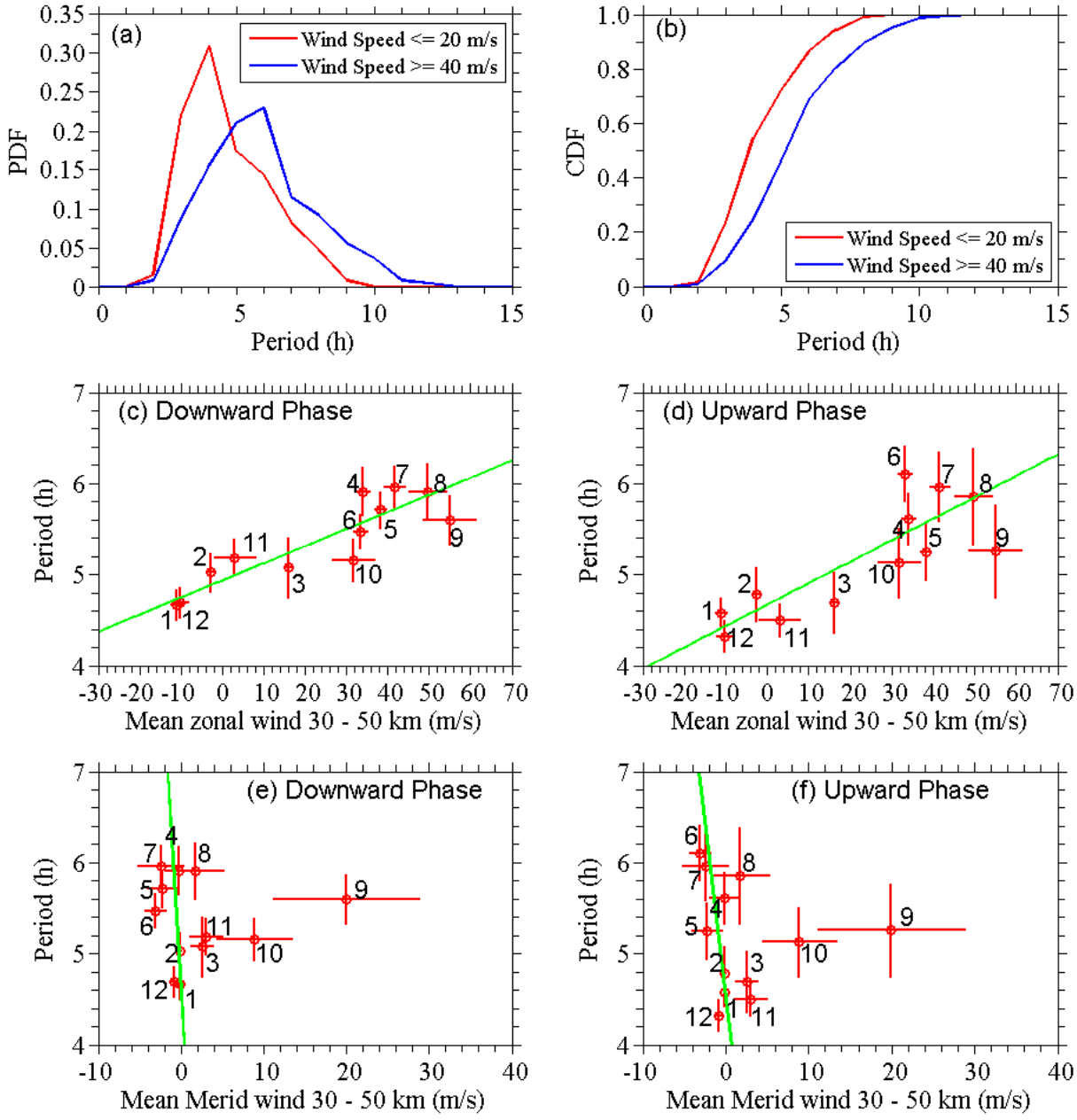

Figure 2.7. Same as Figure 2.6, except for the ground-relative periods.

2.6. Frequency and Vertical Wavenumber Spectra

We perform a frequency spectral study in the upper stratosphere at McMurdo, Antarctica, in order to compare with the observational results of wave frequency spectra in the MLT region reported by Chen et al. [2016]. For this comparison purpose, we take the following procedure identical to that used by Chen et al. [2016] but somewhat different from the one described in Section 2.2. June data over the five

years of 2011–2015 are chosen with the resolutions of 0.5 h by 0.96 km from 30 to 65 km. For this part of the study, we do not divide the observations into 12-h segments, but leave them with their original observational time lengths as we wish to have long time series for frequency spectral analysis. For each observation, the temporal means are derived over the observation length at individual altitudes, and we calculate the relative temperature perturbations by subtracting the temporal means from the raw data and then dividing by the temporal means. The obtained relative perturbations are pre-whitened with a fourth order autoregressive model developed by Chen et al. [2016]. As the longest dataset is 65 h, we zero pad the shorter data to 65 h before applying a 1DFFT to the pre-whitened perturbations at individual altitudes. We then smooth the resulted raw PSDs using a Hamming window with the full-width-at-half-maximum (FWHM) of 0.1 h^{-1} to reduce the variability in the spectral slope estimation. Such smoothed PSDs are post-colored to obtain the real PSDs following Chen et al. [2016]. The real PSDs at individual altitudes are then vertically averaged over 5 km interval to reduce the variance in the spectra. Finally, the results for individual observations are weighted by their observational time lengths to obtain weighted-average frequency spectra. The mean smoothed frequency spectra are plotted at increments of 5 km from 32.5 to 62.5 km in Figure 2.8a. The error bars are calculated as the standard errors in the process of weighted-averaging.

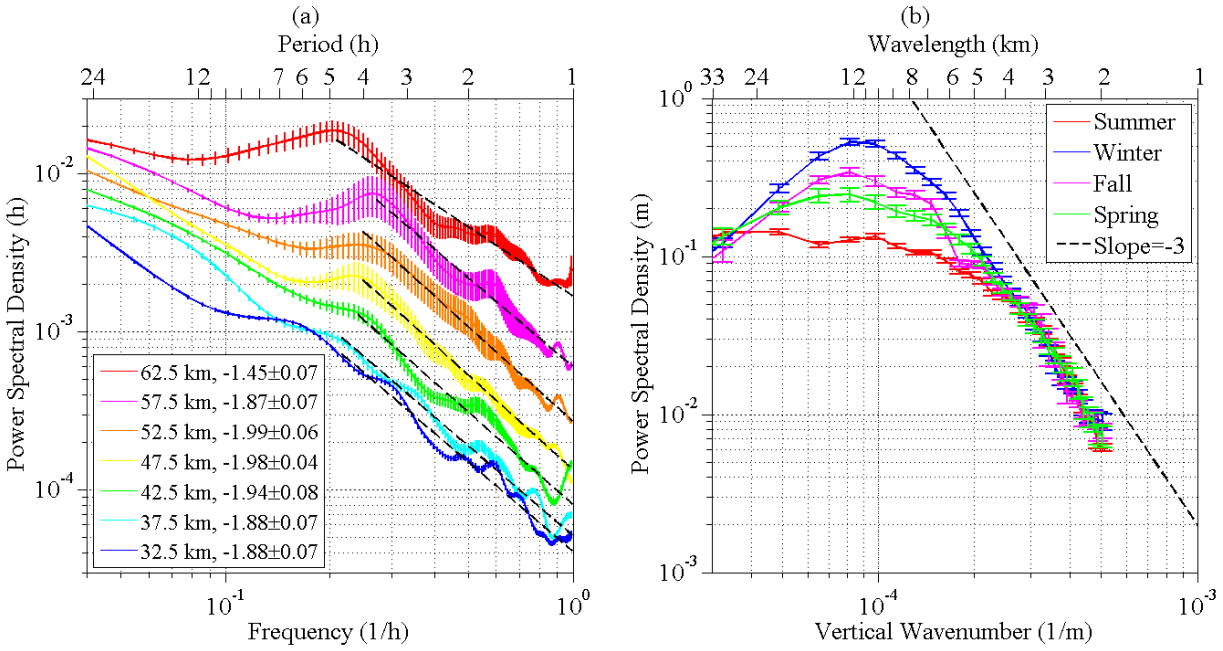

Figure 2.8. (a) Stratospheric gravity wave frequency spectra for different altitudes in June. Different colors represent different altitudes. The dashed black lines are linear fittings with their slopes given in the legends. (b) Power spectral density for vertical wavenumber spectra over different seasons (represented by different colors). The error bars are obtained as standard errors.

The dashed black lines in Figure 2.8a are the least-square fittings of the PSDs in the frequency range from the characteristic frequency to 1 h^{-1} using a linear equation $log_{10}(y) = a + b \times log_{10}(x)$. Here the characteristic frequency is the transition frequency between the regions of positive and negative slopes in the PSD spectra for altitudes above 40 km. As it is difficult to judge where the characteristic frequency is from 32.5 to 37.5 km, the fitting frequency range is set to 0.2–1 h^{-1} for these lower altitudes. The fitted spectral slopes, as indicated in the legends of Figure 2.8a, are about -1.9 from 30 to 60 km, which is close to the value of -1.8 obtained by Hostetler and Gardner [1994] for the upper stratosphere (25–40 km) at Kahalui, Maui. The frequency spectral slopes become much shallower (about -1.45) when the altitudes reach 60–65 km. We notice that the slopes in the stratosphere are generally shallower

than those in the MLT, which range from -2.6 to -1.6 for 85–110 km [Chen et al., 2016]. The slopes in Chen et al. [2016] are derived with a uniform fitting range from 0.1 to 1 h^{-1}. According to Fritts and Alexander [2003], frequency spectra of horizontal wind and temperature typically vary as $PSD(\omega) \sim \omega^{-p}$ (where ω is the wave frequency) with p varying from 1–2 between the inertial and buoyancy frequencies, and most p ≈ -5/3. Our results in the stratosphere are within the 1–2 range but most of the slopes in the MLT fall outside this range [Chen et al., 2016].

Lu et al. [2015a] have studied PSDs in the vertical wavenumber domain for winters from 2011 to 2013 at McMurdo. We expand the study to four seasons from 2011 to 2015. The vertical wavenumber spectra over different seasons are plotted in Figure 2.8b. T'_{Rel} is used to calculate the vertical wavenumber spectra (data resolution of 2 h by 0.96 km from 30 to 50 km over 5 years). We follow the procedures from Dewan and Grossbard [2000] and Lu et al. [2015a]. The power spectral density is derived using Discrete Fourier Transform (DFT) as

$$F(m_k) = \frac{\Delta z \left| f(m_k) \right|^2}{N_z} = \frac{\Delta z}{N_z} \left| \sum_{n=1}^{N_z} x(z_n) e^{-\frac{2\pi(n-1)(k-1)}{N_z} i} \right|^2 \qquad (2.5)$$

where m_k is the k^{th} wave number and $f(m_k)$ denotes its DFT value, Δz is the vertical interval of the data, and N_z is the total points of a single vertical profile. The parameter $x(z_n)$ is the relative temperature perturbation at altitude z_n, which is T'_{Rel} in our case. With current data resolutions, there are no apparent white noise floors at the high wavenumber end of the spectra. Therefore, noise floors were not subtracted from the obtained spectra, similar to the practice by Gardner et al. [1989], Senft and Gardner [1991], and Lu et al. [2015a]. We group the calculated PSDs into four seasons and average them within each season. The average is weighted by observational time length, and the error bars are calculated as the standard errors.

The characteristic vertical wave number $m*$, which is the turning point of the slopes of the spectra, is ~0.08 km^{-1} in winter, fall, and spring, corresponding to a characteristic vertical wavelength of 12.5 km. There is no apparent characteristic vertical wavenumber in summer. For the short vertical wavelength spectrum, the PSDs remain almost the same over different seasons. The seasonal variations of PSDs show up substantially in the long vertical wavelength portion with a lowest level in summer, a highest level in winter, and medium levels in spring and fall. Since there exists an enhancement of PSDs in long vertical wavelength portion in winter, we would have more chances to observe gravity waves with longer vertical wavelengths, which is the case as elaborated in Section 2.4. The dashed black line in Figure 2.8b is an indication of slope -3. The slopes of PSDs in all four seasons are close to -3 in the shorter vertical wavelength range. This result is similar to the studies done by Allen and Vincent [1995] and Lu et al. [2015a].

2.7. Discussion

We explore the linear correlations of λ_z and τ of gravity waves with the mean stratospheric winds in Figures 2.6 and 2.7, and attempt to infer the horizontal wavelengths and intrinsic periods of dominant gravity waves in the stratosphere. Our purpose is to use the information obtained to help judge if the dominant gravity waves in the stratosphere are the same persistent waves as observed in the MLT [Chen et al., 2016; Chen and Chu, 2017] and to infer the dominant wave propagation directions. A Doppler frequency shift exists between the intrinsic frequency ω and the ground-relative frequency ω:

$$\hat{\omega} = \omega - \vec{k} \cdot \vec{\bar{u}} = \omega - k\bar{u}\cos\theta = \omega\left(1 - \frac{\bar{u}\cos\theta}{c_h}\right), \qquad (2.6)$$

where $\vec{k}$ is the wave vector, $\vec{\bar{u}}$ is the mean background wind, θ is the angle between the wave propagation direction and the background wind, and $c_h = \omega / k = \lambda_h / \tau$ is the horizontal phase speed of the gravity wave. Here $\tau = 2\pi / \omega$ is the ground-relative

period, and $\lambda_h = 2\pi/k$ is the horizontal wavelength. As the intrinsic period $\hat{\tau}$ is given by

$$\hat{\tau} = \frac{2\pi}{\hat{\omega}} = \frac{2\pi}{\omega - k\bar{u}\cos\theta} = \frac{\tau}{1 - \frac{\bar{u}\cos\theta}{c_h}}, \tag{2.7}$$

the ground-relative period τ can be expressed precisely as

$$\tau = \hat{\tau}\left(1 - \frac{\bar{u}\cos\theta}{c_h}\right) = \hat{\tau}\left(1 - \frac{k\bar{u}\cos\theta}{\omega}\right). \tag{2.8}$$

Equation (2.8) gives roughly a linear relationship between the ground-relative period and the mean background wind $\bar{u}$, if $\hat{\tau}$, $\cos\theta$, and c_h do not change much, explaining the observations shown in Figures 2.7c and 2.7d. However, because the intrinsic period $\hat{\tau}$ can vary quite a bit under different background winds, deviations from the linear correlation are fairly visible in these two plots. Consequently, the slopes $S_\tau = -\frac{\hat{\tau}}{c_h}\cos\theta = -\frac{2\pi k\cos\theta}{\hat{\omega}\cdot\omega}$ in Figures 2.7c and 2.7d are subject to bias caused by the varied $\hat{\tau}$.

According to the dispersion relation of inertial gravity waves considering the Coriolis force but neglecting the term of $1/4H^2$ [e.g., Marks and Eckermann, 1995; Vadas, 2013], the intrinsic angular frequency $\hat{\omega}$ obeys

$$(\hat{\omega})^2 = \frac{k^2N^2}{m^2} + f^2, \tag{2.9}$$

where N is the Brunt-Väisälä frequency in the unit of radian/s, k and m are the horizontal and vertical wavenumbers, and $\lambda_h = 2\pi/k$ $\lambda_z = 2\pi/m$ are respectively the horizontal and vertical wavelengths. The Earth's rotational frequency is $f = 2\Omega\sin\Theta$, where Ω is Earth's rotation rate and Θ is latitude. For McMurdo, the inertial period is $2\pi/f = 12.24$ h. Combining equations (2.6) and (2.9), we express the vertical wavelength as

$$\lambda_z = \frac{2\pi}{N}\sqrt{(c_h - \bar{u}\cos\theta)^2 - (f/k)^2}, \tag{2.10}$$

Equation (2.10) indicates the relationship between λ_z and the mean background wind is rather complicated; however, we will demonstrate below that for the observed periods of 4–6 h in Figure 2.7, the Coriolis terms in equations (2.9) and (2.10) can be neglected, leading to a linear correlation between λ_z and the background wind. For this purpose, we define

$$\hat{\omega}_0 = N\left|\frac{k}{m}\right| = N\left|\frac{\lambda_z}{\lambda_h}\right| \text{ and } \hat{\tau}_0 = \frac{2\pi}{\hat{\omega}_0} \tag{2.11}$$

as the intrinsic angular frequency and intrinsic period, respectively, when the Coriolis term in equation (2.9) is omitted. Hence, the full intrinsic period can be written as

$$\hat{\tau} = \frac{2\pi}{\hat{\omega}} = \frac{2\pi}{\sqrt{\hat{\omega}_0^{\,2} + f^2}} = \frac{1}{\sqrt{\left(1/\hat{\tau}_0\right)^2 + \left(f/2\pi\right)^2}} = \frac{1}{\sqrt{\left(1/\hat{\tau}_0\right)^2 + \left(1/12.24\text{ h}\right)^2}}, \tag{2.12}$$

Equation (2.12) demonstrates that $\hat{\omega} > \hat{\omega}_0$ and $\hat{\tau} > \hat{\tau}_0$, but the corrections are quite small within the observed range in Figure 2.7. For example, for $\hat{\tau}_0 = 4$ h, $\hat{\tau} = 3.80$ h so the correction is ~5%; for $\hat{\tau}_0 = 6$ h, $\hat{\tau} = 5.39$ h and the correction is ~10%. Such small corrections for wave periods no more than 6 h are due to the square relation in equation (2.9), as the correction is approximately proportional to half of the square of period ratio:

$$\frac{\Delta\hat{\omega}}{\hat{\omega}_0} = \frac{\hat{\omega} - \hat{\omega}_0}{\hat{\omega}_0} \sim \frac{1}{2}\left(\frac{\hat{\tau}_0}{12.24\text{ h}}\right)^2. \tag{2.13}$$

Consequently, it is reasonable to neglect the Coriolis term in equations (2.9) and (10) for the waves with periods of ≤ 6 h, so we obtain a linear relationship between λ_z and the background wind

$$\lambda_z \approx \frac{2\pi}{N}\left(c_h - \bar{u}\cos\theta\right). \tag{2.14}$$

Equation (2.14) helps explain the observed linear correlation in Figures 2.6c and 2.6d. The slopes $S_{\lambda_z} = -\dfrac{2\pi\cos\theta}{N}$ in Figures 2.6c and 2.6d depend on N and $\cos\theta$.

To estimate the horizontal wavelengths and intrinsic periods of the stratospheric gravity waves, we use the following equations

$$\lambda_h = c_h \cdot \tau \approx c_h(\bar{u}\cos\theta = 0)\cdot\tau \tag{2.15}$$

$$\hat{\tau} = \frac{2\pi}{\hat{\omega}_0} = \frac{2\pi}{\left(\frac{\lambda_z}{\tau}\frac{N}{c_h}\right)} \approx \frac{2\pi}{N}\frac{c_h(\bar{u}\cos\theta = 0)}{\lambda_z/\tau} \tag{2.16}$$

An assumption made in equations (2.15) and (2.16) is that the horizontal phase speed c_h does not change much for various background winds, so it is approximated as a constant at the point where the wind projection along the wave vector $\vec{k}$ is zero ($\bar{u}\cos\theta = 0$). This assumption is reasonable considering the vertical phase speed c_z remains nearly constant through the year in Figure 2.4c. Under this assumption, c_h can be derived from equation (2.14) as

$$c_h(\bar{u}\cos\theta = 0) = \frac{N\cdot\lambda_z(\bar{u}\cos\theta = 0)}{2\pi} \tag{2.17}$$

Therefore, the final equations to derive the horizontal wavelengths and intrinsic periods are

$$\lambda_h \approx \frac{N\cdot\lambda_z(\bar{u}\cos\theta = 0)\cdot\tau}{2\pi} \tag{2.18}$$

$$\hat{\tau} \approx \frac{\lambda_z(\bar{u}\cos\theta = 0)}{\lambda_z/\tau} \tag{2.19}$$

The $\cos\theta$ in equation (2.14) can be estimated from the fitted slopes in Figures 2.6c and 2.6d

$$\cos\theta = -\frac{S_{\lambda_z}\cdot N}{2\pi}. \tag{2.20}$$

The horizontal and vertical group velocities relative to the ground can be estimated using the intrinsic properties and background winds as [Fritts and Alexander, 2003; Chen et al., 2013]

$$c_{gh} = \frac{\partial\omega}{\partial k_h} = \frac{k_h\left(N^2 - \hat{\omega}^2\right)}{\hat{\omega}\left(k_h^2 + m^2 + 1/4H^2\right)} + u_h\cos\theta \tag{2.21}$$

$$c_{gz} = \frac{\partial \omega}{\partial m} = -\frac{m\left(\hat{\omega}^2 - f^2\right)}{\hat{\omega}\left(k_h^2 + m^2 + 1/4H^2\right)} + w, \tag{2.22}$$

where u_h and w are respectively the horizontal and vertical background winds, and the vertical wavenumber m has negative and positive signs for the downward and upward phase progression waves, respectively. That is, when the background wind w is negligible, the energy propagation direction is opposite to the phase progression direction.

It is worth clarifying the convention usage on k, $\bar{u}$ and $\cos\theta$ in the above derivations. Our convention is that k is the (positive) magnitude of wave vector $\vec{k}$, the background wind $\bar{u}$ contains the sign information (not the absolute value), while $\cos\theta$ does not have the sign change and θ is defined as the angle from the positive wind direction to the wave vector $\vec{k}$ direction, regardless of the wind sign. In principle $\bar{u}$ in equation (2.6) is the total wind. Considering the vertical winds provided by ECMWF are at least 2–3 orders of magnitude smaller than the horizontal winds in the stratosphere, the vertical winds are neglected in all equations except equation (2.22). Also, because the meridional winds provided by ECMWF are in general much smaller than the zonal winds, we choose to use the zonal winds to represent the stratospheric horizontal winds. Defining the eastward as the positive zonal wind direction, the zonal winds plotted in Figures 2.6 and 2.7 have both positive and negative signs. Under such a convention, θ in above equations is defined as the angle from the eastward to the wave vector $\vec{k}$, while u_h in equation (2.21) is the zonal wind with positive and negative signs, but not the magnitude only.

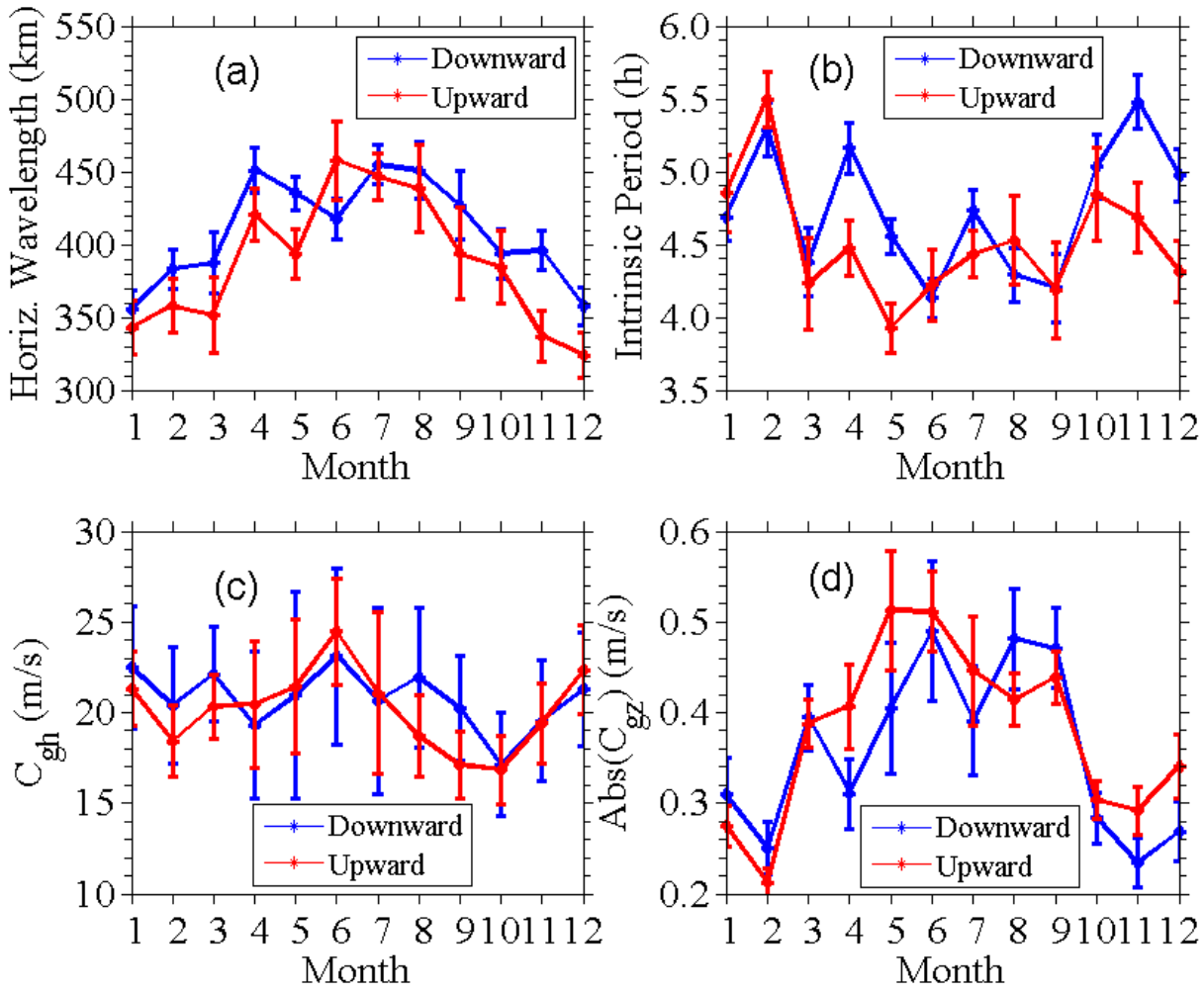

Figure 2.9. Inferred (a) horizontal wavelength, (b) intrinsic period, (c) ground-relative horizontal group velocity, and (d) the absolute value of vertical group velocity of the stratospheric gravity waves through 12 months of a year at McMurdo.

With equations (2.18–2.22) and the convention stated above, we estimate the intrinsic properties of gravity waves via plugging in actual numbers. As the intercepts in the zonal and meridional winds are comparable in Table 2.5 but the meridional cases have much larger uncertainties than the zonal cases, we infer $\lambda_z(\bar{u}\cos\theta = 0)$ only from the zonal wind plots. The vertical wavelengths at the zero wind points are inferred to be $\lambda_z(\bar{u}\cos\theta = 0) = 6.22$ km and $\lambda_z(\bar{u}\cos\theta = 0) = 6.11$ km for downward and upward phase progression waves, respectively. Taking the mean value of $N = 2.14\times10^{-2}$ radian/s derived from the lidar temperature data, we calculate the

horizontal phase speeds to be $c_h(\bar{u}\cos\theta = 0) = 21.2$ m/s and $c_h(\bar{u}\cos\theta = 0) = 20.8$ m/s for the downward and upward phase progression waves, respectively. Substituting the monthly mean vertical wavelengths and ground-relative periods (as shown in Figures 2.4a and 2.4b) along with $\lambda_z(\bar{u}\cos\theta = 0)$ and $N = 2.14\times10^{-2}$ radian/s into equations (2.18) and (2.19), we estimate the horizontal wavelengths and intrinsic periods for all 12 months. The results are plotted in Figures 2.9a and 2.9b. The horizontal wavelengths are around 350–460 km, while the intrinsic periods are about 4–5 h. Using the slopes S_{λ_z} derived from the zonal wind plots (see Table 2.5), $\cos\theta$ derived from equation (2.20) are –0.14 and –0.16 for downward and upward phase progression waves, respectively. The horizontal and vertical group velocities are then estimated from equations (2.21) and (2.22) using the intrinsic properties and background winds, and the results are plotted in Figures 2.9c and 2.9d. The elevation angles of the energy propagation of the gravity waves calculated from the ratio of vertical to horizontal group velocities are about 1.1°, comparable to the case studies by Chen et al. [2013] in the MLT. Note that, for downward (upward) phase progression waves, this elevation is for upward (downward) energy propagation.

The derived intrinsic periods are shorter than the corresponding ground-relative periods in most of the 12 months except the summer four months (November through February). When the zonal winds are eastward (positive) from March through October, negative $\cos\theta$ indicates that in the non-summer months most gravity waves that survive at a given altitude (i.e., are not filtered out by the background wind) propagate against the background winds, leading to higher intrinsic frequencies. Consequently, the intrinsic periods are shorter than the ground-relative periods outside the summer months. During December, January and February, the zonal winds are westward (negative). Negative $\cos\theta$ indicates that the dominant waves propagate along the background winds in summer, leading to lower intrinsic frequencies thus longer intrinsic periods than the ground-relative

counterparts. Because the horizontal winds in the winter stratosphere at McMurdo are dominated by the zonal winds, the small absolute values of $\cos\theta$ (0.14–0.16) indicate that the stratospheric gravity waves have their predominant propagating directions nearly along the north-south (meridional) directions.

In this section, the Doppler shift between the intrinsic and ground-relative frequencies as well as the dispersion relationship neglecting the Coriolis term help explain the linear correlations observed between the monthly mean vertical wavelengths/ground-relative periods and the monthly mean zonal wind velocities in the stratosphere. The linear relationships may reflect the critical-level filtering effects, e.g., the orographic waves may be largely filtered out during summer months when the winds cross the zero point, while the strong eastward winds in winter may allow the orographic waves to propagate upward. Another factor is the extra wave generation in the stratosphere by the strong-wind-induced jet streams or unbalanced flow. Besides such speculations, the search of wave sources is beyond the scope of this study.

The derived horizontal wavelengths in the stratosphere are typically shorter than 500 km, which are significantly shorter than the dominant horizontal wavelengths in the MLT that are typically over 1000 km and up to several thousands of kilometers [Chen et al., 2013, 2016; Chen and Chu, 2017]. If we assume that the horizontal wavelength of a gravity wave does not change significantly as the wave propagates upward in altitude, it is likely that the persistent gravity waves in the MLT region do not come directly from the stratospheric dominant waves with horizontal wavelengths of less than 500 km. Of course, we cannot rule out the possibility that some gravity waves in the tail of the lognormal distribution can propagate to the MLT region. Another possibility is that some of the gravity waves propagating upward in the stratosphere dissipate between 40 and 65 km, which overlap with part of the gap region between the currently studied altitudes and the

MLT altitudes studied by Chen et al. [2016]. Such dissipation results in horizontal body forces that generate upward and downward-propagating secondary gravity waves [Vadas et al., 2003; Vadas and Becker, 2018]. These secondary waves usually have a wide spectral range, and, once propagating into the MLT, they become important sources for the persistent and dominant 3–10 h waves discovered by lidar in the MLT region [Chen et al., 2016; Chen and Chu, 2017], as shown via modeling [Becker and Vadas, 2018; Vadas and Becker 2018]. This speculation of possible wave sources in the lower to middle mesosphere may help explain the observation that the frequency spectral slopes in the lower mesosphere are much shallower than those in the MLT. This is because the upward propagating waves originated from the lower to middle mesosphere still carry some source spectra and have not dissipated sufficiently, thus contributing to the steeper slopes from 85–100 km (not far from the source) [Chen et al., 2016]; while the downward propagating waves originating from this source overlap with the upward propagating waves (from the lower atmosphere) around 60–65 km (right below the source region) with sufficient amplitudes, which may contribute to the shallower slopes as observed in this study. More observations are needed, especially in the gap region of 65–82 km, to investigate this possibility.

2.8. Summary

Five years of atmospheric temperature data, measured with the Fe Boltzmann lidar by the University of Colorado Boulder group from 2011 to 2015 at Arrival Heights near McMurdo, Antarctica, are used to characterize the vertical wavelengths, periods, vertical phase speeds, frequency spectra, and vertical wavenumber spectra of stratospheric gravity waves from 30–50 km altitudes. A total of 1062 dominant gravity waves are identified from the data, allowing us to quantify the distributions of basic wave parameters. The seasonal distributions of vertical wavelengths, periods, and vertical phase speeds in summer, winter, and spring/fall are found obeying lognormal distributions. The lognormal probability density function fittings to these

distributions have nearly 100% correlation coefficients, making it unequivocal that the observed vertical wavelengths, periods, and vertical phase speeds are log-normally distributed at McMurdo. Both the downward and upward phase progression gravity waves are observed by the lidar, and the fractions of gravity waves with downward phase progression increase from summer ~59% to winter ~70%, similar to the lidar results for the South Pole and Rothera by Yamashita et al. [2009]. The seasonal and monthly mean vertical wavelengths and periods exhibit clear seasonal cycles with vertical wavelength growing from summer ~5.5 km to winter ~8.5 km, and period increasing from summer ~4.5 h to winter ~6 h.

By analyzing the ECMWF wind data from 2011 through 2015, we discover statistically significant linear correlations between the monthly mean vertical wavelengths/periods and the mean zonal wind velocities in the stratosphere. With higher stratospheric winds, gravity waves tend to have longer vertical wavelengths and longer ground-relative periods. Such linear correlations may be explained through the Doppler frequency shift of the intrinsic frequency and the gravity wave dispersion relationship when neglecting the Coriolis term. By taking the vertical wavelengths at zero zonal wind points to estimate the horizontal phase speeds of ~21 m/s and assuming nearly constant monthly mean horizontal phase speeds throughout the year, the monthly mean horizontal wavelengths, intrinsic periods, and horizontal and vertical group velocities are estimated for the stratospheric gravity waves through 12 months. The gravity waves reaching the McMurdo stratosphere (30–50 km) tend to have vertical wavelengths of ~6–8 km, horizontal wavelengths of ~350–460 km, intrinsic periods of ~4–5 h, and group velocities of ~22 m/s and ~0.45 m/s for the horizontal and vertical directions, respectively. Two interesting results further inferred are that the stratospheric gravity waves at McMurdo propagate along nearly the north-south direction, and the wave energy propagates at a very shallow elevation of ~1.1°. Gravity wave frequency spectra show that the slopes between the

characteristic frequency and 1 h^{-1} change from about -1.9 at 30–60 km to -1.45 around 60–65 km. The vertical wavenumber spectra show that the power spectral densities at vertical wavelengths of ~5–20 km decrease from the winter maximum to the summer minimum.

Because the typical horizontal wavelengths of the dominant gravity waves in the stratosphere are substantially shorter than those of the dominant and persistent waves in the MLT region, we conclude that the typical dominant waves with horizontal wavelengths of ~350–460 km in the stratosphere are most likely not the direct source of the persistent gravity waves (3–10 h ground-relative periods, 20–30 km vertical wavelength, and over 1000 km horizontal wavelengths) in the MLT observed by Chen et al. [2016] and Chen and Chu [2017]. However, it is possible that these stratospheric dominant waves break in the upper stratosphere and mesosphere, and generate secondary waves. These secondary waves as well as the gravity waves in the tail of the lognormal distribution may survive into the MLT, providing sources for the persistent waves in the MLT.

Overall, in this chapter, through charactering basic gravity wave parameters in the stratosphere and comparing with the MLT wave properties, we have ruled out the possibility that the MLT persistent gravity waves come directly from the dominant gravity waves in the stratosphere. The ultimate source for it is a mystery at this point (In Chapter 4, we will invoke the secondary gravity wave generation theory to explore the source). Furthermore, although we explain the characteristics of certain gravity wave parameters with the background winds from ECMWF, the possible sources responsible for the analyzed stratospheric gravity waves are still unclear. In the next chapter, we aim to investigate the sources of the stratospheric gravity waves via characterizing gravity wave potential energy densities and their seasonal variations.

CHAPTER III

CHARACTERIZATION OF THE STRATOSPHERIC GRAVITY WAVES: 2. POTENTIAL ENERGY DENSITY AND SOURCE SEARCH

3.1. Introduction

To search for the possible sources of the stratospheric gravity waves, five years of Fe Boltzmann lidar's Rayleigh temperature data from 2011 through 2015 at McMurdo are used to characterize gravity wave potential energy mass density (E_{pm}), potential energy volume density (E_{pv}), vertical wavenumber spectra, and static stability N^2 in the stratosphere 30–50 km. A spectral proportion method was developed along the way in order to handle observational data with low signal-to-noise ratios (SNRs) in the Antarctic summers.

E_{pm} (E_{pv}) profiles increase (decrease) with altitude and the scale heights of E_{pv} indicate stronger wave dissipation in winter than in summer. Altitude-mean $\overline{E}_{pm}$ and $\overline{E}_{pv}$ obey lognormal distributions and possess narrowly clustered small values in summer but widely spread large values in winter. $\overline{E}_{pm}$ and $\overline{E}_{pv}$ vary significantly from observation to observation, but exhibit repeated seasonal patterns with summer minima and winter maxima. The winter maxima in 2012 and 2015 are higher than in other years, indicating inter-annual variations. Altitude-mean $\overline{N^2}$ varies by ~30–40% from the mid-winter maxima to minima around October, and exhibits a nearly bi-modal distribution. Monthly-mean vertical wavenumber power spectral density for vertical wavelengths of 5–20 km increases from summer to winter.

Using MERRA-2 data we find that large values of $\overline{E}_{pm}$ during wintertime occur when McMurdo is well inside the polar vortex. Monthly-mean $\overline{E}_{pm}$ are anti-correlated with wind rotation angles but positively correlated with wind speeds at 3

and 30 km. Corresponding correlation coefficients are –0.62, +0.87, and +0.80, respectively. Results indicate that the summer-winter asymmetry of $\overline{E}_{pm}$ is mainly caused by critical level filtering that dissipates most gravity waves in summer. $\overline{E}_{pm}$ variations in winter are mainly due to variations of gravity wave generation in the troposphere and stratosphere, and Doppler shifting by the mean stratospheric winds. This work is published in Chu, Zhao et al. [2018].

3.2. Methodology

3.2.1. Lidar Observations at Arrival Heights

A lidar observational campaign has been ongoing at Arrival Heights observatory (77.84°S, 166.69°E) near McMurdo, Antarctica since December 2010 via the collaboration between the United States Antarctic Program (USAP) and Antarctica New Zealand (AntNZ) [Chu et al., 2011a,b]. An Fe Boltzmann temperature lidar [Chu et al., 2002; Wang et al., 2012] deployed by the University of Colorado Boulder has been recording multiple parameters of the atmosphere from ~15 km to nearly 200 km [e.g., Chu et al., 2011a,b, 2016; Yu et al., 2012; Chen et al., 2013, 2016; Lu et al., 2013, 2015, 2017; Fong et al., 2014, 2015; Chen and Chu, 2017; Chu and Yu, 2017; Zhao et al., 2017]. Analyzed here are the five years of lidar temperature data from the pure Rayleigh scattering region (~30–70 km) from 1 January 2011 through 31 December 2015. Over these five years, around 5000 h of data were collected owing to the dedication of our winter-over lidar students (Yu, Roberts, Fong, Chen, and Zhao) and summer-season lidar researchers. The data collection was mainly dictated by weather conditions, and the data distributions in overall 60 months have been tabulated in Zhao et al. [2017]. After data screening and division process as described in Zhao et al. [2017], the actual data used in this study totals to 3798 h. All the data segments used in this study are between 6 h and 12 h in duration. There are a total of 354 data segments and the statistics of segment durations are summarized in Table

3.1. Among them 216 segments have data duration of 12 h and 32 segments of 11 h duration, occupying 70% of the data points. Segmenting the data in this way enriches sufficient statistical samples while preserving the persistent gravity waves with periods of 3–10 h discovered by Chen et al. [2016] in the MLT above McMurdo (see Zhao et al. [2017] for detailed explanations).

Table 3.1. Statistics on Observational Segments from 2011 to 2015 Employed in the Study

Data Duration	Total	12-h	11-h	10-h	9-h	8-h	7-h	6-h
Number of Segments	354	216	32	17	26	22	28	13
Proportion in Data Points	100%	61.0%	9.0%	4.8%	7.4%	6.2%	7.9%	3.7%
Number of Hours (h)	3798	2592	352	170	234	176	196	78
Proportion in Data Length	100%	68.2%	9.3%	4.5%	6.2%	4.6%	5.2%	2.0%

The raw photon counts were collected with temporal and altitude resolutions of 1 min and 48 m, respectively. To achieve sufficient signal-to-noise ratios (SNRs) while keeping resolutions as high as possible, the Rayleigh temperature data were retrieved with a temporal integration window of 2 h and a spatial binning window of 0.96 km. The data retrieval was done with the over-sampling method to have display resolutions of 1 h and 0.96 km. We chose the altitude range of 30–50 km in order to achieve sufficiently high SNRs in summer (when solar background noise is largest) for the studies covering the entire year.

3.2.2. Derivation of Potential Energy Densities

The derivation of E_{pm} and E_{pv} contains two major courses. The first course is to extract gravity wave perturbations from the raw temperature data that are affected by a variety of atmospheric waves (e.g., planetary, tidal and gravity waves), and the second course is to accurately estimate the energy of gravity waves from the derived gravity wave perturbations that also include noise contributions. Here we use summer lidar observations on 31 December 2014 as an example to illustrate the data processing procedures employed. The 12 temperature profiles from 30 to 50 km measured on that day are plotted in Figure 3.1A$_1$, while the corresponding temperature contour is plotted in Figure 3.1A$_2$. Temperatures generally increase from ~240 K at 30 km to ~280–290 K at 50 km. Typical measurement errors increase from ~0.2 K to 6 K from 30 to 50 km in summer but are much smaller (~0.1 K to 1 K) in winter. Two-dimensional fast Fourier transform (2DFFT) applied to the raw temperature data reveals dominant DC components (corresponding to mean background fields with zero frequency and zero vertical wavenumber) in the power spectral density (see Figure 3.1A$_3$).

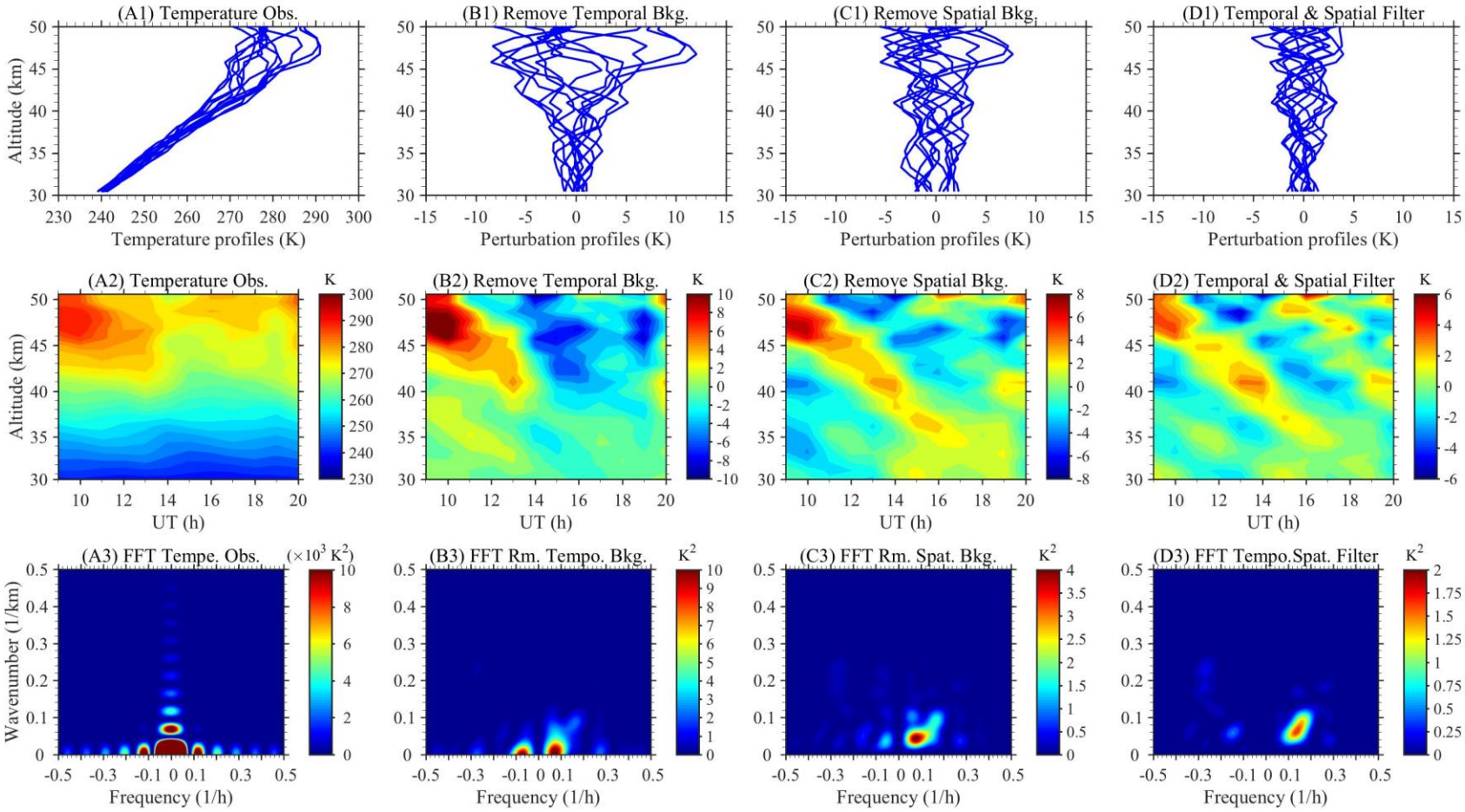

Figure 3.1. Lidar observations on 31 December 2014 at McMurdo are used as an example to illustrate the procedures of extracting gravity wave perturbations from the raw temperature data. (1st row) Temperature profiles, (2nd row) temperature contours, and (3rd row) power spectral density (PSD) of 2D FFT. (A column) Raw temperature profiles along with their corresponding temperature contour and 2D FFT PSD. (B column) Temperature perturbation profiles, contour, and 2D FFT PSD after temporal background subtraction. (C column) Same as B column except after temporal and spatial background subtractions. (D column) Same as B column except after temporal and spatial filtering following the temporal and spatial background subtractions. Note that the color scales for (A3) are 3–4 orders of magnitude larger than those in (B3)–(D3).

3.2.2.1. Extracting Gravity Wave Perturbations from Raw Temperatures

The main goals of the first course are to remove the DC background to reveal wave perturbations and remove waves outside the gravity wave spectra interested via a combination of background subtractions and applications of filters. Briefly, temperature perturbations $\Delta T(z,t)$ are calculated by first subtracting the temporal background $T_0(z)$ of the observational segment at each altitude, and then subtracting the spatial background over the altitudes of 30–50 km at each time grid for every observational segment. Following the background subtractions, at each individual altitude, a high-pass filter is applied to the temperature perturbations obtained above to remove waves with periods longer than 11 h. At each individual time grid, such temporally filtered temperature perturbations are run through a high-pass filter to remove waves with vertical wavelengths longer than 30 km. The obtained final temperature perturbations $T'(z,t)$ are converted to the filtered relative temperature perturbations of every observational segment as

$$T'_{Rel}(z,t) = T'(z,t) / T_0(z) \tag{3.1}$$

Let us illustrate how each step in the procedure described above affects the wave spectra.

1) The temporal background $T_0(z)$ is estimated as the temporal mean over the time span of the observation segment. Subtraction of $T_0(z)$ leads to the removal of DC term and long-period (low-frequency) waves in the frequency spectra. Note that the revealed wave spectra in Figure 3.1B$_3$ have much smaller power densities than those of the DC and long-period spectra in Figure 3.1A$_3$. The corresponding temperature perturbations and contour are plotted in Figures 3.1B$_1$ and 3.1B$_2$. Comparing two contour plots in Figures 3.1A$_2$ and 1B$_2$, we can see that the raw temperatures are dominated by long-period waves (most likely planetary waves [Lu et al., 2013, 2017]), while shorter-period waves begin to show up in Figure 3.1B$_2$.

2) The spatial background is estimated as the running mean over the altitude range of 30–50 km obtained with a 30-km sliding window. For any altitude within the 30–50 km range, the arithmetic mean over its vicinity of ±15 km (i.e., 30-km window width) is taken as the spatial background at this altitude. (When taking the average within the 30-km window, altitude bins outside the observational range are ignored.) As the window slides along the altitude, different altitudes may have different spatial backgrounds, which are different from the traditional practice where a vertical mean over the observational range is used as the spatial background for all altitudes within the 30–50 km range [e.g., Zhao et al., 2017]. Subtraction of the spatial background removes the DC term and long-vertical-wavelength (low-vertical-wavenumber) waves in the vertical wavenumber spectra (Figure 3.1C$_3$). Figure 3.1C$_1$ shows more confined gravity wave perturbation profiles than Figure 3.1B$_1$, while gravity wave signatures become clear in Figure 3.1C$_2$.

3) The temporal and spatial filters applied after the background subtractions effectively remove waves with periods longer than 11 h and vertical wavelengths longer than 30 km that still show considerable energy before filtering (Figure 3.1C$_3$). The choices of filters are to preserve the IGW spectra [Chen et al., 2013, 2016; Chen and Chu, 2017] as fully as possible while significantly removing planetary wave components with periods over 1 d and vertical wavelengths over 40 km that are dominant in the stratosphere [Lu et al., 2013, 2017]. As shown in Table 3.1, ~70% data segments have durations of 11–12 h, which allow the inclusion of full energy of gravity waves with periods of 4–11 h, and partial energy of gravity waves with periods of 2–4 h (because the temporal integration is 2 h for raw temperatures) and from the segment lengths up to twice the segment lengths. The short duration (6–7 h) segments occupy only ~7% of the total data length (see Table 3.1), but they still include partial energy of gravity waves with periods from 6–7 h up to twice the segment lengths (12–14 h). Similarly, the 20-km data window from 30 to 50 km contains the full energy of gravity waves with vertical wavelengths of 2–20 km (the spatial integration is 0.96 km for raw temperatures) and partial energy of gravity waves with vertical wavelengths from 20 km up to twice the window length (40 km). Therefore, the filtered relative temperature perturbations (Figures 3.1D$_1$ and 3.1D$_2$) derived above represent the perturbations induced mainly by gravity waves with periods of 3–11 h and vertical wavelengths of 2–30 km (see Figure 3.1D$_3$).

3.2.2.2. Accurately Estimating Gravity Wave Potential Energy Densities

The second course is to further analyze these gravity wave perturbations to derive E_{pm} and E_{pv}, removing the contamination from noise contributions. For each observational segment, the altitude profile of gravity wave potential energy mass density $E_{pm}(z)$ is calculated from the filtered temperature perturbations $T'(z,t)$ using equation (3.2) [e.g., Duck et al., 2001; Yamashita et al., 2009]

$$E_{pm}(z) = \frac{1}{2}\frac{g^2}{N^2(z)}\overline{\left(\frac{T'_{GW}(z,t)}{T_{Bkg}(z)}\right)^2} = \frac{1}{2}\frac{g^2}{N^2(z)}\frac{1}{N_p}\sum_{i=1}^{N_p}\left(\frac{T'_{GW}(z,t_i)}{T_{Bkg}(z)}\right)^2 \quad (3.2)$$

where the over bar denotes taking the mean over the time span of the observational segment, g = 9.7 m/s^2 is the gravitational acceleration corresponding to the stratosphere, z and t respectively represent altitude and time, T_{Bkg} is the monthly mean background temperature, and N_p is the number of temperature perturbation profiles within the segment time span. Here $N(z)$ is the buoyancy frequency calculated from the segment temporal mean temperatures $T_0(z)$ through

$$N^2(z) = \frac{g}{T_0(z)}\left(\frac{dT_0(z)}{dz} + \frac{g}{C_p}\right) \quad (3.3)$$

where C_p = 1004 J/K/kg is the specific heat of dry air at constant pressure. In equation (3.2) T'_{GW} represents the pure temperature perturbations induced by gravity waves, and its mean square should be calculated, in principle, from $T'(z,t)$ by removing the noise variance

$$\overline{[T'_{GW}(z,t)]^2} = \overline{[T'(z,t)]^2} - \overline{[\sigma_T(z,t)]^2} \quad (3.4)$$

where the over bars represent taking the mean over the time span of the observational segment, $T'(z,t)$ is the filtered temperature perturbation as derived in Section 3.2.2.1, and $\overline{[\sigma_T(z,t)]^2}$ is the noise variance due to the temperature uncertainty caused by photon noise.

How to accurately estimate noise variance deserves careful considerations. Duck et al. [2001] calculated the noise variance as

$$\overline{[\sigma_T(z,t)]^2} = \sigma_T^2(z) = \frac{1}{N_p}\sum_{i=1}^{N_p}[\delta T(z,t_i)]^2 \quad (3.5)$$

where $\delta T(z,t_i)$ is the filtered uncertainty of the measured temperature profile. Such subtraction method is best suited for nighttime lidar measurements when the

temperature measurement uncertainty is small compared to the gravity-wave-induced temperature perturbation. In this case, the error in measurement uncertainty estimation is negligible and will not substantially alter the derived $E_{pm}(z)$. However, under full sunlight in Antarctic summer, the magnitude of temperature uncertainty can become so large that it is close to or even exceeds that of the gravity-wave-induced temperature perturbation at high altitudes. Due to the error in the uncertainty estimation, the noise variance as calculated by Eq. (3.5) can become unreliable and even lead to negative values for $E_{pm}(z)$ when using Eqs. (3.4) and (3.2). To overcome this issue, we develop a spectral proportion method that combines Fourier spectral analysis at each altitude with Monte Carlo simulations to estimate the proportion of gravity wave energy occupying the total spectral energy, and then scale the $T'(z,t)$ square with this proportion. Let $p(z)$ represent this gravity-wave proportion at altitude z, the mean square of pure temperature perturbations induced by gravity waves and the noise variance are then computed as

$$\begin{aligned} \overline{[T'_{GW}(z,t)]^2} &= \overline{[T'(z,t)]^2} \bullet p(z) \\ \overline{[\sigma_T(z,t)]^2} &= \overline{[T'(z,t)]^2} \bullet [1-p(z)] \end{aligned} \tag{3.6}$$

The uncertainty associated with $E_{pm}(z)$ is calculated following the same procedure as in Whiteway and Carswell [1995] but the estimate of noise variance is updated with Eq. (6) to:

$$\delta E_{pm}(z) = \frac{1}{2}\frac{g^2}{N^2(z)}\overline{\left[\frac{\sigma_T(z,t)}{T_{Bkg}(z)}\right]^2}\frac{1}{\sqrt{N_p}} = \frac{1}{2}\frac{g^2}{N^2(z)}\overline{\left[\frac{T'(z,t)}{T_{Bkg}(z)}\right]^2}[1-p(z)]\frac{1}{\sqrt{N_p}} \tag{3.7}$$

Here the gravity-wave proportion $p(z)$ is estimated with the following Monte Carlo procedure. We first construct 1000 sets of 2-D temperature map (temperature versus time and altitude) with Gaussian white-noise added to the lidar-measured raw temperatures, i.e., the standard deviation of the added noise is equal to the raw

measurement error at each grid point. Second, we run each of the so-constructed 2-D temperature maps through the same processing procedure as described in Section 3.2.2.1 (background subtraction and filtering processes) to obtain filtered temperature perturbation fields. Then at each altitude we calculate the 1D FFT power spectra for each time series of the 1000 filtered temperature perturbations, and take the mean of these 1000 spectra of 1D FFT to estimate the spectral noise floor. An example is shown in Figure 3.2a for lidar observations on 31 December 2014 at an altitude of 48.64 km. The dashed red line indicates the noise floor at various frequencies. Finally, we integrate the power spectral density (PSD) above and below the spectral noise floor to obtain the wave and noise areas, respectively, and then derive the gravity-wave proportion as

$$p(z) = \frac{wave\ area\ in\ PSD}{wave\ area + noise\ area} \tag{3.8}$$

As $p(z)$ varies only between 0 and 1 (see Figure 3.2b), the derived $E_{pm}(z)$ using Eqs. (3.2) and (3.6) will never be negative (see Figure 3.2c), thus overcoming the issues associated with the noise-variance subtraction method. Tests with forward modeled data and then with real lidar data have shown that for Antarctic winter cases with small error bars, the spectral proportion method and the noise-variance subtraction method give nearly identical profiles of $E_{pm}(z)$, demonstrating the effectiveness of this new method in handling both small and large error cases.

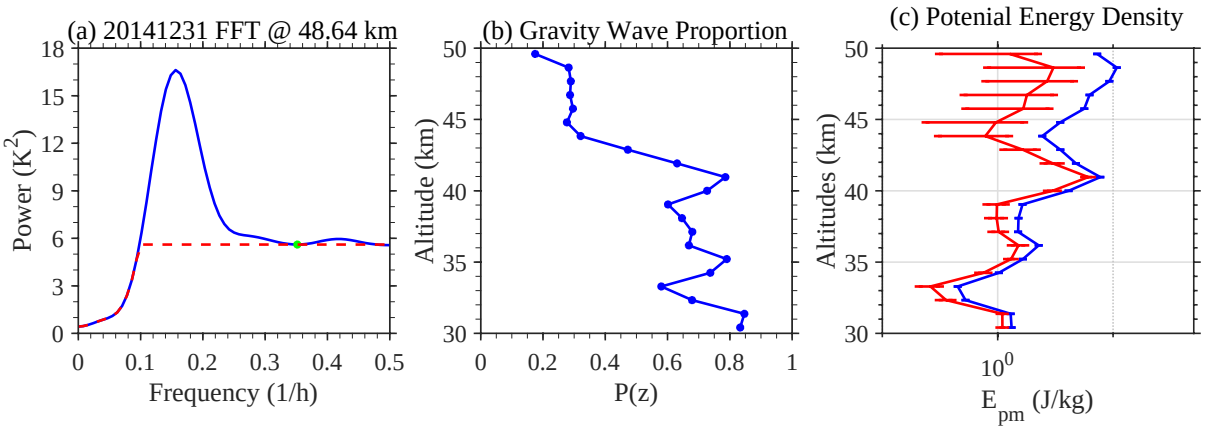

Figure 3.2. Illustration of the spectral proportion method using lidar observations on 31 December 2014 at McMurdo as an example. (a) The mean power spectral density over 1000 Monte Carlo simulated time series at an altitude of 48.64 km. The red dashed line marks the noise floor over the frequency spectra. (b) The gravity wave proportion p(z) profile determined with Eq. (3.8) for this day of lidar observations. (c) Blue curve represents the total potential energy mass density including noise contributions, and the red curve represents the pure gravity wave $E_{pm}(z)$ profile obtained with the spectral proportion method. The horizontal lines denote the error bars associated with the derived $E_{pm}(z)$.

We also compute the altitude profile of gravity wave potential energy volume density $E_{pv}(z)$, for each observational segment, by multiplying $E_{pm}(z)$ with the background atmospheric density $\rho_0(z)$:

$$E_{pv}(z) = \rho_0(z) E_{pm}(z) \tag{3.9}$$

Here the background atmospheric density can be taken from an empirical model such as NRLMSISE-00 [Picone et al., 2002] or determined from the lidar measurements of Rayleigh scattering signals. The latter choice involves first deriving the relative atmospheric density via taking Rayleigh normalization [Chu and Papen, 2005] at a pre-chosen altitude z_N, say 45 km, and then converting to the absolute atmospheric density via scaling the relative density profile with atmospheric density at z_N taken

from the empirical model NRLMSISE-00 [e.g., Chu and Papen, 2005; Gardner et al., 1989; Hauchecorne et al., 1992]. Although the absolute density values still depend on an empirical model, the relative density profiles, which determine the shape of E_{pv} profiles, are given by the real lidar measurements. The corresponding error bars are estimated as:

$$\delta E_{pv}(z) \approx \sqrt{\rho_0^2(z)\delta E_{pm}^2(z) + E_{pv}^2(z)\frac{N_{Total}(z)}{\left[N_{Total}(z) - B\right]^2}} \tag{3.10}$$

where N_{Total} is the total photon count at altitude z and B is the estimated background count. Our tests show that the empirical model density and the lidar-measured density give similar results of E_{pv} profiles, so we choose to use the empirical model density and simplify the error estimation to

$$\delta E_{pv}(z) \approx \rho_0(z)\delta E_{pm}(z) \tag{3.11}$$

3.3. Vertical Profiles of E_{pm} and E_{pv} along with their Scale Heights

$E_{pm}(z)$ vertical profiles are derived for all the qualified segments following the procedures described above, and then $E_{pm}(z)$ in the same months are averaged over the five-year span to obtain the monthly-mean $E_{pm}(z)$ profiles for each of the 12 months through a year. The errors of monthly-mean $E_{pm}(z)$ profiles are computed from the errors of individual $E_{pm}(z)$ profiles via error propagation. The results are shown as three separated groups in the left column of Figure 3.3: Summer (Nov through Feb), winter (May through Aug), and spring/fall (Sep-Oct/Mar-Apr). Similar analyses are done for $E_{pv}(z)$ and $N^2(z)$ and the results are shown in the middle and right columns of Figure 3.3. Using logarithmic scales, $E_{pm}(z)$ profiles increase with altitude while $E_{pv}(z)$ profiles decrease with altitude in general. The increasing and decreasing rates vary from month to month. We quantify such variations via the scale heights H_{pm} and H_{pv} of $E_{pm}(z)$ and $E_{pv}(z)$ as defined below [e.g., Lu et al., 2015]:

$$E_{pm}(z) = E_{pm}(z_0)\exp\left(\frac{z - z_0}{H_{pm}}\right) \Rightarrow \ln[E_{pm}(z)] - \ln[E_{pm}(z_0)] = (z - z_0)/H_{pm} \quad (3.12)$$

$$E_{pv}(z) = E_{pv}(z_0)\exp\left(\frac{z - z_0}{H_{pv}}\right) \Rightarrow \ln[E_{pv}(z)] - \ln[E_{pv}(z_0)] = (z - z_0)/H_{pv} \quad (3.13)$$

Table 3.2. Scale Heights* (H_{pm} and H_{pv}) of E_{pm} and E_{pv} Profiles

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
H_{pm} (km)	9.9	10.5	12.0	12.8	11.0	21.7	14.1	12.6	11.4	10.4	12.2	9.0
	±0.9	±1.0	±1.2	±0.9	±1.0	±2.2	±1.1	±1.1	±0.9	±1.0	±1.4	±1.0
H_{pv} (km)	-34.8	-25.0	-16.4	-13.0	-14.4	-8.8	-12.3	-15.7	-20.6	-27.0	-20.3	-51.8
	±7.3	±5.8	±3.8	±1.8	±2.8	±0.8	±1.6	±3.1	±4.5	±6.3	±4.5	±6.1
H (km)	-7.73	-7.44	-6.99	-6.52	-6.28	-6.33	-6.61	-7.02	-7.36	-7.56	-7.69	-7.77
	±0.12	±0.11	±0.12	±0.13	±0.15	±0.17	±0.18	±0.19	±0.18	±0.16	±0.15	±0.13

* Positive (negative) scale heights denote the cases of corresponding parameters increasing (decreasing) with increasing altitude. The signs of ± denote the standard errors.

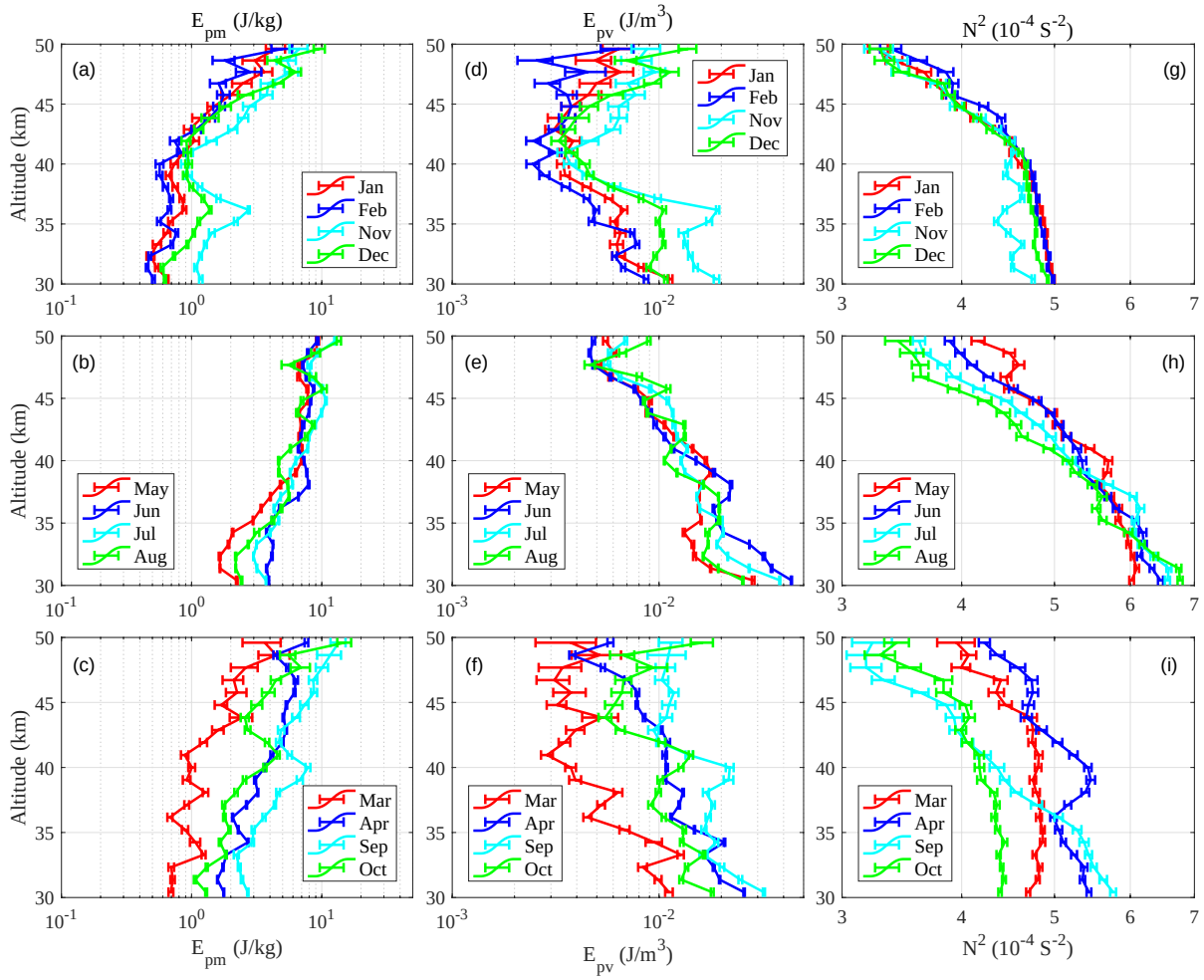

Figure 3.3. Five-year (2011–2015) monthly mean vertical profiles of E_{pm} (left column), E_{pv} (middle column), and N² (right column) with error bars for (top) summer months (November through February), (middle) winter months (May through August), and (bottom) fall/spring months (March, April, September, and October) at McMurdo, Antarctica. Error bars represent the errors of monthly-mean profiles, which are computed from the errors of individual segment profiles via error propagation.

Figure 3.4 illustrates the linear fits to the logarithmic $E_{pm}(z)$ and $E_{pv}(z)$ profiles using equations (3.12) and (3.13). The scale heights determined from the linear fits are summarized in Table 3.2 along with the atmospheric density scale height H. Under the definitions of equations (3.13), negative H_{pv} indicates the

decrease of $E_{pv}(z)$ with increasing altitude. These scale heights are indicators for the dissipation of gravity wave energy: Assuming that the background wind and the static stability do not change with altitude (constant vertical wavelength), E_{pv} remains constant with altitude for a conservative gravity wave, and decreases with altitude when wave dissipation occurs [Wilson et al., 1991; Whiteway and Carswell, 1995]. Thus, for a monochromatic upward propagating gravity wave, the absolute magnitude of H_{pv} is infinity if the wave is non-dissipative and its vertical wavelength is constant with height, but H_{pv} is negative and finite if the wave dissipates. The smaller the absolute magnitude of H_{pv}, the faster the wave dissipates with altitude. Likewise, a positive H_{pm} larger than the atmospheric density scale height H means that E_{pm} grows with altitude more slowly than a monochromatic conservative wave, with larger positive values of H_{pm} indicating stronger dissipation [Mze et al., 2014; Lu et al., 2015]. However, the actual situations are much more complicated than this simple picture because the observed potential energy densities are the sum of contributions from multiple gravity waves generated by various sources at multiple heights and locations. Furthermore, when the mean wind varies slowly with height (which is usually the case), only are the energy flux density and momentum flux density constant with altitude for a monochromatic and conservative gravity wave, but the potential and kinetic energy densities are not necessarily constant [Becker and Vadas, 2018]. The obtained $E_{pm}(z)$ and $E_{pv}(z)$ in Figure 3.3 show complicated profiles, e.g., $E_{pv}(z)$ profiles in summer months decrease with altitude first and then increase or become nearly vertical, reflecting the interplay of multiple factors. Because of this complication we choose not to further study the turning features in the summer $E_{pv}(z)$ profiles, but rather apply simple linear fits to quantify the slopes. The H_{pv} values in Table 3.2 indicate that the wave dissipation in winter is much more severe than that in summer in the altitude range of 30–50 km.

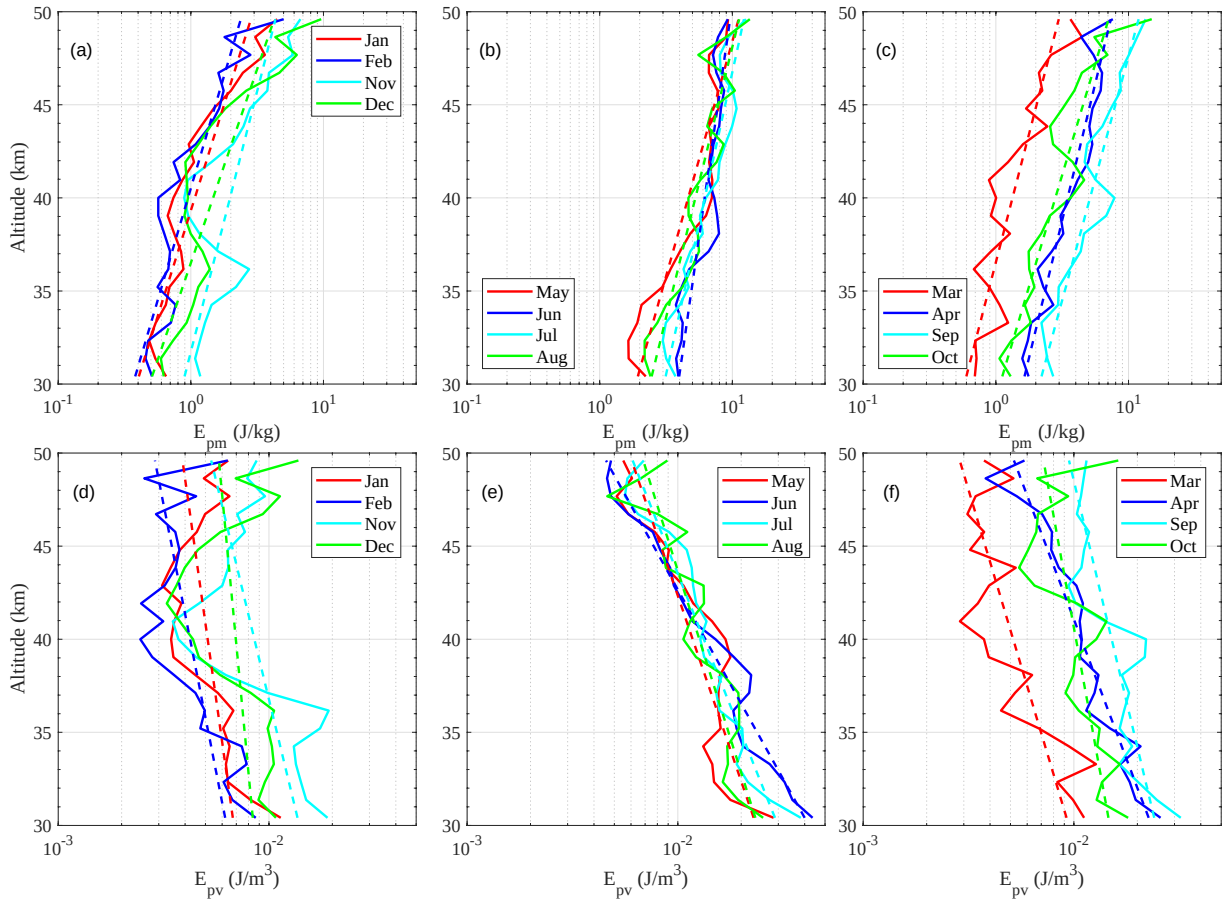

Figure 3.4. Linear fits (dashed lines) to the vertical profiles (solid lines) of E_{pm} (top panel) and E_{pv} (bottom panel) for summer months (November through February) on the left, winter months (May through August) in the middle, and fall/spring months (March, April, September, and October) on the right at McMurdo, Antarctica.

3.4. Statistics and Lognormal Distributions of E_{pm} and E_{pv}

The altitude mean $\overline{E}_{pm}$ of gravity wave potential energy mass density from 30 to 50 km is derived for each observational segment as

$$\overline{E}_{pm} = \frac{1}{N_z}\sum_{k=1}^{N_z} E_{pm}(z_k) \tag{3.14}$$

and the associated error bar is reduced to

$$\delta\overline{E}_{pm} = \frac{1}{N_z}\sqrt{\sum_{k=1}^{N_z}[\delta E_{pm}(z_k)]^2} \tag{3.15}$$

where N_z is the number of data points within an altitude profile. The altitude mean $\overline{E}_{pv}$ of gravity wave potential energy volume density from 30 to 50 km and associated error are derived for each observational segment as

$$\overline{E}_{pv} = \frac{1}{N_z}\sum_{k=1}^{N_z} E_{pv}(z_k) \tag{3.16}$$

$$\delta\overline{E}_{pv} = \frac{1}{N_z}\sqrt{\sum_{k=1}^{N_z}[\delta E_{pv}(z_k)]^2} \tag{3.17}$$

We obtain 354 data points of $\overline{E}_{pm}$ for all the qualified data through five years, and the same number of data points for $\overline{E}_{pv}$. Here each data point represents one observational segment. The mean $\overline{E}_{pm}$ and $\overline{E}_{pv}$ along with their standard deviations are summarized in Table 3.3 for the entire dataset and for summer, winter, and spring/fall. $\overline{E}_{pm}$ and $\overline{E}_{pv}$ histograms for all 354 data points are plotted in Figures 3.5a and 3.5f, respectively. These distributions are skewed and deviate significantly from normal distributions. $\overline{E}_{pm}$ and $\overline{E}_{pv}$ histograms are fitted with the lognormal distribution

$$h(x) = \frac{A}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \tag{3.18}$$

Table 3.3. Data Statistics and Lognormal Fitting Parameters to $\overline{E}_{pm}$ and $\overline{E}_{pv}$ Histograms

$\overline{E}_{pm}$ (J/kg)	Total Data Points	Mean ± Std. Deviation	μ	σ	A	Correlation	MPV (J/kg)
Total	354	3.82±3.42	0.253	0.81	93.98	97%	1.29
Summer	142	1.84±1.35	0.055	0.60	55.56	99%	1.06
Winter	122	6.16±3.82	1.158	0.63	15.28	82%	3.18
Spring + Fall	90	3.76±3.16	0.406	0.75	21.97	92%	1.50
$\overline{E}_{pv}$(10^{-2} J/m^3)	Total Data Points	Mean ± Std. Deviation	μ	σ	A	Correlation	MPV
Total	354	1.08±0.85	-5.25	0.65	104.71	99%	0.52
Summer	142	0.71±0.46	-5.48	0.60	56.57	99%	0.42
Winter	122	1.48±0.99	-4.86	0.49	24.01	91%	0.77
Spring + Fall	90	1.13±0.87	-5.25	0.72	26.48	95%	0.52

The fitting parameters μ, σ and A are listed in Table 3.3. The correlation coefficients for the lognormal fittings are very high, 0.97 and 0.99 for $\overline{E}_{pm}$ and $\overline{E}_{pv}$, respectively, at 95% confidence levels. Therefore, the mean $\overline{E}_{pm}$ and $\overline{E}_{pv}$ between 30 and 50 km are log-normally distributed at McMurdo. The most probable value (MPV) is given by $x_{MPV} = e^{\mu}$, which is 1.29 J/kg and 5.2×10^{-3} J/m^3 for $\overline{E}_{pm}$ and $\overline{E}_{pv}$, respectively, as summarized in Table 3.3. $\overline{E}_{pm}$ ranges from less than 1 to nearly 20 J/kg, and $\overline{E}_{pv}$ ranges from 1.4×10^{-3} to 4.5×10^{-2} J/m^3; however, the lognormal distribution indicates that most $\overline{E}_{pm}$ and $\overline{E}_{pv}$ values are small.

Dividing the $\overline{E}_{pm}$ data into seasons, we plot the histograms for summer, spring/fall, and winter distributions in Figures 3.5b–3.5d. The differences among seasons are striking—the summer $\overline{E}_{pm}$ are clustered in a narrow band of low values

(with a MPV of 1.06 J/kg and ranging mainly from 0.2 to 4 J/kg) with a nearly perfect lognormal distribution at 0.99 correlation, while the winter $\overline{E}_{pm}$ shifts the MPV to ~3.2 J/kg and exhibits a very wide distribution range. The spring/fall distribution lies in between the summer and winter cases with intermediate MPV and distribution width. It is worth pointing out from Figure 3.5 that most of the large $\overline{E}_{pm}$ values (>9 J/kg) in the overall lognormal distribution (Figure 3.5a) occur in the winter season. In contrast, the small $\overline{E}_{pm}$ values (<1 J/kg) in Figure 3.5a receive very minor contributions from the winter season but occur mainly during summer. Lognormal distributions given by equation (3.18) are also fit to the seasonal histograms, and the fitting parameters are summarized in Table 3.3. Again, the correlation coefficients are high, ranging from 0.82 in winter to 0.99 in summer. The annual cycle of $\overline{E}_{pm}$ (with summer minima and winter maxima) clearly shows up in the most probable values. Because of the large width of the distributions for $\overline{E}_{pm}$ during winter, the 122 data points in this case are less ideal for our statistics, and more data points would increase the correlation. The spring/fall seasons have only 90 data points, which is also less ideal than the summer case with 142 points distributed in a narrow range. The distributions of $\overline{E}_{pv}$ seasonal histograms (Figures 3.5g–3.5i) are similar to those of $\overline{E}_{pm}$ as demonstrated in the middle column of Figure 3.5.

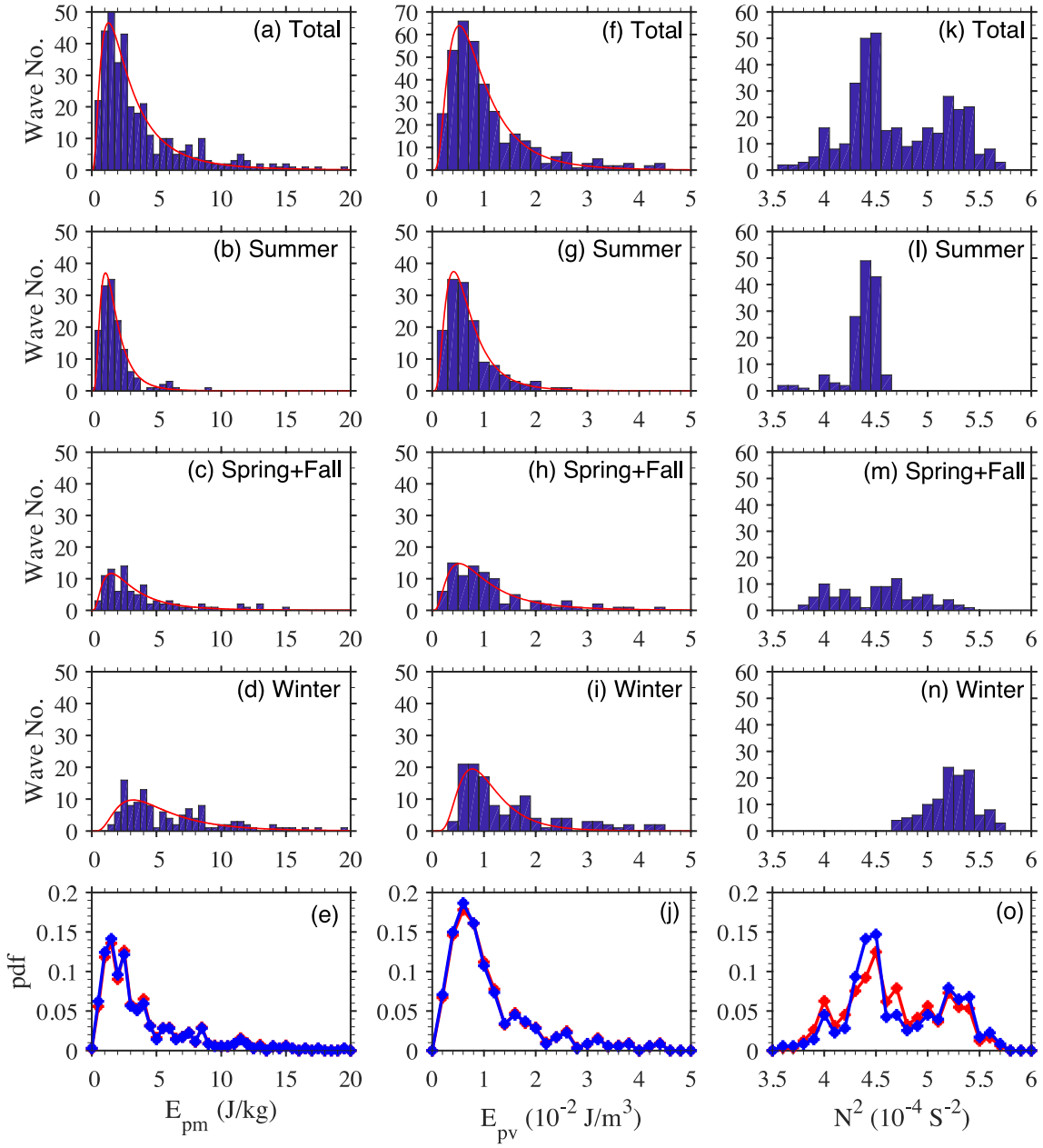

Figure 3.5. Lognormal distribution of altitude-mean $\overline{E}_{pm}$ (left column), lognormal distribution of altitude-mean $\overline{E}_{pv}$ (middle column), and distribution of altitude-mean $\overline{N^2}$ (right column) in the whole year, summer, spring/fall, and winter seasons from the first to the fourth rows. The bottom row shows the comparison of the original (blue) and normalized (red) distributions of $\overline{E}_{pm}$, $\overline{E}_{pv}$, and $\overline{N^2}$.

Histograms of altitude-mean $\overline{N^2}$ of total 354 data points and of summer, winter, and spring/fall seasons are plotted in the right column of Figure 3.5. Apparently, $\overline{N^2}$ does not obey a lognormal distribution. A striking feature in the year-round distribution is the two independent peaks corresponding to summer and winter, respectively. Summer $\overline{N^2}$ clusters narrowly around 4.4×10^{-4} s^{-2}, whereas winter $\overline{N^2}$ distributes in a much wider range from 4.7×10^{-4} to 5.7×10^{-4} s^{-2}, centering around 5.25×10^{-4} s^{-2}. The two peaks in Figure 3.5m correspond to spring and fall. Spring has lower $\overline{N^2}$ values than the summer peak, while fall $\overline{N^2}$ values are between summer and winter. The fall peak of $\overline{N^2}$ is higher than that during spring by ~0.6×10^{-4} s^{-2}. Considering that different seasons have different numbers of data, which may affect the overall distribution of $\overline{N^2}$, we normalize each season with its number of data points and then combine all seasons together. Such normalized results are compared with the original PDF in Figure 3.5o. The $\overline{N^2}$ distribution does change somewhat from the original PDF but it is still not a lognormal distribution. The similar normalization procedure is applied to $\overline{E}_{pm}$ and $\overline{E}_{pv}$, and the results are compared with the original $\overline{E}_{pm}$ and $\overline{E}_{pv}$ in Figures 3.5e and 3.5j. Obviously, the normalized $\overline{E}_{pm}$ and $\overline{E}_{pv}$ distributions are nearly identical to the original PDFs and closely resemble the lognormal distributions. Lognormal distributions of $\overline{E}_{pm}$ and $\overline{E}_{pv}$ observed by the lidar are in good agreement with radiosonde and satellite observations that reveal lognormal distributions of gravity wave potential energy density and momentum flux [Baumgaertner and McDonald, 2007; Alexander et al., 2008; Hertzog et al., 2012; Murphy et al., 2014]; however, we are not aware of previous observations of $\overline{N^2}$ distributions.

3.5. Seasonal and Inter-Annual Variations of $\overline{E}_{pm}$ and $\overline{E}_{pv}$

We now investigate how $\overline{E}_{pm}$, $\overline{E}_{pv}$, and $\overline{N^2}$ vary with time. All the 354 data points of $\overline{E}_{pm}$ and $\overline{E}_{pv}$ are plotted as a function of day of year (DOY) in Figures 3.6a

and 3.6d, ranging from 1 January 2011 to 31 December 2015. Each data point represents the measurements of one 12-h segment, and the error bars are the associated $\delta\overline{E}_{pm}$ and $\delta\overline{E}_{pv}$ computed with Eqs. (15) and (17). An obvious and repeatable pattern in Figure 3.6 is the seasonal variations of $\overline{E}_{pm}$ and $\overline{E}_{pv}$, with summer minima and winter maxima every year for McMurdo, despite the fact that $\overline{E}_{pm}$ and $\overline{E}_{pv}$ of individual segments could vary substantially from observation to observation. Such seasonal variations are in general similar to the previous lidar observations of stratospheric $\overline{E}_{pm}$ at Rothera [Yamashita et al., 2009] and Davis [Kaifler et al., 2015] in Antarctica. Nevertheless, the multiple years of high-quality lidar data at McMurdo provide the first time series of repeatable seasonal patterns of $\overline{E}_{pm}$ and $\overline{E}_{pv}$ in the Antarctic stratosphere. Similar analyses are done for $\overline{N^2}$ and the results are plotted in Figure 3.7. The altitude-mean $\overline{N^2}$ also exhibits repeated seasonal patterns, which varies from minima (less than 4×10^{-4} s^{-2}) around October to maxima (over 5.5×10^{-4} s^{-2}) in winter. The peak-to-peak variation of $\overline{N^2}$ is ~40%, with the mean around 4.7×10^{-4} s^{-2}. Although the maxima for $\overline{E}_{pm}$ and $\overline{N^2}$ both occur around mid-winter, the minimum $\overline{N^2}$ occurs around October but the minimum $\overline{E}_{pm}$ occurs around late January and early February. Such differences indicate that, even at polar latitudes, the seasonal change of $\overline{N^2}$ is unlikely the major cause of the observed seasonal variations of $\overline{E}_{pm}$ and $\overline{E}_{pv}$, although it may contribute.

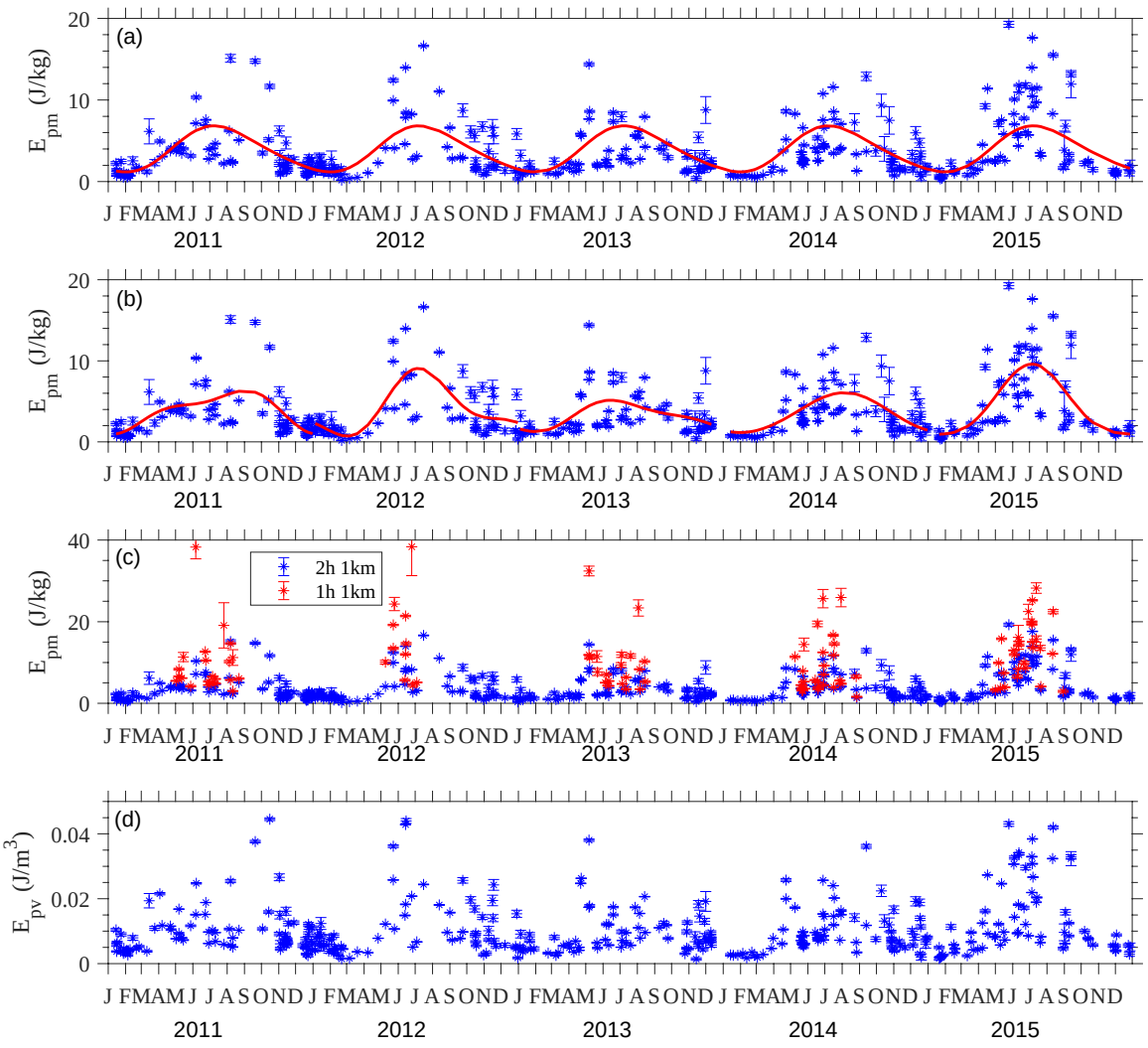

Figure 3.6. (a-c) Altitude-mean $\overline{E}_{pm}$ (J/kg) and (d) altitude-mean $\overline{E}_{pv}$ (J/m^3) averaged over the 30–50 km altitude range for all the observations over 5 years from 2011 through 2015. The blue asterisks in all panels denote the actual $\overline{E}_{pm}$ or $\overline{E}_{pv}$ observation at temporal and spatial resolutions of 2 h and 1 km during single observational segments with their errors calculated using Eqs. (15) and (17). The red lines are (a) overall annual+semianuual fits for 5 years and (b) single year annual+semianuual fits for 5 individual years, respectively. The red dots in (c) are the $\overline{E}_{pm}$ derived under a higher temporal resolution of 1 h (only for winter months).

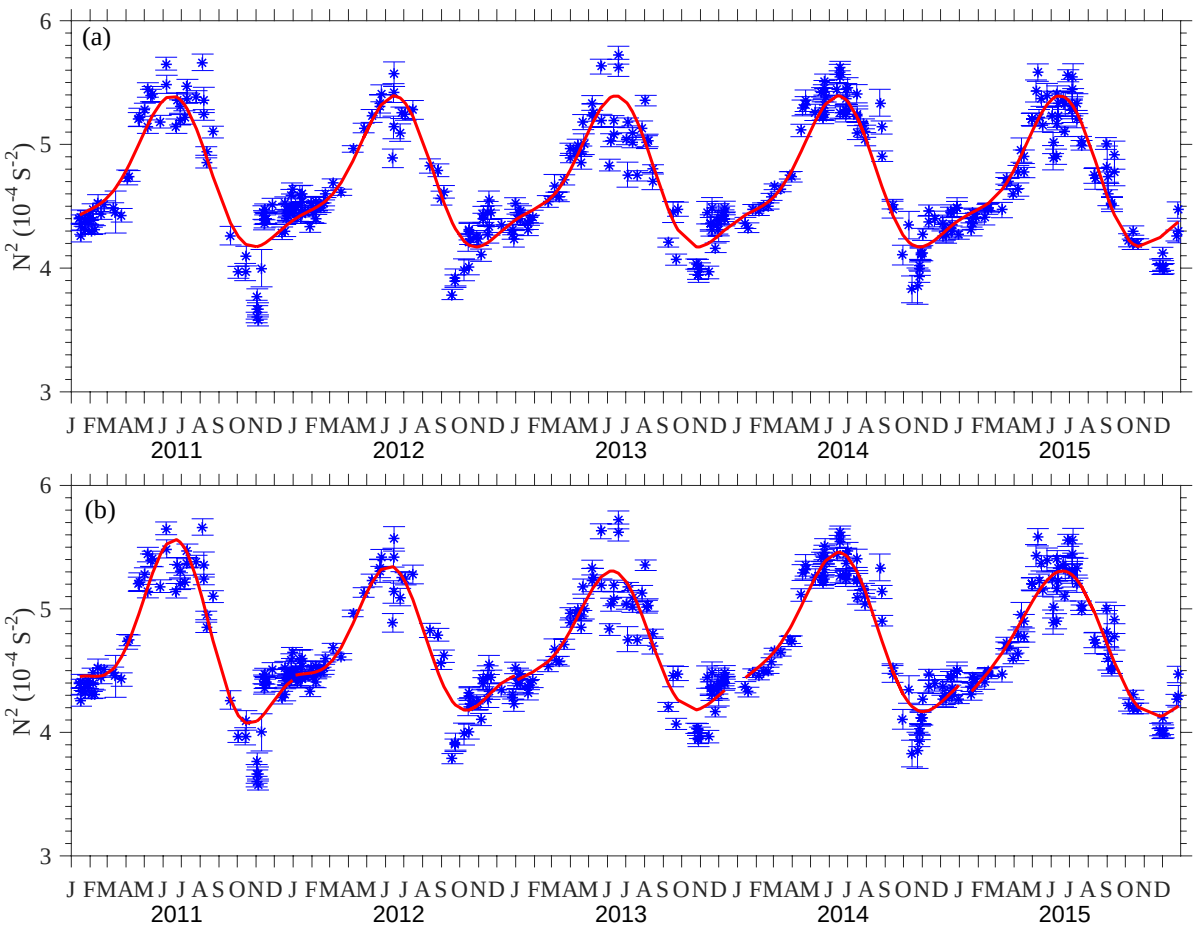

Figure 3.7. Altitude-mean $\overline{N^2}$ from 30 to 50 km through five years of 2011–2015. Each blue asterisk denotes a single observational segment with error bar. The red lines are (a) overall annual+semianuual fits for 5 years and (b) single year annual+semianuual fits for 5 individual years, respectively.

To quantify the seasonal variations of $\overline{E}_{pm}$ and $\overline{N^2}$ to compare with Yamashita et al. [2009], Figure 3.8 shows the 5-year average time series of $\overline{E}_{pm}$ and $\overline{N^2}$ with data binned every 5 days. Then we apply the harmonic fittings to the folded data as well as to the 5-year time series:

$$y = A_0 + A_{12}\cos\left[\frac{2\pi}{365}(x-\Phi_{12})\right] + A_6\cos\left[\frac{2\pi}{365/2}(x-\Phi_6)\right] \tag{3.19}$$

where A_0 is the annual mean, A_{12} and A_6 are the amplitudes of annual and semiannual variations, and Φ_{12} and Φ_6 are the corresponding phases. Such harmonic fittings to multiple years of $\overline{E}_{pm}$ and $\overline{N^2}$ help mitigate the uneven sampling dictated by weather conditions. There are multiple ways to do the fitting for different purposes. To characterize the overall seasonal variations through five years, the harmonic fitting is applied to all the five years of data and yields the red curve shown in Figure 3.6a, which assumes constant amplitudes and continuous phases through these five years. The obtained fitting parameters are summarized in the first row of Table 3.4. The annual variation dominates over the semi-annual variation. The annual variation peaks in the mid-winter (July), while the semi-annual variation peaks about one month earlier (see the first row of Table 3.4). The harmonic fitting to the $\overline{E}_{pm}$ data shown in Figure 3.8a gives results (see the 2nd row in Table 3.4) nearly identical to the first fitting. Similar harmonic fittings are performed on $\overline{N^2}$ as indicated by the red lines in Figure 3.7a and 3.8b. The fitting parameters are summarized in the first two rows of Table 3.5.

Table 3.4. Fitting Parameters and Errors for $\overline{E}_{pm}$ in Figures 3.6 and 3.8

Case	A_0 (J/kg)	A_{12} (J/kg)	Φ_{12} (Day)	A_6 (J/kg)	Φ_6 (Day)
Figure 3.6a	3.89±0.30	2.75±0.40	200±9	0.35±0.43	166±34
Figure 3.8a	3.78±0.34	2.64±0.48	202±11	0.32±0.48	161±43
3.6b (2011)	3.92±0.60	2.26±0.80	210±22	0.87±0.88	88±26
3.6b (2012)	4.26±0.69	3.54±1.00	204±18	1.44±1.01	180±20
3.6b (2013)	3.23±0.58	1.69±0.79	191±29	0.53±0.83	143±44
3.6b (2014)	3.51±0.59	2.46±0.85	216±18	0.12±0.80	226±39
3.6b (2015)	4.35±0.43	4.32±0.57	187±9	0.92±0.57	186±19

Table 3.5. Fitting Parameters and Errors for $\overline{N^2}$ in Figures 3.7 and 3.8

Case	A_0 ($\times 10^{-4} s^{-2}$)	A_{12} ($\times 10^{-4}$ s^{-2})	Φ_{12} (Day)	A_6 ($\times 10^{-4}$ s^{-2})	Φ_6 (Day)
Figure 3.7a	4.70±0.02	0.55±0.02	154±3	0.17±0.02	178±4
Figure 3.8b	4.69±0.03	0.56±0.04	154±5	0.17±0.04	180±8
3.7b (2011)	4.70±0.05	0.61±0.06	156±7	0.28±0.08	178±7
3.7b (2012)	4.67±0.04	0.49±0.05	145±8	0.20±0.06	167±8
3.7b (2013)	4.67±0.04	0.50±0.05	146±8	0.16±0.06	171±12
3.7b (2014)	4.75±0.03	0.59±0.04	153±5	0.15±0.05	181±10
3.7b (2015)	4.67±0.02	0.57±0.02	160±3	0.09±0.02	189±9

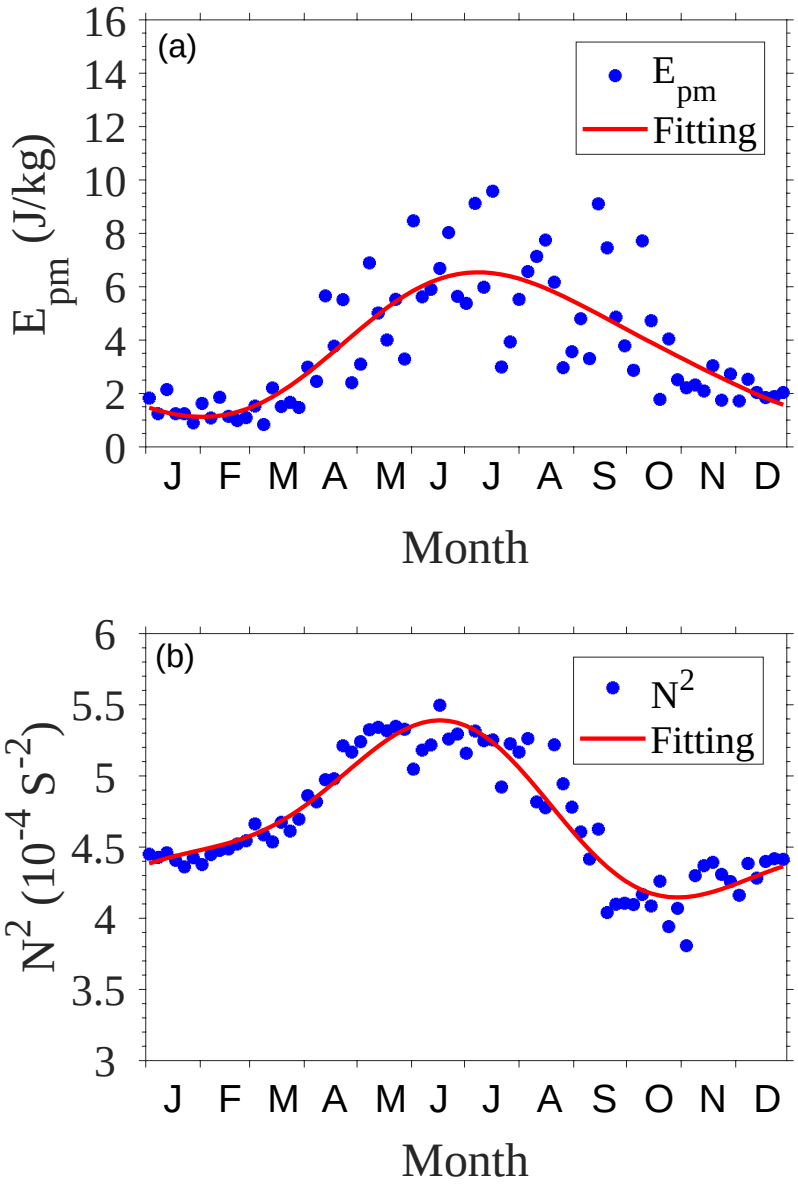

Figure 3.8. Five-year average annual cycles of $\overline{E}_{pm}$ and $\overline{N^2}$ binned every 5 days at McMurdo. The red lines are the harmonic fits given by Equation (3.19).

To assess the inter-annual variability, the harmonic fitting given by Eq. (3.19) is applied to each individual year as illustrated in Figure 3.6b, and the fitting parameters are summarized in the 3rd to 7th rows of Table 3.4. The annual variation amplitudes in 2015 and 2012 are larger than those in 2011, 2013 and 2014. It will be shown later that this difference is likely related to critical level filtering but not to

variations in static stability. The semi-annual variation phase in 2011 is nearly out of phase with that in other years. Overall the five years of lidar data exhibit quite large year-to-year variability, which may also be related to the high variability of the timing of the vortex breakdown in November/December from year to year [e.g., Lübken et al., 2015]. Similar harmonic fitting analyses are applied to $\overline{N^2}$ data in Figure 3.7b, and the fitting parameters are summarized in Table 3.5. The inter-annual variations of $\overline{N^2}$ are much smaller in both amplitude and phase when compared to $\overline{E}_{pm}$. The $\overline{N^2}$ fitting results for the annual and semi-annual variations do not have large changes over the 5 years. The annual component peaks at ~154 DOY while the semi-annual component maximizes at ~178 DOY.

The temperature data resolutions (2 h by 1 km) used above are chosen to achieve sufficient SNRs for both winter and summer. If we consider only winter when the SNR is significantly higher (due to the low solar background), temperature data can be retrieved at 1 h and 1 km resolutions. Indeed, we did such a test in Figure 3.6c. That is, when using resolutions of 1 h and 1 km to derive winter temperatures, the obtained $\overline{E}_{pm}$ can be as large as 40 J/kg. For the original data processing, the integration window is 2 h but the oversampled display resolution is 1 h, so the minimum resolved temporal resolution of waves included in the original $\overline{E}_{pm}$ lies in between 2–4 h. Using 1-h time resolution enables us to include all waves with periods of 2 h and longer. Figure 3.6c shows that including the full wave spectra from 2–4 h substantially enhances (nearly doubles) $\overline{E}_{pm}$ at altitudes of 30–50 km. The absolute values of $\overline{E}_{pm}$ are comparable to lidar observations at both Rothera and Davis with similar seasonal variations [Yamashita et al., 2009; Alexander et al., 2011; Kaifler et al., 2015], but are larger than at the South Pole where $\overline{E}_{pm}$ is nearly constant throughout the year at a low value (~2.8 J/kg) [Yamashita et al., 2009]. Note that different wave spectra are included due to different data resolutions used in these

studies. Gravity waves with periods of ~1–6 h and vertical wavelengths of 2–30 km are included in the study of Yamashita et al. whereas Kaifler et al. put more emphasis on waves with periods longer than 2 h and vertical wavelengths of ~4–20 km. In this study, we include waves with periods of ~2–11 h and vertical wavelengths of ~2–30 km.

3.6. Monthly-Mean Vertical Wavenumber Spectra

In a previous study we investigated the seasonal-mean vertical wavenumber spectra and revealed the power spectral density (PSD) of gravity waves with vertical wavelengths of 5–20 km increasing from summer minima to winter maxima at McMurdo [Zhao et al., 2017]. Such seasonal variations of vertical wavenumber PSD are qualitatively consistent with the seasonal variations of $\overline{E}_{pm}$. Here we expand the study of Zhao et al. [2017] and compute monthly-mean PSDs (see Figure 3.9). The variations throughout a year can be clearly identified. Maximum PSD for vertical wavelengths between 5 and 20 km is evident during the winter months, which corresponds to on average longer vertical wavelengths in winter. The PSD for the shortest vertical wavelengths remains nearly identical throughout the year. The characteristic vertical wavenumber (m*), i.e., the transition vertical wavenumber between the positive and negative slopes in the PSD spectra, is ~1×10^{-4} m^{-1} from March to July and about 6.5×10^{-5} m^{-1} in August and September, whereas there is no distinguishable m* from October to February.

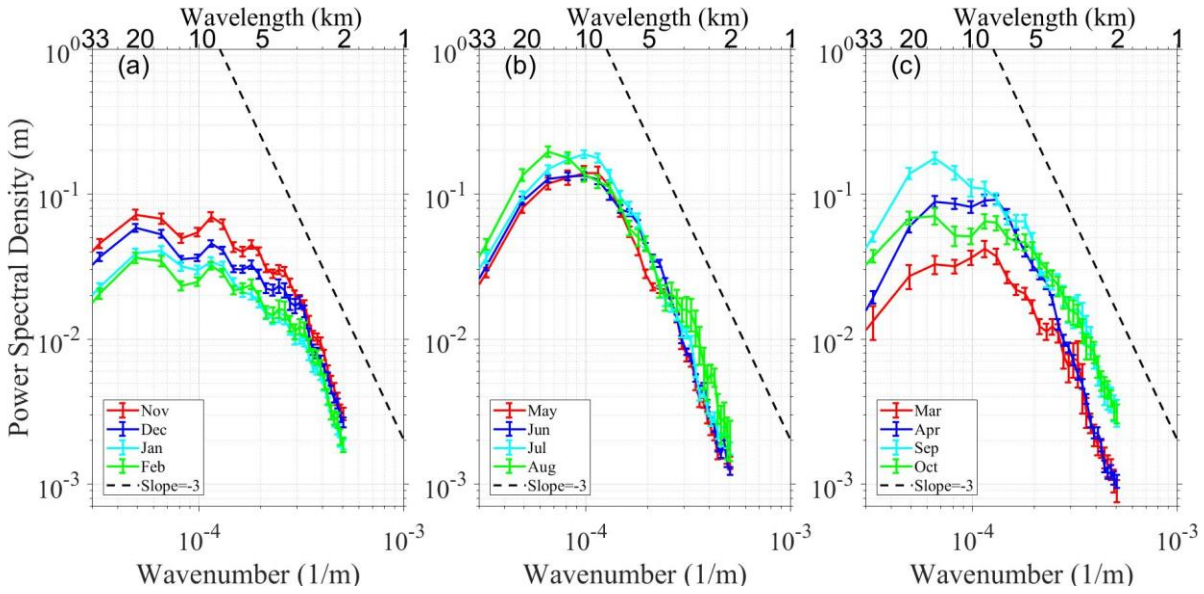

Figure 3.9. Five-year mean distributions of vertical wavenumber spectra for each month of the year. (left) Summer months (November through February), (middle) winter months (May through August), and (right) fall and spring months (March, April, September, and October).

To quantify the slope changes through a year, we perform linear fittings to the vertical wavenumber spectra in the range of 2–10 km. The resulting slopes and associated errors are summarized in Table 3.6. The slopes exhibit distinct seasonal signatures—steep slopes around -3 from April through July, but much shallower around -2 from October through February.

Table 3.6. Slopes of Monthly-Mean Vertical Wavenumber Spectra in Figure 3.9

Range	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2–10 km	-1.76	-1.63	-2.46	-3.17	-3.14	-3.19	-3.20	-2.64	-2.35	-2.00	-1.90	-1.71
	±0.24	±0.30	±0.27	±0.28	±0.19	±0.26	±0.18	±0.23	±0.21	±0.26	±0.29	±0.29

Compared to the winter values, summer PSD values between 5 and 20 km are significantly smaller. As for the fall (March and April) and spring (September and October) seasons, it is interesting that March and September exhibit PSDs nearly identical to the summer and winter PSDs, respectively, while April and October show the transitions from summer to winter and from winter to summer, respectively. Such PSD variations in the m-spectra are consistent with the seasonal variations of $\overline{E}_{pm}$ and $\overline{E}_{pv}$ inferred from Figures 3.6 and 3.8a. The characteristic vertical wavelengths increase from a shorter wavelength in March (representing summer) to a longer wavelength in September (representing winter), and then return to shorter wavelengths in October and summer. Such spectral behaviors are also consistent with the seasonal results in Zhao et al. [2017].

3.7. Correlations of $\overline{E}_{pm}$ with Polar Vortex, Wind Rotation, and Wind Speeds

Whiteway et al. [1997] and Duck et al. [1998] have found that the amount of gravity wave energy in the upper stratosphere at Eureka in the Arctic is related to the position of the stratospheric polar vortex. They showed that gravity wave activity was a maximum within the westerly jet at the edge of the vortex, a minimum inside the vortex near its center, and intermediate outside the vortex [Whiteway et al., 1997]. Baumgaertner and McDonald [2007] have found from the CHAMP/GPS observations that $\overline{E}_{pm}$ at 23–27 km shows gravity wave enhancement within ±10 deg from the vortex edge. For the gravity wave potential energy density from 30 to 45 km at Rothera, Antarctica, Yamashita et al. [2009] have found that all the large $\overline{E}_{pm}$ events were observed between +5 and -10 deg from the vortex edge. To examine the polar vortex effects, we calculate the McMurdo position relative to the polar vortex edge at an altitude of 32 km using global wind and temperature data from the Modern Era Retrospective Analysis for Research and Applications version 2 (MERRA-2)

[Bosilovich et al., 2015]. The Antarctic vortex edge is defined as follows. The "Q diagnostic" measures the relative contribution of strain and rotation in the flow [Malvern, 1969]; Q is negative in regions where rotation is dominant and positive where there is strong shear. Starting at the center of cyclonic wind systems, the Q diagnostic is integrated around concentric stream function (ψ) contours (aligned parallel to the rotational wind) and nodes in Q are vortex edge candidates. The vortex edge is defined as the candidate ψ contour where line-integrated wind speed maximizes. See Harvey et al. [2002] for a more extensive description of the algorithm. According to the results shown in Figure 3.10a, there are small $\overline{E}_{pm}$ values (<4 J/kg) for all possible positions of McMurdo relative to the polar vortex edge; however, the major high-level gravity wave activity ($\overline{E}_{pm}$ >10 J/kg) occurs when McMurdo is about +3 to -23 deg relative to the vortex edge. In particular, the very large $\overline{E}_{pm}$ values (>15 J/kg) occur when McMurdo lies well inside the polar vortex (about 8–23 deg poleward from the vortex edge). While this result appears to contradict some previous observational results mentioned above, what we may be seeing is simply a wintertime maximum in GW activity. In other words, without another lidar site that is coincident with the vortex edge in midwinter, it is not clear whether even higher gravity wave activity occurs at the edge during the May–August timeframe. Figure 3.10b illustrates the number of lidar observations during the presence of polar vortex. The distribution starts in March (9 points), peaks in May through July, and ends in November (20 points). The Antarctic vortex is not present from December through February (see also Figure 3.13e). Figure 3.10b also illustrates the distance of McMurdo station from the vortex edge, which clearly shows that on average McMurdo stays deep inside the vortex during mid-winter.

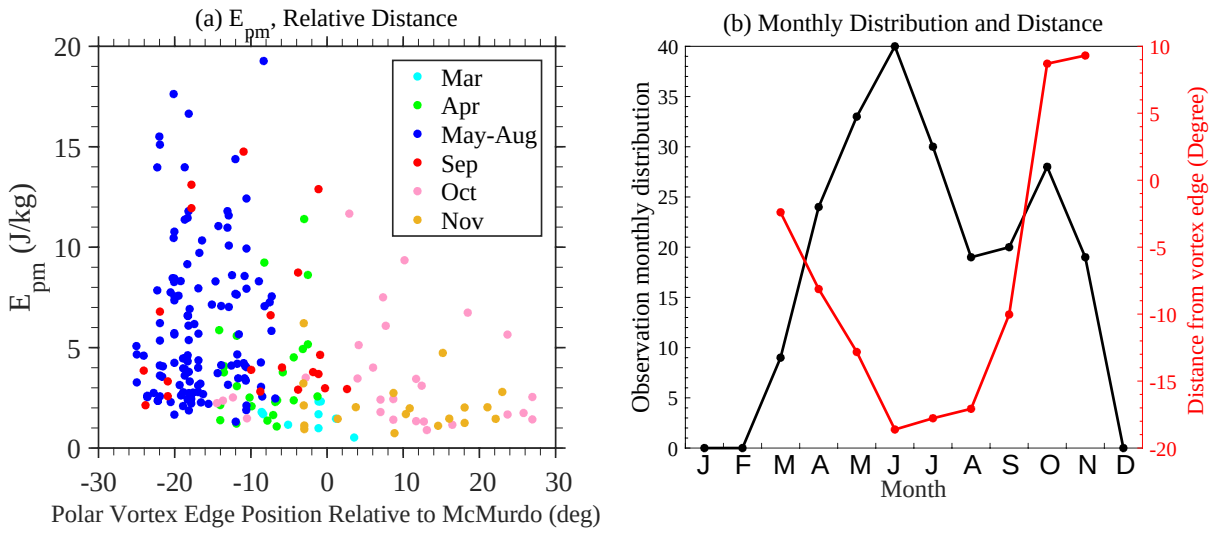

Figure 3.10. (a) $\overline{E}_{pm}$ versus polar vortex edge position relative to McMurdo color-coded by season. Zero degree means that McMurdo is at the polar vortex edge, negative degrees mean that McMurdo is inside the vortex, and positive degrees mean that McMurdo is outside the vortex. (b) Monthly distribution of the number of lidar observations (black line) during the presence of polar vortex in Antarctica and the monthly mean distance between McMurdo and the vortex edge (red line), as given above. The Antarctic vortex is not present from December through February.

Yamashita et al. [2009] suggested that critical level filtering of gravity waves by the background winds, orographic wave generation in the lower troposphere, and *in-situ* generation of gravity waves by large wind shear at the edge of the polar vortex could have contributed to the observed seasonal variations of $\overline{E}_{pm}$ at Rothera. Following these hints, wind data from ECMWF at McMurdo are used here to inspect the correlations between $\overline{E}_{pm}$ and several potential factors. In Figures 3.11a, 3.11b, and 3.11c, we plot the five years of lidar-observed $\overline{E}_{pm}$ along with wind rotation angles from 11 to 30 km, absolute wind speeds at 3 km and at 30 km, respectively, given by the ECMWF reanalysis data. Despite some fluctuations in the data, it is obvious from these three figures that $\overline{E}_{pm}$ are anti-correlated with the wind rotation

angles and positively correlated with the wind speeds at both 3 and 30 km. The correlation coefficients of monthly-mean $\overline{E}_{pm}$ with monthly-mean wind rotation angles, near-surface winds at 3 km, and stratospheric winds at 30 km (see Figure 3.11) are –0.62, +0.87 and +0.80, respectively. All these correlations have 95% confidence levels. In the following we show that these correlations are significant.

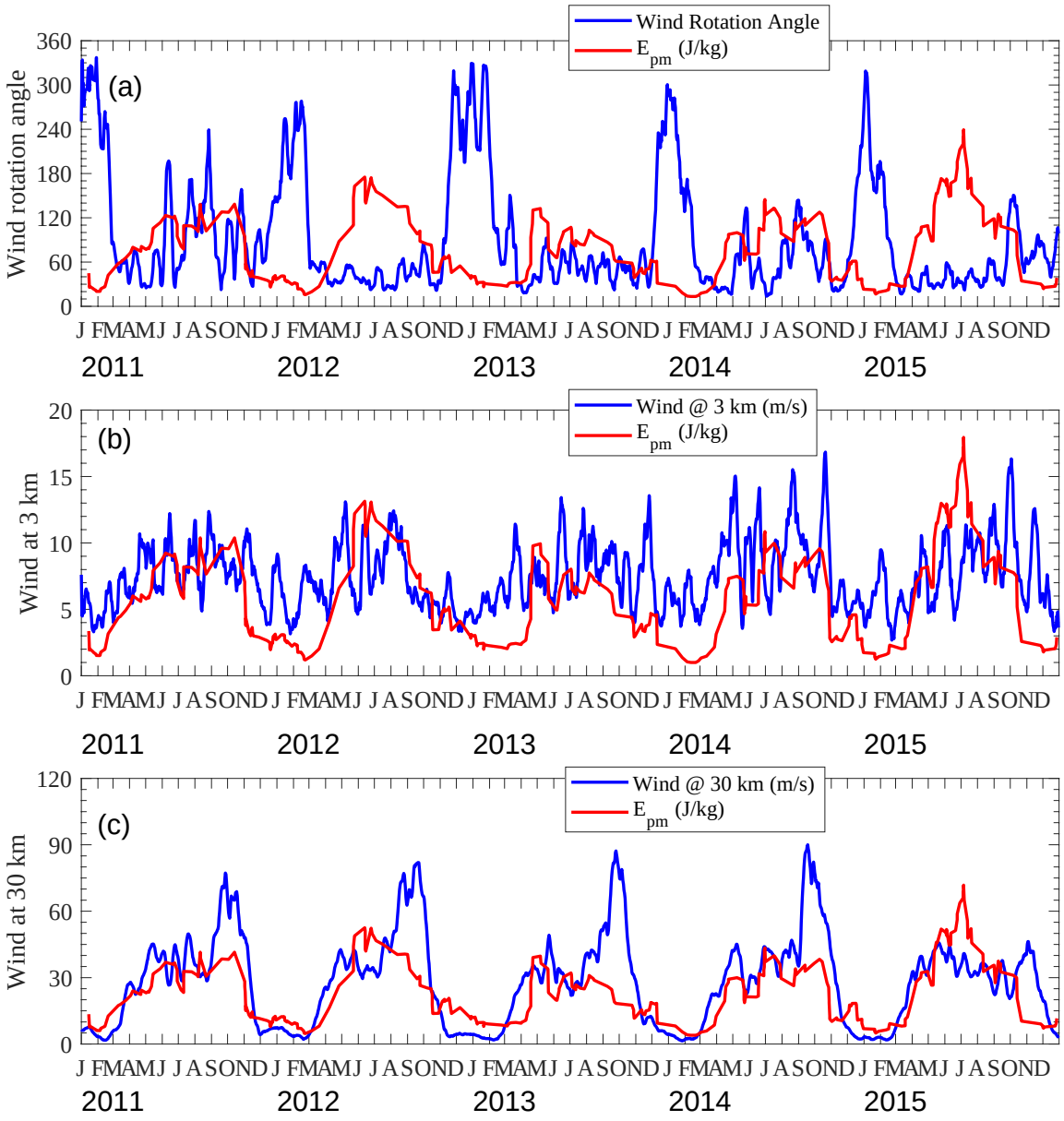

Figure 3.11. Time series of $\overline{E}_{pm}$ versus (a) wind rotation angle between 11 and 30 km, (b) surface wind at 3 km, and (c) stratospheric wind at 30 km given by ECMWF at McMurdo. A 14-day running mean is applied to smooth the ECMWF data (wind rotation angles, 3- and 30-km wind speeds) with a 1-day step, while a 7-point running mean is used to smooth the lidar-measured $\overline{E}_{pm}$ with a 1-point step. $\overline{E}_{pm}$ is multiplied by 20, 1.5, and 6 in (a) to (c), respectively.

Wind rotation angles from lower altitudes to the observed region are taken as a proxy for critical level filtering effects. As pointed out by Yamashita et al. [2009], large wind rotation angles usually indicate that wind profiles cross the zero-wind lines, providing critical levels for orographic gravity waves with nearly zero phase speeds and other low phase-speed waves. Thus, the large wind rotation angles during summer will largely filter out gravity waves that are produced by winds blowing over the trans-Antarctic mountains, islands near the Antarctic continent, or Antarctic Peninsula, or by winds blowing down the Antarctic Ice Sheet. In contrast, the small wind rotation angles in winter indicate that orographic gravity waves and other gravity waves propagating against the tropospheric mean flow (i.e., predominantly westward waves) have much larger probabilities of penetrating through the troposphere and reaching the stratosphere. To quantify this point, we group $\overline{E}_{pm}$ observations with their corresponding wind rotation angles at McMurdo. Probability density functions (PDF) are plotted for wind rotation angles larger than 180° and less than 100° in Figure 3.12a, while the corresponding cumulative distribution functions (CDF) are plotted in Figure 3.12b. Almost all large $\overline{E}_{pm}$ occur during the periods with small wind rotation angles, i.e., in winter. As shown in Figure 3.11, the wind rotation angles in 2012 and 2015 winters are consistently small throughout the winter, in contrast to the other three winters having much more variable wind rotations. Such differences in wind rotation angles are consistent with the observations that $\overline{E}_{pm}$ in 2012 and 2015 winters are higher than those in other three winters.

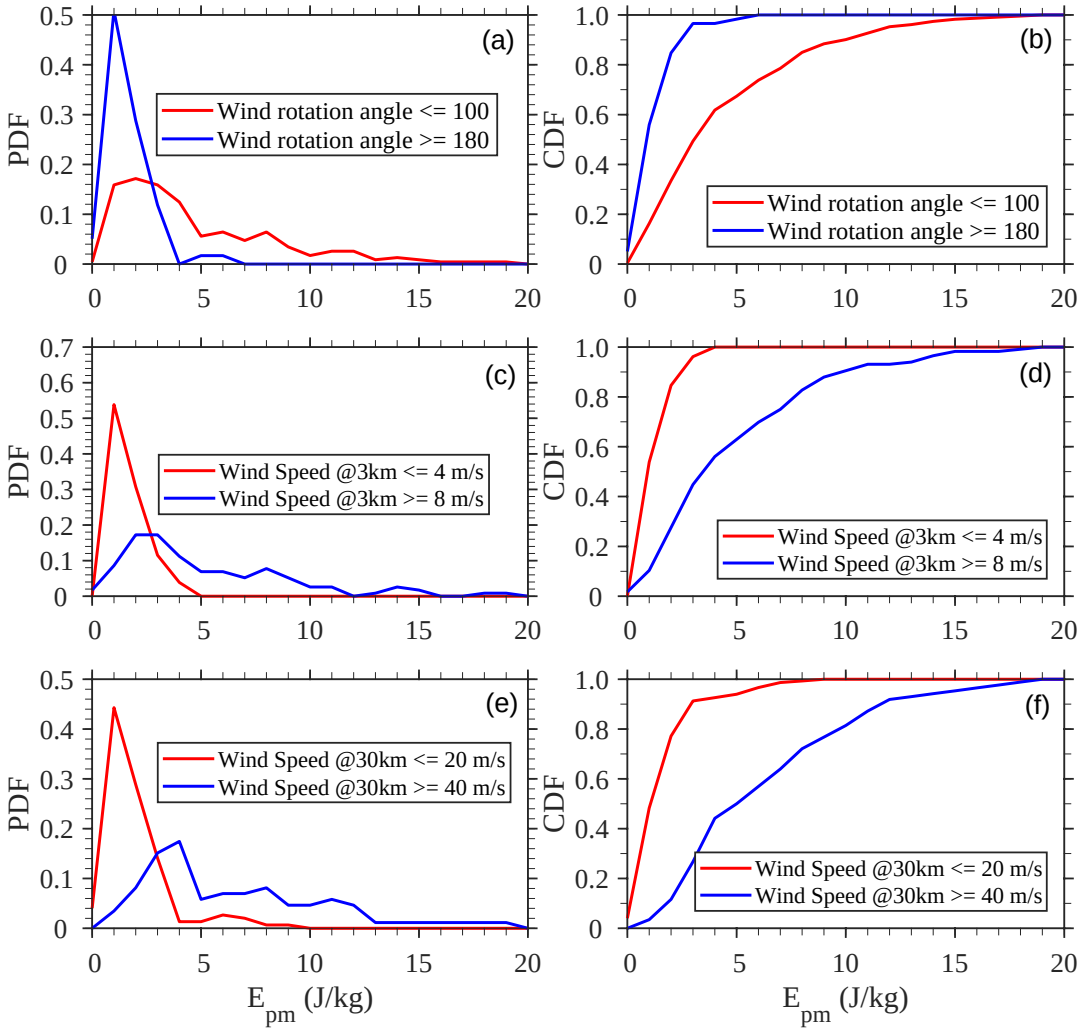

Figure 3.12. (a) Probability density functions (PDF) and (b) cumulative distribution functions (CDF) of $\overline{E}_{pm}$ under different wind rotation angles (i.e., larger than 180° or smaller than 100°). (c) PDF and (d) CDF of $\overline{E}_{pm}$ under different surface wind conditions (i.e., larger than 8 m/s or smaller than 4 m/s). (e) PDF and (f) CDF of $\overline{E}_{pm}$ under different stratospheric wind conditions (i.e., larger than 40 m/s or smaller than 20 m/s). ECMWF wind rotation angles, and 3- and 30-km wind speeds at McMurdo are smoothed with a 7-day running mean with a 1-day step. The lidar-measured raw $\overline{E}_{pm}$ data are not smoothed.

Near-surface wind speed at McMurdo is taken as a proxy of the forcing strength for generation of orographic gravity waves. Faster surface winds in winter than in summer favor the generation of more orographic gravity waves in winter. Hence, the wave sources in winter are expected to be stronger than those in summer. We group the $\overline{E}_{pm}$ observations with corresponding near-surface winds (at 3 km) and show the corresponding PDF and CDF in Figures 3.12c and 3.12d. $\overline{E}_{pm}$ amounts to ~5 J/kg in the slower wind case. Large $\overline{E}_{pm}$ events occur only in the faster wind case, and the positive correlation between $\overline{E}_{pm}$ and the near-surface wind speed is strong. When choosing only the data during small wind rotation angles (<100 deg) and redrawing Figures 3.12c and 3.12d, the positive correlation is even stronger as $\overline{E}_{pm}$ goes up to only ~3 J/kg in the slower wind case and all larger $\overline{E}_{pm}$ occur in the faster wind case (not shown). This result of strong positive correlation is different from that of Yoshiki and Sato [2000] who found a weak correlation between lower stratospheric $\overline{E}_{pm}$ and the surface wind speed at two Antarctic stations: Syowa (69.0°S, 39.6°E) and Casey (66.3°S, 110.5°E). This difference may be caused by differences in the alignment of topography and surface wind direction, as suggested by Watanabe et al. (2006). Westward katabatic winds on the East Antarctic Ice Sheet can excite orographic gravity waves over Syowa and Casey Stations. However, the mean wind in the lower stratosphere is eastward in Antarctic winter and spring, so there is a critical level (zero wind line) for the orographic waves [Yoshiki and Sato, 2000]. Consequently, such orographic gravity waves will not be observed at higher altitudes, losing the correlation between stratospheric gravity wave energy and the surface wind. Contrarily, at McMurdo, most large $\overline{E}_{pm}$ events occurring from May through September are most likely caused by eastward winds blowing down the steep slopes of the East Antarctic Ice Sheet to the west coast of Ross Sea [Watanabe et al., 2006; Vadas and Becker, 2018]. Because 1) critical level filtering is minimal, 2) the enhanced eastward wind can increase the gravity wave vertical wavelength, thereby

allowing orographic gravity waves to propagate upward into the stratosphere before breaking [Vadas and Becker 2018], these effects can lead to the strong positive correlation observed at McMurdo. The McMurdo results along with Syowa and Casey results support the suggestion by Watanabe et al. [2006] and Vadas and Becker [2018] for the roles played by the alignment of the topography relative to the surface wind.

Stratospheric wind speed at 30 km is taken as a proxy of forcing strength for *in-situ* generation of non-orographic gravity waves. Strong winter stratospheric winds, once forming unbalanced flow, can generate gravity waves via geostrophic adjustment [e.g., Zhu and Holton, 1987; Fritts and Luo, 1992; Vadas and Fritts, 2001; Sato and Yoshiki, 2008; Nicolls et al., 2010]. Stratospheric wind speed at 30 km is also a proxy for the Doppler shifting effect. Whiteway et al. [1997] and Duck et al. [2001] have proposed that high stratospheric winds induce Doppler shift, shifting waves towards longer vertical wavelengths. The winter maxima of $\overline{E}_{pm}$ observed at McMurdo are likely related to the increase of vertical wavelength towards winter (Figure 4 in Zhao et al. [2017]), allowing larger wave amplitudes before reaching saturation. This point can be explained by the linear saturation theory—longer λ_z waves can grow to larger wave amplitudes (N^2/m^3) before saturating [Lindzen, 1981; Whiteway et al., 1997], leading to higher $\overline{E}_{pm}$ in winter than in summer. We group $\overline{E}_{pm}$ with their corresponding stratospheric winds at 30 km. PDF and CDF are shown in Figures 3.12e and 3.12f. It is obvious that higher $\overline{E}_{pm}$ correspond to faster stratospheric winds at 30 km. Such strong positive correlation is also true for winds at 40 and 50 km (not shown).

We now discuss the possible causes of the observed seasonal variations of $\overline{E}_{pm}$, λ_z and vertical wavenumber PSD, combining the results from this study with the Part I study by Zhao et al. [2017]. Several monthly-mean wave parameters are plotted in Figure 3.13, complementing the monthly-mean vertical and horizontal wavelengths,

ground-relative and intrinsic periods, vertical phase speed, and horizontal and vertical group velocities given by Figures 4 and 9 in Zhao et al. [2017]. A major conjecture from our observational results is that the large summer-winter asymmetry observed in $\overline{E}_{pm}$ is mainly caused by the critical level filtering of gravity waves by the prevailing wind system (see Figure 4 in Becker [2012]). In our study this mechanism is represented by the wind rotation angle (Figure 3.13b). During summer, most gravity waves are filtered out by the prevailing wind system except the waves with significant eastward phase speeds that become important near the mesopause only. In contrast, under eastward prevailing winds during the wintertime, the majority of gravity waves originating from the lower atmosphere (with westward phase speeds, near zero phase speeds, and low eastward phase speeds) propagate to the observed altitude range (30–50 km), growing to large amplitudes. Orographic events over McMurdo occur when the background wind is large and eastward, thereby causing gravity wave vertical wavelengths to increase with altitude [Vadas and Becker, 2018]. This causes the mountain waves to propagate higher before breaking, thereby leading to larger-amplitude body forces. Larger wave amplitudes and more waves that survive in the winter lead to higher values of $\overline{E}_{pm}$ than in the summer. Because strong critical level filtering in the summer significantly attenuates gravity wave amplitudes before these waves reach 30–50 km, the summer waves in the 30–50 km region experience less severe dissipation than the large amplitude waves in the winter, explaining the H_{pv} seasonal variations in Figure 3.13b. An interesting remark is that the wind rotation angles in 2012 and 2015 winters (Figure 3.11a) appear to be smaller and less variable than those in the winters of 2011, 2013, and 2014. This relation is consistent with the observational results that winter $\overline{E}_{pm}$ in 2012 and 2015 are generally larger than in the other three winters. Hence, critical level filtering likely plays a key role in the observed inter-annual variations of $\overline{E}_{pm}$.

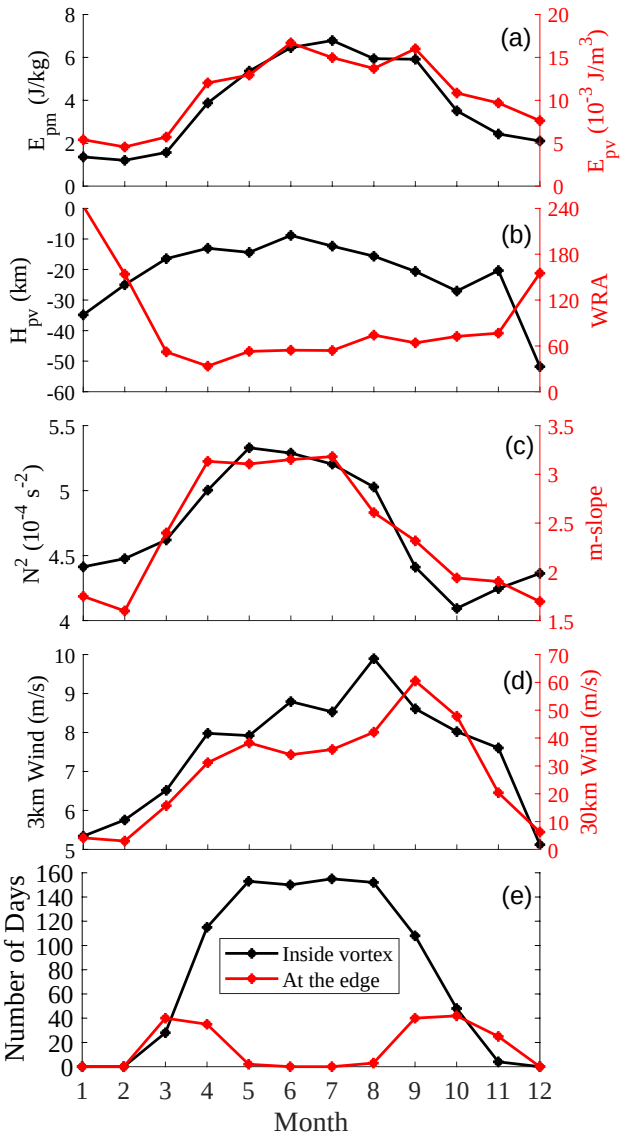

Figure 3.13. Monthly-mean gravity wave parameters and atmospheric conditions at McMurdo. (a) Altitude-mean $\overline{E}_{pm}$ (black) and $\overline{E}_{pv}$ (red), (b) Scale height H_{pv} (black) and wind rotation angle (WRA, red) from 11 to 30 km, (c) monthly-mean $\overline{N^2}$ (black) and slopes (red) of monthly-mean vertical wavenumber spectra in the vertical wavelength range of 2–10 km, (d) ECMWF total wind speeds at 3 km (black) and 30 km (red) altitudes near McMurdo location, and (e) the total number of days during five years of 2011-2015 when McMurdo is inside the polar vortex (<-5 deg, black) and at the vortex edge (within ±5 deg, red).

Another conjecture from the observational results is that the variations of $\overline{E}_{pm}$ from March to October, when the critical level filtering is minimal, are mainly determined by the near-surface winds that are associated with the generation of orographic gravity waves and by the stratospheric winds that are associated with in-situ generation of non-orographic gravity waves (Figure 3.13d) and with Doppler shifting of the vertical wavelengths of gravity waves. The major source of gravity waves in the winter over McMurdo are likely orographic gravity waves produced by downslope eastward winds that blow down from the East Antarctic Ice Sheet to the Ross Sea [Watanabe et al., 2006; Becker and Vadas, 2018; Vadas and Becker, 2018]. The positive correlation between $\overline{E}_{pm}$ and the near-surface wind at 3 km is strong and statistically significant, and so is the positive correlation between $\overline{E}_{pm}$ and the stratospheric wind (see also Figure 3.12). In addition, high stratospheric winds can Doppler shift waves towards longer vertical wavelengths as proposed by Whiteway et al. [1997] and Duck et al. [2001], which allows larger wave amplitudes in winter before reaching saturation as explained above. The increase of the atmospheric stability $\overline{N^2}$ from summer to winter (Figure 3.13c) also allows larger amplitudes of gravity waves during wintertime, but it plays only a secondary role in this context. Moreover, wave dissipation due to the strong stratospheric winds can produce intermittent horizontal body forces that generate secondary waves [Vadas et al., 2003; Becker and Vadas, 2018; Vadas and Becker, 2018]. This process may also change the wave spectra, leading to altered wave saturation. As the dominant waves over McMurdo propagate obliquely with small elevation angles ~1.1^o [Zhao et al., 2017], gravity waves in the McMurdo stratosphere could have originated from orographic wave sources and *in-situ* wave sources (e.g., polar vortex edge) that are far away from McMurdo, explaining the observations of large $\overline{E}_{pm}$ despite McMurdo often being deep inside the polar vortex.

3.8. Summary

With the knowledge of the basic stratospheric gravity wave characteristics obtained in previous chapters, we further investigate the strength (potential energy density) of gravity wave activity at McMurdo and how it varies both throughout a year and over five years from 2011 to 2015. The purpose is to search for potential wave sources of the stratospheric gravity waves.

By developing the spectral proportion method, we conduct a statistical study to characterize E_{pm}, E_{pv}, N^2, and vertical wavenumber spectra in the stratosphere using the Rayleigh temperatures from Fe Boltzmann lidar observations from 2011 to 2015 at McMurdo, Antarctica. This research presents the first lidar observations of repeated seasonal patterns of stratospheric gravity wave potential energy densities and atmospheric static stability N^2 in Antarctica. E_{pm}, E_{pv}, and N^2 profiles in the altitude range of 30–50 km are derived from five years of lidar-measured temperature perturbations and gradients. The scale heights H_{pv} show that wave dissipation in winter is much more severe than in summer. Altitude-mean $\overline{E}_{pm}$ and $\overline{E}_{pv}$ vary significantly from observation to observation, but they exhibit repeated seasonal patterns with winter maximum and summer minimum. The winter maxima in 2012 and 2015 are higher than in other years, exhibiting inter-annual variations. $\overline{E}_{pm}$ and $\overline{E}_{pv}$ obey lognormal distributions and possess narrowly clustered small values in summer, widely spread large values in winter, and intermediate values in spring and fall. Altitude-mean $\overline{N^2}$ also exhibits seasonal variations with maxima in mid-winter but minima around October. The peak-to-peak variation of $\overline{N^2}$ is ~40%, with the mean around 4.7×10^{-4} s^{-2}. The statistical distributions of $\overline{N^2}$ are different from lognormal but nearly bi-modal. The monthly-mean vertical wavenumber power spectral density for vertical wavelengths of 5–20 km increases from summer to winter and decreases back via transition months April and October, respectively. The

seasonal asymmetry in the spectra corresponds to the observed statistically longer vertical wavelengths of gravity waves in winter.

Using MERRA-2 data to determine the polar vortex edge position, we find that largest $\overline{E}_{pm}$ occur in mid-winter when McMurdo is deep inside the polar vortex. We need another lidar site that is situated at the vortex edge in mid-winter to determine whether gravity wave activity peaks at the Antarctic vortex edge as is the case in the Arctic [e.g., Whiteway et al., 1997]. Using ECMWF data to represent the wind fields over McMurdo, we find that large gravity wave $\overline{E}_{pm}$ events occur when the wind rotation angles (representing critical level filtering) are small, and when both the near-surface wind speeds at 3 km (representing the forcing strength of orographic gravity waves) and the stratospheric wind speeds at 30 km (representing the forcing strength of non-orographic gravity waves and Doppler shifting effect) are large. The monthly-mean $\overline{E}_{pm}$ are anti-correlated with ECMWF monthly-mean wind rotation angles while positively correlated with near-surface winds and stratospheric winds with correlations of -0.62, +0.87, and +0.80, respectively, at 95% confidence level. These correlations are strong and statistically significant.

We conclude from these results that the large summer-winter asymmetry observed in $\overline{E}_{pm}$ is mainly caused by the critical level filtering of gravity waves by the prevailing wind system that induces dissipation of most gravity waves in summer around the tropopause and in the lower stratosphere. The variations of $\overline{E}_{pm}$ from March to October, when the critical level filtering is minimal, are mainly determined by the near-surface winds that generate orographic gravity waves and by the stratospheric winds that generate non-orographic gravity waves and Doppler shift gravity waves to longer vertical wavelengths due to the formation of polar vortex. The major source of gravity waves in the winter stratosphere over McMurdo are likely orographic gravity waves produced by downslope winds that travel down the surface slopes of the East Antarctic Ice Sheet. The polar vortex, possessing high-speed

stratospheric winds, has two major effects on gravity waves: it is an important source of non-orographic waves and Doppler shift effects that allow gravity waves to propagate to higher altitudes before reaching wave saturation. The highest PSD in the λ_z range of 5–20 km, the longest vertical wavelengths, and the highest $\overline{E}_{pm}$ in winter are all consistent with each other. Therefore, $\overline{E}_{pm}$ seasonal variations are determined by the combined seasonal changes in critical level filtering from the troposphere to the stratosphere, orographic wave generation near the surface, *in-situ* wave generation in the stratosphere, and modulation of wave saturation by the stratospheric mean winds during wintertime.

Overall, in this chapter, we have discussed the possible sources of the stratospheric gravity waves above McMurdo primarily via characterizing gravity wave potential energy densities and correlating their seasonal variations with various atmospheric parameters. Quantitative determination of contributions from various aforementioned factors requires gravity wave modeling and ray-tracing analysis that are beyond the scope of this work, but it could be the subject of future work. Some remaining questions, for example, include: 1) What do the lognormal distributions of vertical wavelength, ground-relative period, vertical phase speed, and $\overline{E}_{pm}$ imply about wave sources and dissipation? 2) What are the sources of dominant waves in the stratosphere, especially during summer? 3) What causes the inter-annual variations of $\overline{E}_{pm}$? It is worth pointing out that the first part of our statistical study in Chapter 2 focused on dominant gravity waves only, but the potential energy density calculation in this Chapter 3 includes contributions from all gravity waves whose parameters fall within the spectral range determined by the lidar data resolutions. Therefore, this second part of work complements Chapter 2 in terms of both wave activity strength and broader wave spectra. In the future, the combination of wave characteristics, wave energies, and wavenumber and frequency spectra will

help identify sources of persistent gravity waves in the MLT and contribute to studies of wave coupling from the lower to the middle and upper atmosphere. In the next chapter, we explore the possibility of secondary gravity wave generation as the major source for the MLT persistent gravity waves.

CHAPTER IV

SOURCES OF PERSISTENT GRAVITY WAVES: SECONDARY GRAVITY WAVE GENERATION

4.1. **Introduction**

Since the finding that the dominant gravity waves in the stratosphere are not the direct source for the MLT persistent gravity waves in Chapter 2, the search for the source continues. Shibuya et al. [2017] analyzed and simulated the quasi-12 h IGWs in the lower mesosphere observed by the PANSY radar at Syowa Station, and they infer that the IGWs are likely generated by spontaneous emission in the regimes of the midlatitude tropospheric jet and the stratospheric polar night jet. This interpretation was suggested by model simulations showing a predominantly vertical flux of westward momentum in the southern winter mesosphere [Shibuya et al., 2017]. However, the periods of the MLT persistent gravity wave at McMurdo are different from those of IGWs in their case.

Very recently Vadas and Becker [2018] and Becker and Vadas [2018] have interpreted the persistent gravity waves during the Antarctic winter as secondary gravity waves that are generated in the upper stratosphere and lower mesosphere by intermittent and localized horizontal body forces (dubbed "local body forces") that are induced by the dissipation of primary gravity waves. Here the primary gravity waves originate from the lower atmosphere, like orographic gravity waves that account for the required intermittency of the local body forces. This intriguing picture applies in austral winter at the middle and high latitudes. The persistent gravity waves observed in the summer MLT over McMurdo, however, are likely primary non-orographic gravity waves. These waves are probably generated in the troposphere by spontaneous emission or the breakdown of synoptic-scale Rossby waves, and are

filtered by the prevailing wind system in the stratosphere such that only gravity waves with significant eastward phase speeds can propagate into the summer mesosphere (see [Lindzen, 1981; Hoffmann et al., 2010; Becker, 2012]). It remains to be shown whether this speculation for the case of summer is true in GCMs and in observations.

For the Antarctic winter, following results of the theory of secondary gravity wave generation, we have identified two cases of such wave generation from the McMurdo lidar observations which satisfy all 8 criteria needed to conclude that these gravity waves are secondary gravity waves generated from localized body forces. Due to the fact that secondary gravity waves usually have much larger horizontal wavelengths than the corresponding primary gravity waves, this mechanism is likely responsible for generating the large-scale MLT persistent gravity waves with horizontal wavelengths of thousands of kilometers [Vadas and Becker, 2018; Becker and Vadas, 2018]. My contribution to this work was that I have found two cases of lidar observations that contained the secondary gravity wave generation patterns, and analyzed the inherent gravity wave properties in order to determine if they support the theory. This work is published in Vadas, Zhao, et al. [2018] and the following sections within this chapter are adapted from this paper.

4.2. Theory of Secondary Gravity Wave Generation

Vadas et al. [2003, 2018] and Vadas [2013] proposed the secondary gravity wave generation theory. This theory may potentially explain the source of the recently discovered MLT persistent gravity waves at McMurdo. In Vadas et al. [2018], Dr. Vadas derived the compressible, linear solutions that describes the excitation of secondary gravity waves from a local horizontal body force in an isothermal atmosphere with a constant wind. We have briefly reviewed the mathematical formalisms in Chapter 1 (Introduction). The theory proposes that when primary gravity waves such as orographic gravity waves propagate upward, they can dissipate

and deposit their momentum into the background atmosphere if encountering critical levels or convective/shear instabilities. Consequently, local body forces (temporally and spatially localized accelerations of the background mean flow) are created. The acceleration of the background atmosphere generates secondary gravity waves that have horizontal scales much larger than those of the primary gravity waves. It is well known that small-scale secondary gravity waves can be generated directly by gravity wave breaking and nonlinear interactions. However, the larger-scale secondary gravity waves are different from them. The smaller-scale secondary gravity waves usually cannot propagate long before dissipating into the background atmosphere again. In this dissertation, we use the term secondary gravity waves to represent the larger-scale secondary gravity waves generated by local body forces. The large-scale secondary gravity waves can propagate upward and downward, and in and against the body force direction.

Under the assumption that the body force is horizontally-displaced and the gravity wave induced perturbations are scaled by $\sqrt{\rho}$ (square root of the atmospheric density), an important theoretical prediction of the secondary gravity wave generation is the fishbone or ">" structures in altitude-time cross-section as is shown in Figure 4.1a. The "knee" altitude where the phase lines converge (hot phase lines meet cold phase lines, i.e., the phase lines are asymmetric about this altitude) occurs at the body force altitude, which we define as z_{knee}. It is worth to note that such fishbone structures occur only when observing away from the body force center. In the model simulations [Vadas et al., 2018], when the observer is within the force region, there are only a few gravity wave phase lines within the fishbone structure. However, when the observer's location is 2 times the force radius, there are ~ 5 gravity wave phase lines within the structure. When the observer's location is 5.7 times the force radius, the gravity wave phase lines have very small amplitudes close to z_{knee}, resulting in the appearance that the phase lines do not reach z_{knee}. Briefly, to

examine if the perturbation patterns in an altitude-time plot are secondary gravity waves, we need to determine if the gravity wave characteristics such as periods, vertical wavelengths, and density-scaled perturbation amplitudes are roughly the same above and below z_{knee}. In Figure 4.1, the secondary gravity wave generation associated structures in the latitude-longitude cross-section, and altitude-latitude/altitude-longitude cross-section are also illustrated. Secondary gravity wave generation can happen in the stratosphere, the mesosphere, and the thermosphere. As a distinct and general feature, the fishbone structures can be used for the identification of such events throughout the atmosphere.

In Vadas and Becker [2018], high-resolution model data from the Kühlungsborn Mechanistic Circulation Model (KMCM, gravity wave resolved model) was utilized to identify the excitation, propagation, and dissipation of primary and secondary gravity waves during the wintertime above McMurdo. The fishbone structures were examined to discover cases of secondary gravity wave generation. In that paper, via comparing the similarity of gravity waves above and below the knee altitude, they concluded that the gravity waves in the fishbone structure from 30 to 60 km on 9.5 to 10.5 July were secondary gravity waves generated by a local body force. This local body force was ~ 400 km northwest of McMurdo, and the generation happened 2.5 hr earlier than the time these waves were seen above McMurdo. In the previous research in this dissertation, both upward and downward phase progression gravity waves could be detected in the gravity wave field from lidar measurements, which also suggests that secondary gravity wave generation from the primary gravity waves may play a significant role in the gravity wave dynamics in the McMurdo middle and upper atmosphere. In this chapter, we aim to find secondary gravity waves from lidar observations. Indeed, in Vadas et al., [2018], we identified two secondary gravity wave generation cases associated with fishbone structures from

the McMurdo lidar campaign. The analysis of these two cases are introduced in the following section.

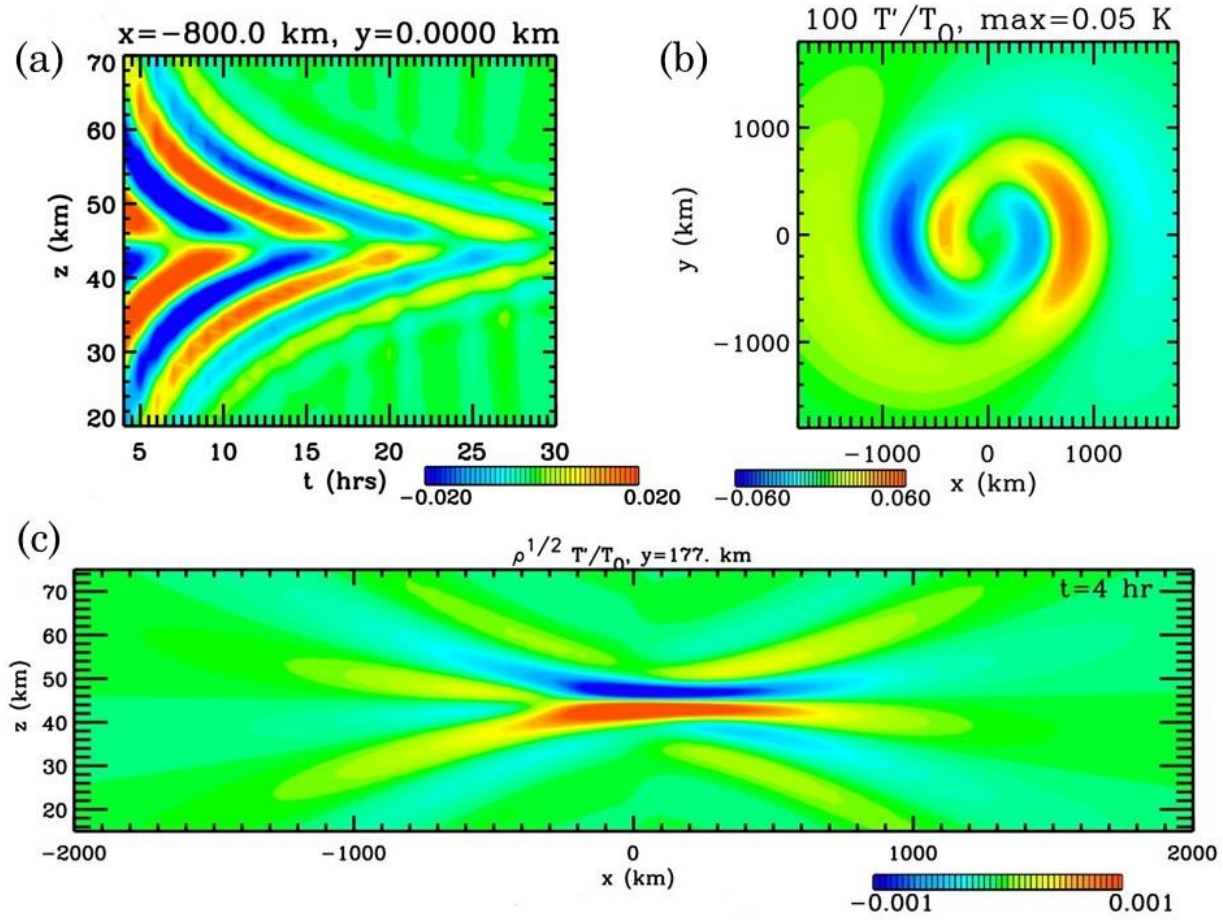

Figure 4.1. (a) Secondary gravity wave generation induced fishbone structure in altitude-time cross-section. (b) Secondary gravity wave generation induced structure in latitude-longitude cross-section. (c) Secondary gravity wave generation induced structure in altitude-latitude/altitude-longitude cross-section. [Vadas et al., 2018].

4.3. Two Cases of Secondary Gravity Wave Generation in McMurdo Fe Lidar Data

Here we enumerate the two lidar observations that contain the fishbone structures associated with secondary gravity wave generation. The lidar temperature data was collected with the Fe Boltzmann lidar at Arrival Heights nearby McMurdo,

Antarctica. The temperature data was retrieved with 1 hr temporal resolution and 1 km altitude resolution via Rayleigh integration technique from 30 to 70 km. From the temperature field (versus time and altitude), we extract the gravity wave induced perturbations via background subtraction and wave filtering process. The background subtraction is to subtract a temporal mean across the observation at every altitude. The wave filtering process refers to at every altitude, we employed 1D FFT, 1DIFFT, and 6th order Butterworth filter to extract gravity waves with periods shorter than 11 hr. We then scale the extracted gravity wave induced temperature perturbations with the square root of the atmospheric density to cancel out the exponential increase of gravity wave amplitudes with respect to altitude. In the scaled temperature perturbations in the following two cases, clear fishbone structures can be identified in the altitude-time cross-section. We will show that the gravity wave parameters such as ground-relative periods, and vertical wavelength are roughly the same below and above z_{knee}, which indicates possible secondary gravity wave generations.

Here we summarized the analysis procedures. The two assumptions we make in advance are that 1) the gravity wave in the fishbone structures are secondary gravity waves, and 2) downward (upward) phase progression corresponds to upward (downward) propagating gravity waves. We will validate the two assumptions at the end of the analysis process. As stated previously, we calculate the scaled gravity wave temperature perturbations $\sqrt{\rho}T'/T$ for gravity waves with periods shorter than 11 hr. We estimate the knee altitude z_{knee} requiring that [Vadas, Zhao et al., 2018]:

1) The structure is asymmetric in z about z_{knee}, i.e., the cold and hot phase lines converge at z_{knee}. An incorrect estimate for z_{knee} (whereby the structure is symmetric in z) yields an incorrect vertical range for the calculated spectra below and above z_{knee}, which results in incorrectly-determined (biased) gravity wave parameters below and above z_{knee}. We then remove all upward (downward)-propagating gravity waves below

(above) z_{knee} to isolate the fishbone structure. We identify by eye the temporal and vertical extent for the structure. Then, we require that the following criteria are met:

2) If upward-propagating gravity waves are present below z_{knee}, they are partially or fully dissipated at least a few km below z_{knee}. This allows for a possible excitation mechanism for the secondary gravity waves; i.e., that the primary gravity waves dissipate and create a body force. However, the center of the body force would need to be horizontally-displaced in order to see the secondary gravity waves;

3) If upward-propagating gravity waves are present below z_{knee}, $|\lambda z|$ does not become extremely large near z_{knee}. This criterion rules out the possibility that the primary gravity waves reflect downward at z_{knee}, which could be mistaken for downward-propagating secondary gravity waves;

4) If downward-propagating gravity waves are present above z_{knee}, they only have small scaled amplitudes relative to the scaled amplitudes of the downward propagating gravity waves below z_{knee}. This criterion helps eliminate overly-complicated cases. We then calculate the spectra below and above z_{knee} separately for the secondary and removed gravity waves, and determine the peak values of period and vertical wavelength. We require that:

5) The peak values of $|\lambda z|$ and τ for the removed gravity waves below z_{knee} are different than that for the secondary gravity waves above z_{knee}. This ensures that the upward-propagating secondary gravity waves are not continuations of the upward-propagating primary gravity waves.

6) The peak values of $|\lambda z|$ and τ for the removed gravity waves above z_{knee} are different than that for the secondary gravity waves below z_{knee}. This ensures that the downward-propagating secondary gravity waves are not continuations of the downward-propagating gravity waves above z_{knee}. We then check the validity of our first assumption, i.e., that the gravity waves in the fishbone structure are secondary gravity waves. Since secondary gravity waves in an unsheared, isothermal

atmosphere have the same τ, $|\lambda z|$ and scaled amplitudes below and above z_{knee}, we require that:

7) The parameters τ and $|\lambda z|$ are similar below and above z_{knee}, and the scaled amplitudes are within a factor of 2-2.5 below and above z_{knee}. (Here, we allow for a significant difference of the gravity wave amplitudes because even small shears can dissipate a large portion of the secondary gravity wave spectrum if it peaks at small to medium intrinsic horizontal phase speeds.)

Finally, we check the validity of our last assumption, i.e., that the GWs in the fishbone structure having upward (downward) phase progression are downward (upward) propagating gravity waves. We do this via requiring that:

8) In the vicinity of z_{knee}, the intrinsic horizontal phase speed c_{IH} is greater than zonal wind $|U|$ and meridional wind $|V|$ (see explanation below). Note that this is an overly conservative estimate if the gravity wave primarily propagates meridionally, because $|U|$ is often much larger than $|V|$. If the GW propagation direction is known, we would instead compare c_{IH} directly with U_H.

If a case satisfies all criteria from 1) to 8), a fishbone structure is regarded as the signature of secondary gravity wave generation that is induced by a horizontally-displaced body force at the knee altitude.

1) Secondary gravity wave generation on 18 June 2014

In order to derive the fishbone structures from the lidar temperature, we first extract wave (planetary wave, tides, gravity wave, etc.) induced temperature perturbations by subtracting a temporal background. Then such perturbations are scaled with atmospheric density to obtain the density-scaled temperature perturbations as is shown in Figure 4.2a,

$$\sqrt{\rho}\,T'/T_{Bkg} = \sqrt{\rho}\,(T - T_{Bkg})/T_{Bkg} \qquad (4.1)$$

where T is the lidar observed temperature versus altitude and time, T' is the wave induced temperature perturbations, and T_{Bkg} is the background temperature averaged across the observational period at each altitude. ρ is the atmospheric density (with the unit of kg/m^3). Here we adopted the monthly averaged (June in this case) atmospheric density profile from NRLMSISE-00 [Picone et al., 2002]. The scaled temperature perturbations are largely dominated by planetary waves with periods of days at McMurdo [Lu et al., 2013]. We apply a 6th order high-pass filter in the temporal domain to extract gravity waves with periods shorter than 11 hr to Figure 4.2a. The derived gravity-wave-induced scaled temperature perturbations are shown in Figure 4.2b. At z < 45 km, constructive and destructive interference patterns are seen for upward and downward-propagating gravity waves. We immediately see that almost all the gravity waves are upward propagating at altitudes higher than 45 km across the whole observation. In seeking for the fishbone structure, we judge that from 5 to 30 UT, upward-phase-progression gravity waves are present at altitudes from 30 to 42 km, downward-phase-progression gravity waves with similar period and vertical wavelengths are present at altitudes from 45 to 60 km. Hence, gravity waves contained in this area are speculated to be part of a fishbone structure. We estimate the knee altitude at ~ 43 km via visual judgement following criterion 1).

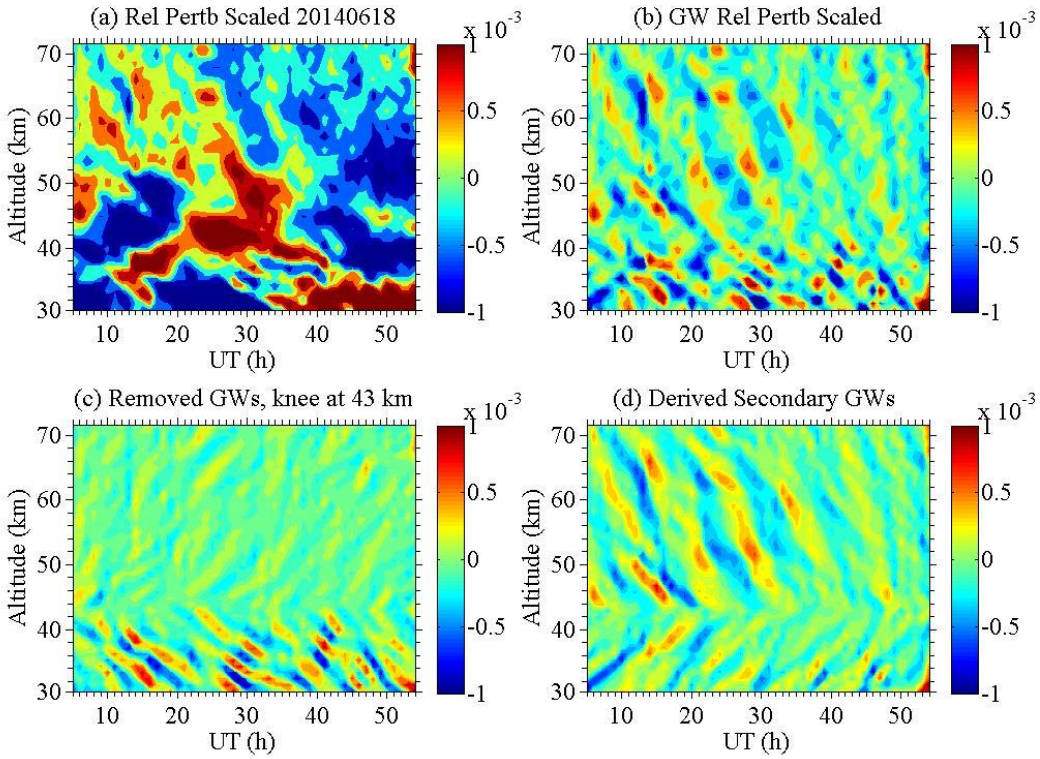

Figure 4.2. (a) Atmospheric density scaled temperature perturbations (including planetary waves, tides, and gravity waves, etc.), $\sqrt{\rho}T'/T$, on 18 June 2014. (b) As in (a), but only retaining gravity waves with periods shorter than 11 hr. (c) Removed gravity waves from (b), obtained by selecting gravity waves with upward phase progression for $z > 43$ km and downward phase progression for $z < 43$ km. (d) Derived secondary gravity waves, obtained by subtracting (c) from (b). Color bars are in units of $\sqrt{kg/m^3}$.

We continue to discuss whether the gravity waves in such fishbone structures are the indications of the secondary gravity wave generation around 43 km. Different Fast Fourier Transform (FFT) filters are applied to the altitude ranges above and below knee individually. We excluded gravity waves with downward phase progression for altitudes below z_{knee} and gravity waves with upward phase progression for altitudes above z_{knee}. Such removed gravity waves are shown in Figure 4.2c. In this panel, at altitudes lower than the knee altitude, large amplitude downward phase progression gravity waves are prominent, which are likely upward

propagating gravity waves from lower altitudes such as mountain waves from the troposphere or *in-situ* generated gravity waves from unbalanced flow in the lower stratosphere. These upward propagating gravity waves encounter severe dissipation around 35 to 40 km which satisfies criterion 2). The hot and cold phase lines also do not become vertical when approaching the knee altitude such as satisfying criterion 3). We also notice that the upward-phase-progression gravity waves above the knee altitude have small-amplitudes, thereby satisfying the requirements of criterion 4). The obtained secondary gravity wave patterns are shown in Figure 4.2d, which is derived by subtracting Figure 4.2c from Figure 4.2b. We immediately see visible fishbone structures from 0 to 30 UT and 30 to 60 km that might be associated with secondary gravity wave generation. Note that the scaled gravity wave amplitudes are smaller below the knee altitude after 20 UT. In order to investigate if such fishbone structures are associated with secondary gravity wave generation, the wave parameters of the secondary and removed gravity waves need to be characterized. Via visual judgements, we locate the fishbone structure extension to be from 5 to 26 UT and from 35 to 50 km. Two dimensional FFT (2DFFT) was taken on the determined fishbone structures (density-scaled gravity-wave-induced temperature perturbations) for the removed and secondary gravity waves below and above the knee altitude separately. We denote the output of the 2DFFT as $\sqrt{\rho}T'/T$. No window was applied to the fishbone structure before performing the transform. Power spectral density (PSD) of the removed and secondary gravity waves is calculated via $\left(\sqrt{\rho}T'/T\right)\left(\sqrt{\rho}T'/T\right)^*$, where "*" represents the complex conjugate. We present the PSD results in Figure 4.3. For the secondary gravity waves, a single dominant peak occurs in each PSD. We performed Monte Carlo simulations with 500 iterations to determine the parameters of the dominant waves and their error bars in the secondary and removed gravity waves. For each simulation, the temperature field

versus altitude and time is constructed according to the lidar observed temperature and the associated temperature errors. The constructed temperature at every grid point consists of the sum of the lidar observed temperature at this grid point and a deviation. This deviation is generated by drawing a random number from a Gaussian distribution with a mean of 0 and a standard deviation equal to the lidar observed temperature uncertainty at this grid point. We then separate the secondary and removed gravity waves, calculate the PSDs below and above the knee altitude on this simulated temperature field as described previously. After 500 simulations, the peak periods and vertical wavelengths are then calculated by taking their PSD weighted average. The error in the temperature uncertainty is calculated by taking the PSD weighted standard deviations for the 500 simulations. The final error bar for each gravity wave parameter includes this Monte Carlo temperature uncertainty error, the temporal or vertical binning resolution error, and the FFT resolution error via taking the square root of their squared sum. Table 4.1 enumerates the results for the dominant wave parameters in the secondary and removed gravity waves for this case, which are used to justify that this fishbone structure is indeed induced by secondary gravity wave generation. Above the knee altitude, the secondary gravity waves have periods of 8.26 ± 0.52 h and vertical wavelengths of 13.62 ± 2.20 km. In contrast, below the knee altitude, the removed gravity waves have periods of 8.09 ± 0.53 h and vertical wavelengths of 4.67 ± 0.52 km. Due to the large difference of the vertical wavelength between the secondary gravity waves and the removed gravity waves, we argue that the upward-propagating secondary gravity waves above the knee altitude are not the continuations of the upward-propagating gravity waves below the knee altitude, which satisfies criterion 5). In addition, we also notice that the secondary gravity waves below the knee altitude have periods of 9.54 ± 0.57 h and vertical wavelengths of 13.55 ± 1.22 km. However, the removed gravity waves above the knee altitude have periods of 6.82 ± 0.53 h and vertical wavelengths of 3.98 ± 0.54 km.

Since the vertical wavelengths still differ by a huge amount, we deduce that the downward-propagating secondary gravity waves below knee are not the continuations of the downward-propagating removed gravity waves above knee. Hence, criterion 6) is met. In the case of the derived secondary gravity waves, the periods and vertical wavelengths below and above the knee altitude are rather similar. Importantly, the density scaled gravity wave amplitudes below and above the knee altitude are (0.25 – 0.6) and (0.25 – 1.0) $\sqrt{kg/m^3}$, respectively. The variations in the scaled amplitudes are within a factor of 2 to 2.5, which satisfies criterion 7).

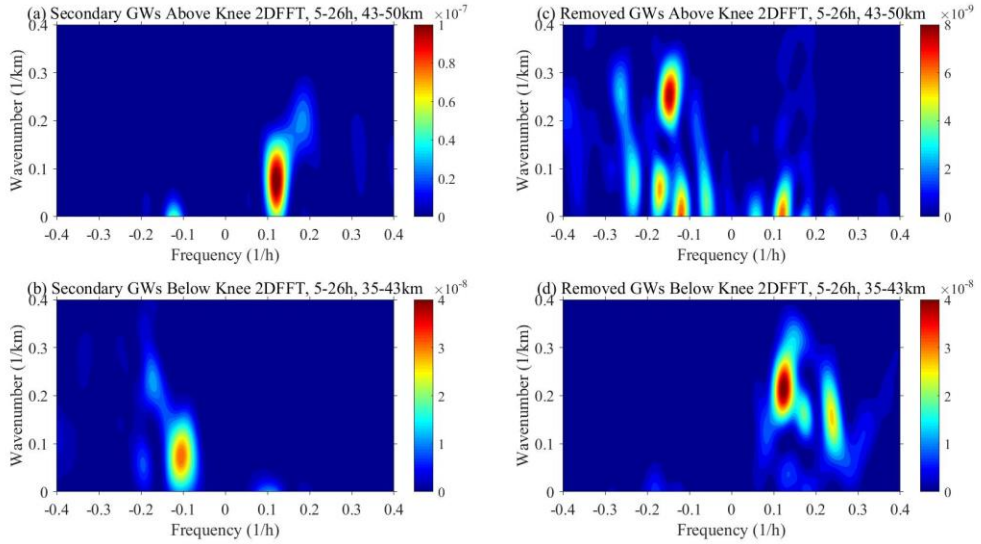

Figure 4.3. (a-b) PSD of $\sqrt{\rho}\,T'/T$ for the derived secondary gravity waves from Figure 4.2 for 5 to 26 UT as a function of wavenumber and frequency: (a) Above the knee using data for z = 43 to 50 km, (b) Below the knee using data for z = 35 to 43 km, Negative (positive) frequency denotes upward (downward) phase progression. (c-d) Same as (a-b), but for the removed gravity waves in Figure 4.2.

Table 4.1. Parameters of the GWs on 18 June 2014

Date	Below Knee		Above Knee	
Secondary GWs	τ (h)	|λz| (km)	τ (h)	|λz| (km)
	9.54±0.57	13.55±1.22	8.26±0.52	13.62±2.20
Removed GWs	8.09±0.53	4.67±0.52	6.82±0.53	3.98±0.54

At this step, we validate the assumption that the secondary gravity waves with upward (downward) phase progression corresponds to downward (upward) energy propagation. An upward-propagating gravity wave has upward phase lines in altitude-time cross-section if it is propagating against background wind with $U_H < 0$ and $|U_H| \geq c_{IH}$ (i.e., the gravity wave propagates against the background wind, and it is swept downstream in the direction of the wind) [Fritts and Alexander, 2003; Dornbrack et al., 2017]. The opposite is true for a downward-propagating gravity wave. The equation reveals such a scenario is,

$$c_{IH} = c_H - U_H \tag{4.2}$$

If $c_H < 0$, then upward-propagating (downward-propagating) gravity waves have upward (downward) phase progression (Stationary mountain waves have horizontal phase speed of 0). We followed the convention that $k_H \geq 0$ and $c_{IH} \geq 0$ by definition (otherwise the gravity wave would have encountered a critical level), so $c_H < 0$ if $U_H < 0$ and $|U_H| > c_{IH}$. In reality, such situations happen when the gravity wave propagate against the background winds and the background winds accelerate rapidly. Hence the gravity wave is swept downstream. An analogy to such a phenomenon is if a swimmer is swimming upstream in a river where the downstream flow accelerates significantly. Even though the swimmer is swimming upstream against the flow, he is actually swept downstream relative to the bank [Vadas and Becker, 2018].

Unfortunately, the wind observation that overlaps with the lidar observation is not available. Hence, MERRA-2 (Modern-Era Retrospective analysis for Research

and Applications - 2) data are utilized to provide the zonal wind (U) and meridional wind (V) information to confirm the presumed criterion 8). Figure 4.4(a-b) presents the zonal and meridional winds during the lidar observations at McMurdo. We immediately see that above the knee altitude, southeastward wind exists within the fishbone structure with an amplitude of 20 to 70 m/s. Whereas below the knee altitude, from 5 to 12 UT and from 20 to 26 UT, the wind is southeastward, and from 12 to 20 UT, the wind is northeastward with an amplitude of 10 to 40 m/s.

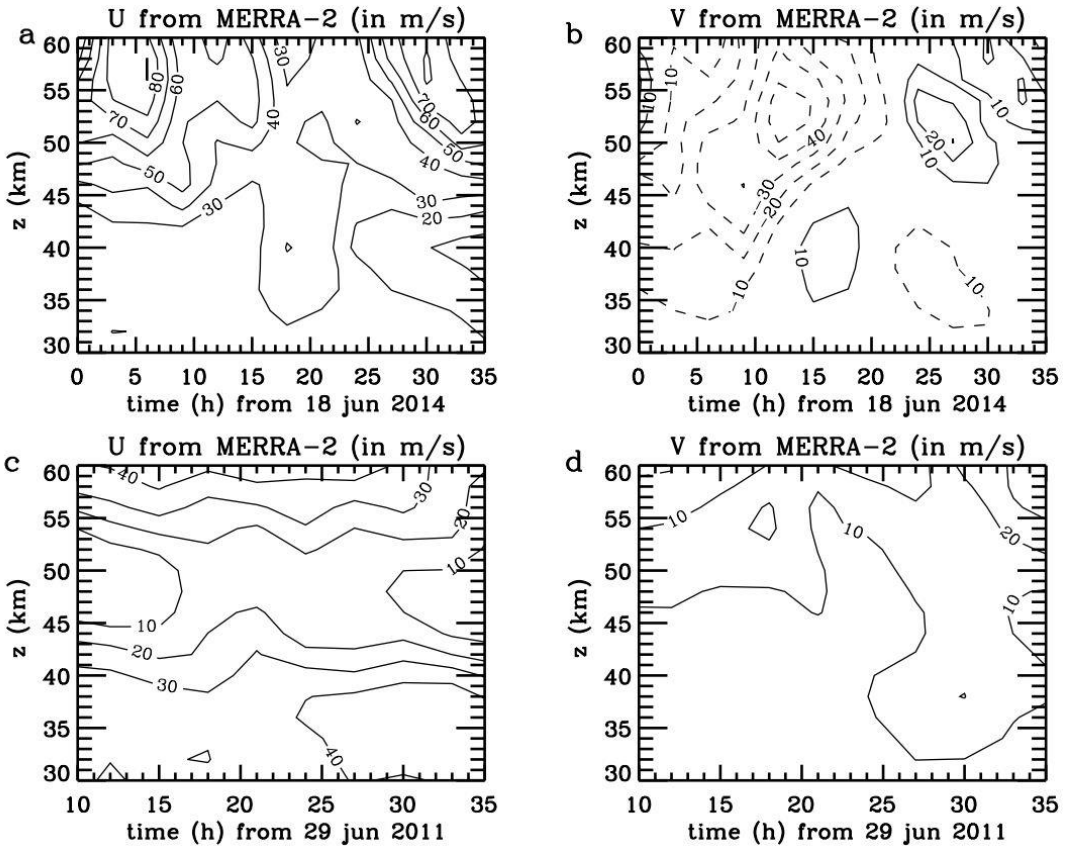

Figure 4.4. Background wind from MERRA-2 at McMurdo. (a) U and (b) V on 18 June 2014. (c-d): Same as (a-b) but on 29 June 2011. Solid (dash) lines represent positive (negative) values.

The intrinsic horizontal phase speed of the secondary gravity wave needs to be extracted from the observation. According to the gravity wave dispersion relation for medium-frequency gravity waves, the gravity wave's intrinsic phase speed is,

$$c_{IH} = \omega_I / k_H = N_B / m = |\lambda_z| / \tau_B \qquad (4.3)$$

where $\tau_B = 2\pi / N_B \simeq 5.0$ min represents the buoyancy period according to MERRA-2 for the fishbone structure extension. Based on the information of the vertical wavelength from Table 4.1, we obtain $c_{IH} = 45$ m/s for the derived secondary gravity waves. As previously stated, to confirm the presumed criterion 8), we have to make sure within the structure extension, $c_{IH} > \sqrt{U^2 + V^2}$. In this case, from Figure 4.4(a-b), the above condition satisfies below the knee altitude. Hence, the secondary gravity waves below the knee altitude with upward phase progression are actually downward-propagating gravity waves. However, above the knee altitude, the situation is more complicated. Generally, the above condition satisfies. However, $c_{IH} > \sqrt{U^2 + V^2}$ except for altitudes from 46 to 50 km and from 5 to 11 UT if the gravity waves have significant eastward propagation component. In contrast, if the secondary gravity waves propagate mainly meridionally, they would be upward propagating and would have downward phase progression within the entire structure. Therefore, we confirmed that the secondary gravity waves were upward propagating from 43 to 46 km at 5 to 26 UT. We notice that the phase lines in Figure 4.2d do not change slope significantly around 46 km (they would change slope dramatically due to the significantly changed vertical wavelengths if they were propagating zonally and came across the strong eastward wind shear as indicated in Figure 4.4a from 5 to 10 UT). Therefore, we conjecture that the upward-propagating secondary gravity waves from 5 to 26 UT continue to propagate upward from 46 to 50 km, and that they propagate significantly in the meridional direction. Hence, criterion 8) is satisfied, that is, the secondary gravity waves above the knee altitude are upward-propagating gravity waves with downward phase progression, and the secondary gravity waves below the

knee altitude are downward-propagating gravity waves with upward phase progression. Since we have validated all 8 criteria that the secondary gravity wave generation theory requires, we conclude that the gravity waves in the derived fishbone structure on 18 June 2014 are secondary gravity waves generated from a horizontally-displaced local body force.

2) Secondary gravity wave generation on 29 June 2011

Figure 4.5a presents the atmospheric-density-scaled temperature perturbations induced by waves (including planetary, tidal, and gravity waves, etc.) on 29 June 2011. In Figure 4.5b, we present the scaled gravity-wave-induced temperature perturbations after the 6^{th} order high-pass Butterworth filter with cutoff frequency at 1/11 (h^{-1}). A possible fishbone structure can be seen with the knee altitude at 52 km from 10 to 25 UT and from 45 to 65 km. Based on criterion 1), we determine that the knee altitude is at 52 km. The removed gravity waves are shown in Figure 4.5c. The gravity waves below the knee altitude propagate upwards with large amplitudes and encounter severe wave dissipation around 43 to 45 km, which satisfies criterion 2). The gravity wave vertical wavelength does not become extremely large around the knee altitude which satisfies criterion 3). Criterion 4) requires that the gravity waves have small amplitudes above the knee altitude which is the case in Figure 4.5c. The derived secondary gravity waves are shown in Figure 4.5d. We can easily outline the fishbone structure with similar periods, vertical wavelengths, and scaled wave amplitudes below and above the knee altitude.

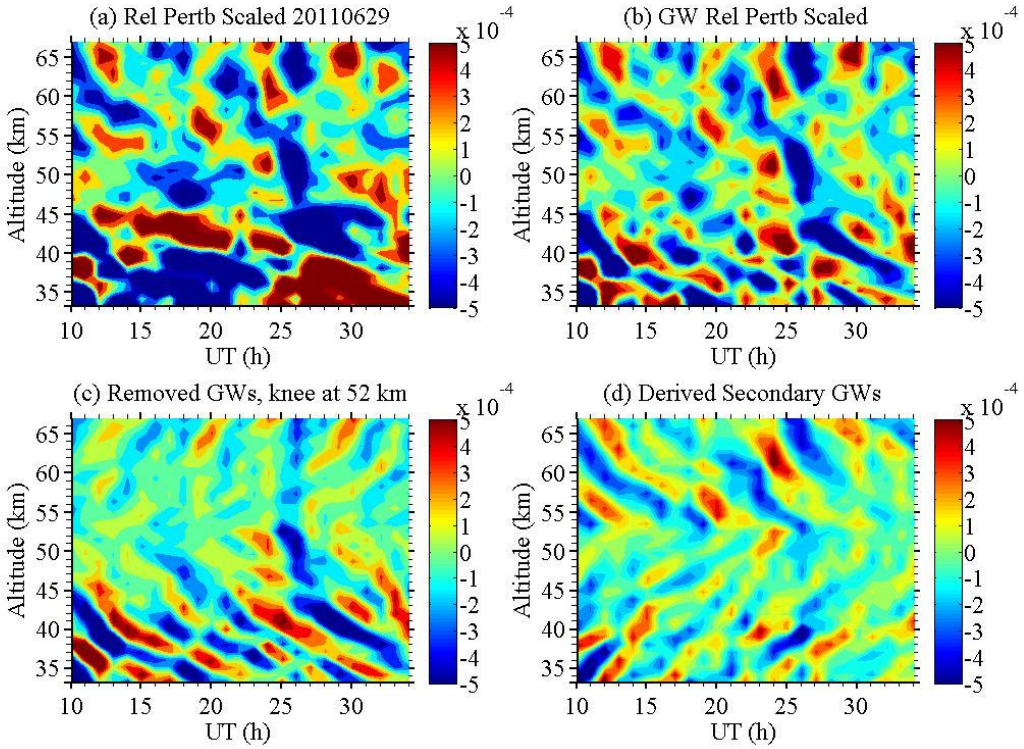

Figure 4.5. Same as Figure 4.2, but for 29 June 2011 with z_{knee} = 52 km.

We decide the fishbone structure extent to be from 10 to 22 UT in the time domain and from 45 to 60 km in the altitude domain to characterize the gravity wave parameters for the secondary gravity waves and the removed gravity waves. The corresponding PSDs for the secondary and removed gravity waves below and above the knee altitude are plotted in Figure 4.6. Single peaks occur in Figure 4.6(a-c). In Figure 4.6d, a dominant peak (gravity wave #1) and a secondary peak (gravity wave #2) show up, indicating there are two upward-propagating gravity waves from altitudes below the knee. The associated gravity wave parameters are summarized in Table 4.2. The secondary gravity waves above the knee altitude have periods of 7.96 ± 0.63 h and vertical wavelengths of 8.10 ± 1.04 km. However, below the knee, the removed gravity wave #1 has period of 4.89 ± 0.54 h and vertical wavelength of 9.93 ± 1.52 km, the removed gravity wave #2 has period of 10.07 ± 0.79 h and vertical wavelength of 38.91 ± 17.08 km. Owing to the differences of the gravity wave parameters between the secondary gravity waves and the removed gravity wave #1

and #2, the case satisfies criterion 5). Additionally, the secondary gravity waves below the knee altitude have periods of 6.10 ± 0.64 h and vertical wavelengths of 6.28 ± 0.83 km. In contrast, the removed gravity waves above the knee altitude have periods of 3.44 ± 0.52 h and vertical wavelengths of 20.27 ± 6.99 km. Due to no similarity is found between these gravity wave parameters, the case satisfies criterion 6). Furthermore, the peak periods and vertical wavelengths below and above the knee altitude are quite similar for the derived secondary gravity waves. We also notice that the scaled gravity wave amplitudes below and above the knee altitude are (1.2 to 2.0) and (1.0 to 4.0) $\sqrt{kg/m^3}$ in Figure 4.6d, whose variations are within a factor of 2 of each other. Hence, the case satisfies criterion 7).

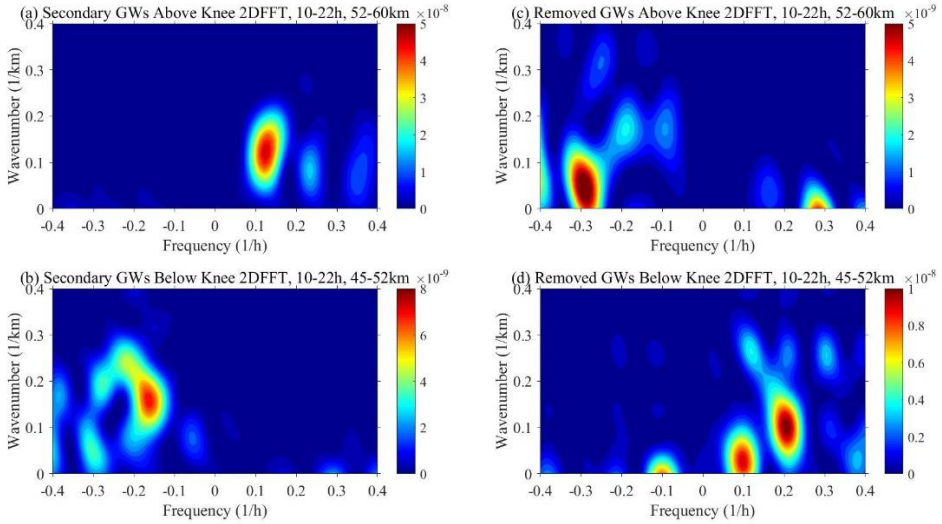

Figure 4.6. Same as Figure 4.3, but for June 29, 2011 at 10-22 UT using data for z = 52 to 60 km above the knee and for z = 45 to 52 km below the knee.

Table 4.2. Parameters of the GWs on 29 June 2011

Date	Below Knee		Above Knee	
Secondary GWs	τ (h)	|λz| (km)	τ (h)	|λz| (km)
	6.10±0.64	6.28±0.83	7.96±0.63	8.10±1.04
Removed GWs #1	4.89±0.54	9.93±1.52	3.44±0.52	20.27±6.99
Removed GWs #2	10.07±0.79	38.91±17.08		

Finally, we validate the presumed criterion 8). According to the secondary gravity wave parameters in Table 4.2, the intrinsic horizontal phase speed is calculated to be c_{IH} = 21 and 27m/s below and above the knee altitude via equation (4.3). We utilize zonal and meridional winds from MERRA-2 to do the validation as is plotted in Figure 4.4(c-d). Across the whole fishbone structure extension, the wind is northeastward. Below the knee altitude, U and V are both less than 21. Above the knee altitude, V is less than 20 m/s. But U is less than 27 m/s only from 52 to 55 km, and is larger than 27 m/s above 55 km. However, due to the slope of the gravity wave phase lines do not change significantly around 55 km in Figure 4.5d (which would occur due to the significantly changed vertical wavelengths if the upward-propagating secondary gravity waves propagate zonally), we regard that the upward-propagating secondary gravity waves continue to propagate upward after 55km, and they should have a significant meridionally-propagating component. By this step, we confirmed the presumed criterion 8), i.e., in the fishbone structure, the gravity waves with upward phase progression below the knee altitude are downward propagating, and the gravity waves with downward phase progression above the knee altitude are upward propagating. Overall, the fishbone structure on 29 June 2011 are likely formed due to the excited secondary gravity waves from a horizontally-displaced local body force.

4.4. Summary

To answer the question of the source of the mysterious MLT persistent gravity waves, we utilized the secondary gravity wave generation theory [Vadas et al., 2003]. Two case studies (18 June 2014 and 29 June 2011) utilizing temperature data from McMurdo lidar campaign are performed to provide the observational basis for such a theory. Clear fishbone structures in the gravity wave induced relative temperature perturbations are seen, which indicate the excitation of secondary gravity waves.

To extract the fishbone structures in the temperature relative perturbations (scaled by the square root of the atmospheric density), we first excluded all the waves with periods longer than 11 hr. Then utilizing selective Fast Fourier Filtering above and below the knee altitude, the fishbone structures can be retrieved. We select the knee altitude, verify the upward-propagating gravity-wave-dissipation around the knee altitude, rule out the possibility of gravity wave reflection around the knee altitude, confirm weak downward-propagating gravity waves above the knee altitude to satisfy the criteria for a secondary gravity wave generation case. We calculate PSDs for the removed and secondary gravity waves in the body force extension region via 2DFFT and compared the similarity of (or lack of similarity of) the gravity wave vertical wavelengths and ground relative periods above and below the knee altitude. Combined with the MERRA-2 winds, we confirm that the upward- (downward-) phase-progression gravity waves are indeed downward- (upward-) propagating gravity waves. By fitting the fishbone structures into the 8 criteria that secondary gravity wave generation theory requires, we conclude that the derived fishbone structure in the gravity-wave-induced temperature perturbations are indeed from secondary gravity waves excited from horizontally-displaced local body forces.

Overall, based on McMurdo lidar temperature measurements, this study presents the first observational evidence for the secondary gravity wave generation theory. It is exciting that we observed such events in reality that upon the break of the primary gravity waves, the body force created by the deposited momentum can excite secondary gravity waves with larger horizontal wavelengths in the upper stratosphere, which can freely propagate well into the mesosphere due to their large spatial scales. As written in Vadas et al. [2018], "this novel picture, that primary gravity waves propagate upward and dissipate, which excites secondary gravity waves that propagate upward, has opened a new door in aeronomy which involves complex intertangled coupling processes from the lower atmosphere to the upper

thermosphere". Since the excited secondary gravity waves have much larger horizontal scales compared to the corresponding primary gravity waves, we conclude that the generation of secondary or higher-order gravity waves in the upper stratosphere or in the lower mesosphere can provide the origin for the discovered large-scale MLT persistent gravity waves with horizontal wavelengths of thousands of kilometers.

By this far, for the gravity waves in the stratosphere and in the MLT above McMurdo, we have 1) investigated the possible sources for the stratospheric gravity waves, 2) ruled out the possibility that the MLT gravity waves come directly from the dominant stratospheric gravity waves, and 3) concluded that the secondary gravity wave generation provides the possible sources for the MLT persistent gravity waves. A much clearer vertical picture of gravity waves above McMurdo, Antarctica is formed.

CHAPTER V

SPECTRAL PROPORTION METHOD AND ITS APPLICATIONS TO THE ACCURATE ESTIMATION OF WAVE ENERGY

5.1. Introduction

In this chapter, we introduce the spectral proportion method that employed in Chapter 3 from the perspective of rigorous mathematical formalisms and forward modeling validations. The spectral proportion method can be utilized to calculate various gravity wave parameters, such as kinetic energy per unit mass $E_{km,GW}$, potential energy per unit mass $E_{pm,GW}$, which is essential for physically resolving gravity waves in GCMs and CCMs [Kim et al., 2003; McLandress et al.,2006; Richter et al., 2010; Becker, 2012]. $E_{km,GW}$ and $E_{pm,GW}$ have been regarded as indicators of the strength of gravity wave activity levels for decades [Wilson et al., 1991; Whiteway et al., 1997; Yoshiki and Sato, 2000; Duck et al., 2001; Yamashita et al., 2009; Alexander et al., 2011; Chu et al., 2018]. Statistical and case studies have been performed on $E_{km,GW}$ from wind observations and $E_{pm,GW}$ from temperature or atmospheric density measurements. The calculations for both $E_{km,GW}$ and $E_{pm,GW}$ mainly involve solving for the mean square of gravity-wave-induced perturbations [e.g., Tsuda et al., 2000; Yoshiki and Sato, 2000; Wilson et al., 1991; Yamashita et al., 2009]. That is,

$$E_{km,GW}(z) = \frac{1}{2}(\overline{u_{GW}^{'2}(z,t)} + \overline{v_{GW}^{'2}(z,t)}) \tag{5.1}$$

$$E_{pm,GW}(z) = \frac{1}{2}\frac{g^2}{\overline{N^2(z)}}\overline{\left(\frac{T_{GW}^{'}(z,t)}{T_{Bkg}(z)}\right)^2} = \frac{1}{2}\frac{g^2}{\overline{N^2(z)}}\frac{1}{N_p}\sum_{i=1}^{N_p}\left(\frac{T_{GW}^{'}(z,t_i)}{T_{Bkg}(z)}\right)^2 \tag{5.2}$$

where $u_{GW}^{'}(z,t)$, $v_{GW}^{'}(z,t)$, and $T_{GW}^{'}(z,t)$ are perturbations in the zonal wind (u), meridional wind (v), and temperature (T) induced purely by gravity waves, z and t represent altitude and time, respectively, N_p is the number of temperature perturbation profiles within the observation, g is the gravitational acceleration, T_{Bkg}

is the background temperature, $N(z)$ is the buoyancy frequency. The overbars represent temporal means.

Note that the gravity-wave-induced perturbations are always accompanied by noise because the wave perturbations are derived from the observed total perturbations that contain noise. Consequently, how to remove the noise contribution is a key to the calculations of gravity wave kinetic and potential energy densities. Here we use the calculation of $E_{pm,GW}$ from lidar temperature observations as an example to illustrate this noise issue. The total temperature perturbations $T'_{total}(z,t)$, obtained after the background subtraction and wave filtering process (see details in Chu et al. [2018]), is composed of gravity-wave-induced perturbations $T'_{GW}(z,t)$ and noise-induced perturbations $T'_{noise}(z,t)$:

$$T'_{total}(z,t) = T'_{GW}(z,t) + T'_{noise}(z,t) \tag{5.3}$$

Traditionally, $\overline{T'_{GW}(z,t)}$ should be derived according to equation (5.4):

$$\overline{[T'_{GW}(z,t)]^2} = \overline{[T'_{total}(z,t)]^2} - \overline{[\sigma_T(z,t)]^2},\ \overline{[\sigma_T(z,t)]^2} = \frac{1}{N_p}\sum_{i=1}^{N_p}[\delta T(z,t_i)]^2 \tag{5.4}$$

where $\overline{[\sigma_T(z,t)]^2}$ is the noise variance, and $\delta T(z,t_i)$ is the uncertainty of measured temperature. For the convenience of comparison, we define "total potential energy density" $E_{pm,total}(z)$ and "noise-induced potential energy density" $E_{pm,noise}(z)$ as equations (5.5) and (5.6), respectively, so that $E_{pm,GW}(z) = E_{pm,total}(z) - E_{pm,noise}(z)$:

$$E_{pm,total}(z) = \frac{1}{2}\frac{g^2}{\overline{N^2(z)}}\frac{1}{N_p}\sum_{i=1}^{N_p}\left(\frac{T'_{total}(z,t_i)}{T_{Bkg}(z)}\right)^2 \tag{5.5}$$

$$E_{pm,noise}(z) = \frac{1}{2}\frac{g^2}{\overline{N^2(z)}}\frac{1}{N_p}\sum_{i=1}^{N_p}\left(\frac{\sigma_T(z,t_i)}{T_{Bkg}(z)}\right)^2 \tag{5.6}$$

An example of $E_{pm,GW}(z)$ calculation following this traditional method is shown in Figure 5.1. Figure 5.1a is the lidar temperature observations on 24 June 2011. The overall temperature field ranging from 200 K to 260 K is composed of planetary waves, tides, gravity waves, and other perturbations. The corresponding temperature

uncertainties are shown in Figure 5.1b. The amplitude of the temperature uncertainty increases from 0.10 K to ~1.5 K with altitudes. The temperature uncertainties are relatively small due to the fact that the lidar observations were performed in Antarctic winter when 24-h darkness enabled high signal-to-noise ratios (SNRs). Figure 5.1c shows the total temperature perturbations $T'_{total}(z,t)$. Clear gravity wave perturbation signatures can be identified in this panel. Again $T'_{total}(z,t)$ in Figure 5.1c includes not only $T'_{GW}(z,t)$, but also $T'_{noise}(z,t)$ according to equation (5.3). Figure 5.1d shows the calculated $E_{pm,GW}(z)$ following the traditional calculation procedures. In the above case, $\delta T(z,t_i)$ is small compared to $T'_{total}(z,t)$. Hence, noise variances $\overline{[\sigma_T(z,t)]^2}$ remains small for the entire observation.

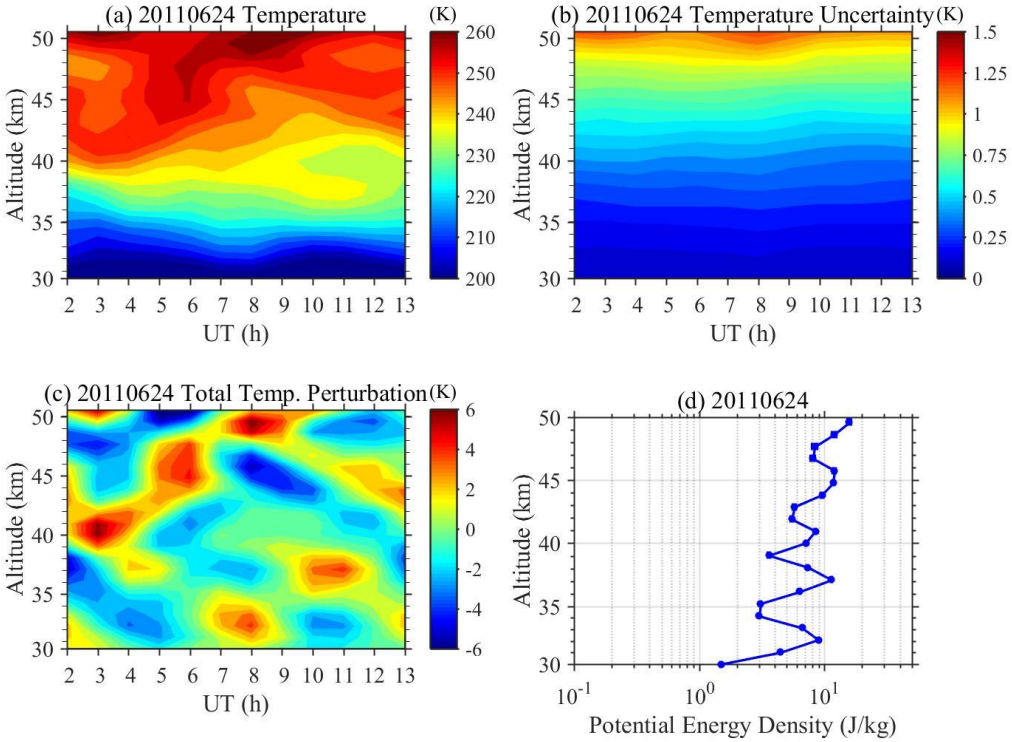

Figure 5.1. (a) Lidar observation of stratospheric temperature on 24 June 2011 from 30 km to 50 km. (b) Temperature uncertainties of the same lidar observation. (c) The derived total temperature perturbations $T'_{total}(z,t)$. (d) The derived gravity wave potential energy density vs. altitude via traditional method.

However, due to sunrise, the lidar observation signals will be considerably contaminated by the solar background. Thus, the derived $\overline{[\sigma_T(z,t)]^2}$ via traditional method can approach and even exceed $\overline{[T'_{total}(z,t)]^2}$. This unrealistic situation can lead to negative $E_{pm,GW}(z)$ according to equation (5.2) and (5.4). Figure 5.2 shows such a case on 01 December 2015 when McMurdo experiences 24-h sunlight. Figures 5.2a and 5.2b illustrate the lidar temperatures and corresponding measurement uncertainties, respectively. The general temperature errors (0.2 K to 7.0 K with increasing altitudes) are much larger than those observed in winter. The extracted $T'_{total}(z,t)$ is shown in Figure 5.2c, while the derived $E_{pm,total}(z)$ and $E_{pm,noise}(z)$ are respectively shown as the black and red lines in Figure 5.2d. $E_{pm,noise}(z)$ begins to exceed $E_{pm,total}(z)$ above ~ 40 km which leads to unrealistic negative $E_{pm,GW}(z)$. Because gravity wave signatures are clearly evident in Figure 5.1c with mostly downward phase progression, theoretically realistic $E_{pm,GW}(z)$ profiles should be derived from such observations. This example demonstrates the disadvantage of the traditional method and consequently, measurements with relatively large temperature uncertainties are excluded due to the resulted unrealistic negative $E_{pm,GW}$, leading to a waste of viable datasets when the desired signals are partially obscured by noise. It is usually impossible to derive exact errors for the observed parameters without any approximations. Thus, the unrealistic negative $E_{pm,GW}$ most likely results from the inevitable over-estimated errors for the observations with low SNRs. These over-estimated errors are inadequate for the calculation of $E_{pm,GW}$ via the traditional method.

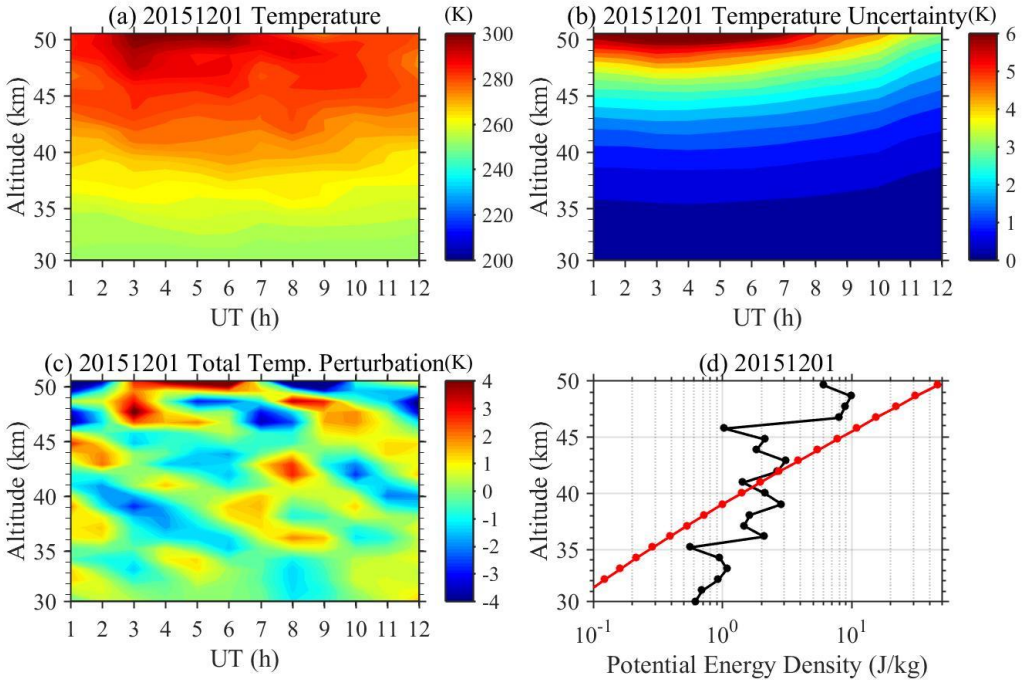

Figure 5.2. (a) Lidar observation of stratospheric temperature on 01 December 2015 from 30 km to 50 km. (b) Temperature uncertainties of the same lidar observation. (c) The derived total temperature perturbations $T'_{total}(z,t)$. (d) The derived "total potential energy density" $E_{pm,total}(z)$ (black) and "noise-induced potential energy density" $E_{pm,noise}(z)$ (red) vs. altitude.

To overcome such a constraint, we proposed a spectral proportion method in Chu et al. [2018] and utilized it to characterize $E_{pm,GW}$ from lidar observations at McMurdo, Antarctica. In the spectral proportion method, we quantify how much energy is in the demanded wave signals relative to all the energy detected via spectra analysis and Monte Carlo simulations. Since the derived wave energy is a portion of the total energy, we avoid the unrealistic results of negative $E_{pm,GW}$ for observations with overestimated uncertainties. Spectral proportion method enables the accurate extraction of gravity wave information from observations with considerably large uncertainty and can potentially utilize large amounts of observations. The main

purpose of this chapter is to provide a rigorous mathematical derivation and forward modeling validation of the spectral proportion method.

5.2. Methodology and Mathematical Derivation of the Spectral Proportion Method

5.2.1. Recapitulation of the Spectral Proportion Method

The spectral proportion method was originally proposed in Chu et al. [2018]. Here we summarize its detailed Monte Carlo calculation procedures.

1) 1000 sets of 2-D temperature map are reconstructed with the simulated temperature field $T_{simu}(z,t)$ centered on the corresponding observed $T(z,t)$ at each grid point but random Gaussian white noise is added. The Gaussian white noise has a mean of 0 and a standard deviation equal to the temperature uncertainty $dT(z,t)$ at this grid point.
2) Run the constructed 1000 $T_{simu}(z,t)$ through the background subtraction and wave filtering process (see details in Chu et al. [2018]) to obtain 1000 total temperature perturbations $T'_{simu}(z,t)$. Now we have 1000 time series of $T'_{simu}(z,t)$ at each altitude.
3) Apply 1D FFT (1 Dimensional Fast Fourier Transform) to these 1000 time series of $T'_{simu}(z,t)$ individually, and obtain 1000 corresponding FFT spectra. Take a mean over these spectra to obtain an averaged FFT spectrum, which will be used to estimate the noise floor and gravity wave area proportion at this altitude.
4) Determine the spectral noise floor on the averaged FFT spectrum by first locating all the minima in the averaged FFT spectrum with frequency larger than 0.15 (1/h). We choose this range because a high-pass filter with cut-off frequency at 0.09 (1/h) is applied during the wave filtering process and we wish to avoid the

influence of the filter on the noise floor estimation. Then we find the vicinity around each minimum. The search for vicinity begins from each minimum and is conducted in both directions in the frequency domain. The search in either direction stops when either the value is larger than 1.04 times of the corresponding minimum or the value begins to decrease. The threshold of 1.04 is chosen empirically considering the performance of the algorithms. We then take the mean of the vicinity for each minimum. Finally, we take the average of the smaller half of the vicinity means as the single noise floor level for frequency range larger than 0.15 (1/h). Because some minima are troughs that well above the noise floor, removing the larger half of the vicinity means likely kicks out such troughs, helping reach the real noise floor. The averaged FFT spectrum below frequency 0.15 (1/h) is taken directly as the noise floor levels for the low frequency range, considering the wave filtering process in step 2).

5) In the averaged FFT spectrum, take the summation of the discrete energy values below the spectral noise floor to obtain the noise energy ($Energy_{Noise}$) and below the entire FFT spectrum to obtain the total energy ($Energy_{Total}$). The gravity wave energy proportion $P(z)$ at this altitude is then computed as

$$P(z) = \frac{Energy_{Total} - Energy_{Noise}}{Energy_{Total}} \tag{5.7}$$

Repeat $P(z)$ calculation for every altitude starting from step 3).

6) $\overline{T'_{GW}(z,t)}$ and $\overline{[\sigma_T(z,t)]^2}$ are computed as

$$\overline{[T'_{GW}(z,t)]^2} = \overline{[T'_{total}(z,t)]^2} \bullet P(z) \tag{5.8}$$

$$\overline{[\sigma_T(z,t)]^2} = \overline{[T'_{total}(z,t)]^2} \bullet [1 - P(z)] \tag{5.9}$$

Note that $T'_{total}(z,t)$ in equations (5.8) and (5.9) is the same total temperature perturbations as in Section 5.1 (e.g., equation (5.3)).

7) $E_{pm,GW}(z)$ are calculated with equation (5.10) (derived from equation (5.2)) and the gravity wave potential energy density errors $\delta E_{pm,GW}(z)$ are calculated with equation (5.11)

$$E_{pm,GW}(z) = \frac{1}{2}\frac{g^2}{\overline{N^2(z)}}\overline{\left(\frac{T'_{GW}(z,t)}{T_{Bkg}(z)}\right)^2} = \frac{1}{2}\frac{g^2}{\overline{N^2(z)}}\overline{\left(\frac{T'_{total}(z,t)}{T_{Bkg}(z)}\right)^2} \bullet P(z) \tag{5.10}$$

$$\delta E_{pm,GW}(z) = \frac{1}{2}\frac{g^2}{\overline{N^2(z)}}\overline{\left(\frac{\sigma_T(z,t)}{T_{Bkg}(z)}\right)^2}\frac{1}{\sqrt{N_p}} = \frac{1}{2}\frac{g^2}{\overline{N^2(z)}}\overline{\left(\frac{T'_{total}(z,t)}{T_{Bkg}(z)}\right)^2}[1-P(z)]\frac{1}{\sqrt{N_p}} \tag{5.11}$$

Since the gravity wave area proportion $P(z)$ varies only between 0 and 1, the derived gravity wave perturbations will always be a portion of the total temperature perturbations. Thus, the spectral proportion method avoids unrealistic negative calculation results for $E_{pm,GW}(z)$.

Figure 5.3 shows the same case for 01 December 2015 but the proposed spectral proportion method was employed to calculate $E_{pm,GW}(z)$. 3 of the 1000 simulated total temperature perturbations $T'_{simu}(z,t)$ (red lines) along with the original total temperature perturbations $T'_{total}(z,t)$ (blue line) at 48.68 km are shown in Figure 5.3a. The temperature uncertainty introduces variations on the total temperature perturbations. After the Monte Carlo simulations for 48.64 km and taking the average of the 1000 FFT spectra, the mean spectrum is shown in Figure 5.3b by the blue line. We use green dots to indicate the regional minima found on the FFT curve. The final determined noise floor is denoted by the red dashed line. Notice this spectrum went through a high-pass Butterworth filter with cutoff frequency 1/11 (1/h). The gravity wave area proportion $P(z)$ for every altitude is illustrated in Figure 5.3c. It is obvious that as the temperature uncertainties increase with increasing altitudes, $P(z)$ generally decreases due to increasing noise floor level. We obtain the $E_{pm,GW}(z)$

which is illustrated by the red line along with its error bars in Figure 5.3d. The blue line shows $E_{pm,total}(z)$ as a reference. As the altitude increases, due to the increase of temperature uncertainty, $E_{pm,GW}(z)$ occupies a decreasing amount of $E_{pm,total}(z)$.

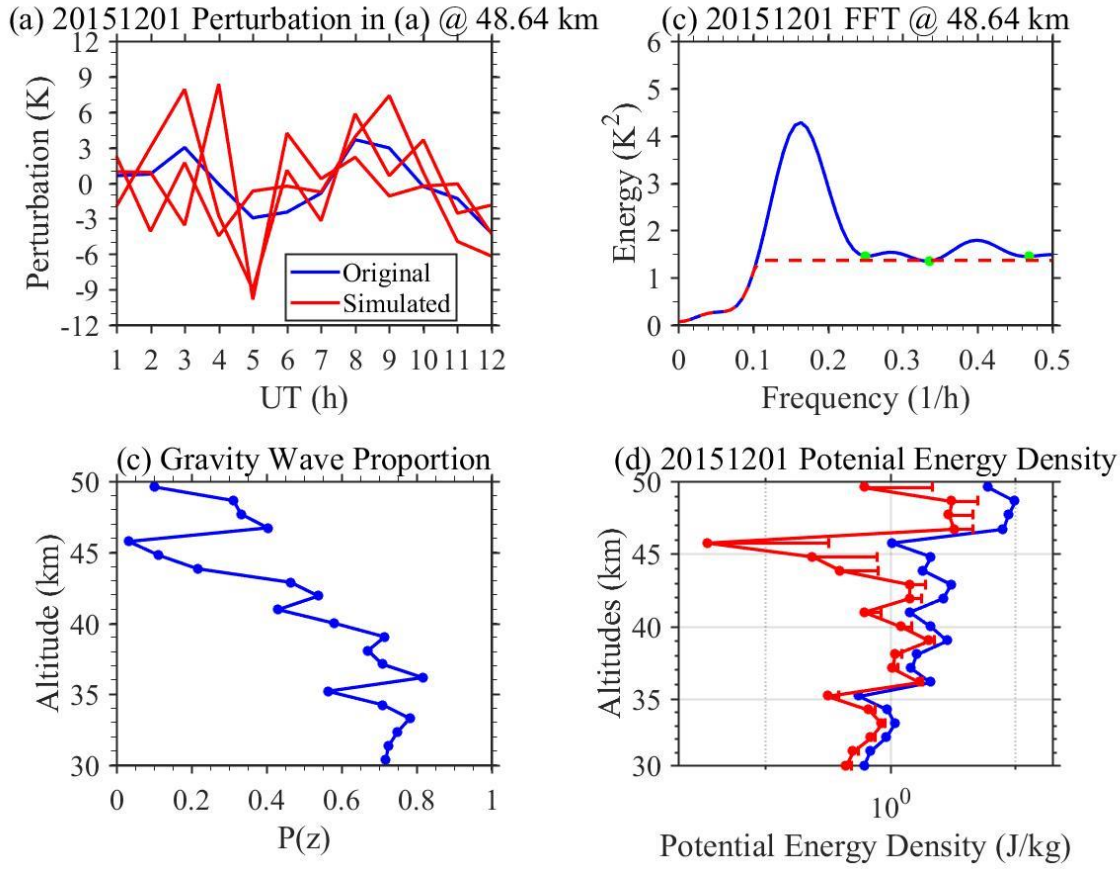

Figure 5.3. (a) Original (blue) and 3 of the 1000 simulated (red) total temperature perturbations at 48.64 km on 01 December 2015. (b) The obtained mean FFT spectrum after 1000 simulations at 48.64 km. The green dotes are the regional minimums at frequency range larger than 0.15 (1/h). The red dashed line is the determined noise floor. (c) The derived gravity wave area proportion. (d) The calculated $E_{pm,GW}(z)$ (red) via spectral proportion method and the $E_{pm,total}(z)$ as a reference (blue). Error bars are only shown at the positive side of the profile for better readability.

5.2.2. Rigorous Mathematical Formalism of the Spectral Proportion Method

In this section, we develop the mathematical formalism behind this methodology and a key relation applied is the Parseval's theorem for Fourier transform (the more detailed derivations are also attached in Appendix I). Briefly, the Parseval's theorem states the conservation of energy between the signals in the time domain and in the frequency domain upon Fourier transform. In Continuous Fourier Transform (CFT), the Parseval's theorem states

$$\int_{-\infty}^{+\infty} \left|T'(t)\right|^2 dt = \int_{-\infty}^{+\infty} \left|T'^{,FFT}(\xi)\right|^2 d\xi \tag{5.12}$$

where t and ξ denote time and frequency, respectively, $T'(t)$ is a continuous temperature perturbation series in the time domain, and $T'^{,FFT}(\xi)$ is the corresponding CFT in the frequency domain. For the case of Discrete Fourier Transform (DFT), the Parseval's theorem has the form of

$$\sum_{i=0}^{N_p-1} \left|T'_{total}(t_i)\right|^2 = \frac{1}{N_p} \sum_{k=0}^{N_p-1} \left|T'^{,FFT}_{total}(\xi_k)\right|^2 \tag{5.13}$$

where t_i and ξ_k denote discrete time and frequency, respectively, N_p is the length of the temperature observation in the time domain. Notice here for simplicity, we only consider one single altitude, so we ignore the altitude parameter z and define $T'_{total}(t_i) = T'_{total}(z,t_i)$. $T'^{,FFT}_{total}(\xi_k)$ is the corresponding DFT of $T'_{total}(t_i)$. Notice that the major difference of the Parseval's theorem in CFT and DFT is that in CFT the theorem is in the form of integration whereas in DFT the theorem is in the form of summation. Recall in equation (5.10), we define gravity wave proportion P as on average, how much the energy of the gravity wave induced perturbations $\sum_{i=0}^{N_p-1} \left|T'_{GW}(t_i)\right|^2$ occupy the total energy of the total temperature perturbations $\sum_{i=0}^{N_p-1} \left|T'_{total}(t_i)\right|^2$ in time domain. Hence,

$$P = \sum_{i=0}^{N_p-1} \left|T'_{GW}(t_i)\right|^2 \Big/ \sum_{i=0}^{N_p-1} \left|T'_{total}(t_i)\right|^2 \tag{5.14}$$

The key is to estimate the gravity wave proportion P so that we know everything on the right-hand side of equation (5.10). It is difficult to estimate P in the time domain since the perturbations induced by gravity waves and temperature uncertainties are combined. We attempt to solve for P in the frequency domain according to equation (5.13). Here we confirm that if we obtain gravity wave proportion P_f in the frequency domain, then the gravity wave proportion P in the time domain should be exactly equal to P_f in the frequency domain via mathematical inductions. We ignore the altitude parameter z in equation (5.3),

$$T'_{total}(t_i) = T'_{GW}(t_i) + T'_{noise}(t_i) \tag{5.15}$$

According to the linearity property of FFT,

$$T'^{,FFT}_{total}(\xi_k) = T'^{,FFT}_{GW}(\xi_k) + T'^{,FFT}_{noise}(\xi_k) \tag{5.16}$$

where $T'^{,FFT}_{total}(\xi_k), T'^{,FFT}_{GW}(\xi_k), T'^{,FFT}_{noise}(\xi_k)$ are the corresponding FFT of $T'_{total}(t_i), T'_{GW}(t_i), T'_{noise}(t_i)$. Here we invoke Parseval's theorem (i.e., equation (5.13)) so that,

$$\sum_{i=0}^{N_p-1} \left|T'_{total}(t_i)\right|^2 = \frac{1}{N_p^{FFT}} \sum_{k=0}^{N_p^{FFT}-1} \left|T'^{,FFT}_{total}(\xi_k)\right|^2 \tag{5.17}$$

$$\sum_{i=0}^{N_p-1} \left|T'_{GW}(t_i)\right|^2 = \frac{1}{N_p^{FFT}} \sum_{k=0}^{N_p^{FFT}-1} \left|T'^{,FFT}_{GW}(\xi_k)\right|^2 \tag{5.18}$$

$$\sum_{i=0}^{N_p-1} \left|T'_{noise}(t_i)\right|^2 = \frac{1}{N_p^{FFT}} \sum_{k=0}^{N_p^{FFT}-1} \left|T'^{,FFT}_{noise}(\xi_k)\right|^2 \tag{5.19}$$

Since we employed zero padding before applying FFT on the original time series, here N_p represents the length of the original time series, N_p^{FFT} represents the length of the time series after zero padding. Substitute equation (5.15) and (5.16) into the left and right side of equation (5.17), respectively, we have

$$\sum_{i=0}^{N_p-1}\left|T'_{GW}(t_i)\right|^2+\sum_{i=0}^{N_p-1}\left|T'_{noise}(t_i)\right|^2+\sum_{i=0}^{N_p-1}\left[2\bullet T'_{GW}(t_i)\bullet T'_{noise}(t_i)\right]$$
$$=\frac{1}{N_p^{FFT}}\sum_{k=0}^{N_p^{FFT}-1}\left|T'^{,FFT}_{GW}(\xi_k)\right|^2+\frac{1}{N_p^{FFT}}\sum_{k=0}^{N_p^{FFT}-1}\left|T'^{,FFT}_{noise}(\xi_k)\right|^2+\frac{1}{N_p^{FFT}}\sum_{i=0}^{N_p^{FFT}-1}\left[2\bullet T'^{,FFT}_{GW}(\xi_k)\bullet T'^{,FFT}_{noise}(\xi_k)\right] \tag{5.20}$$

Considering equation (5.18), (5.19), and (5.20), we have

$$\sum_{i=0}^{N_p-1}\left[2\bullet T'_{GW}(t_i)\bullet T'_{noise}(t_i)\right]=\frac{1}{N_p^{FFT}}\sum_{i=0}^{N_p^{FFT}-1}\left[2\bullet T'^{,FFT}_{GW}(\xi_k)\bullet T'^{,FFT}_{noise}(\xi_k)\right] \tag{5.21}$$

We recall the Monte Carlo simulation was performed previously where we simulated the temperature observation field for 1000 times using the combination of the observed temperature and temperature uncertainties. After the background subtraction and the wave filtering process, we obtain the simulated total temperature perturbation field $T'_{simu,j}(t_i)$. Here $j=1,2,...,1000$ denotes the 1000 instances of simulations. In this case, we have

$$T'_{simu,j}(t_i)=T'_{GW,j}(t_i)+T'_{noise,j}(t_i) \tag{5.22}$$

Since $T'_{GW,j}(t_i)$ are the perturbations induced purely by gravity waves, theoretically, they should be constant across different simulations (i.e., for every j). $T'_{noise,j}(t_i)$ are the perturbations induced by noise (temperature uncertainty). We apply FFT to the 1000 simulated perturbation time series $T'_{simu,j}(t_i)$ individually, and obtain a mean FFT spectrum from the 1000 FFT spectra. Once we transform the perturbations into the frequency domain (i.e., the mean FFT spectrum), it is relatively easy to estimate gravity wave proportion since we can intuitively determine the noise floor level as described previously. In order to make the plotted averaged FFT spectrum more illustrative, we only showed the positive portion of the FFT (e.g. Figure 5.3b) since FFT gives symmetric results for positive and negative frequencies with respect to frequency 0. We actually plotted $\frac{1}{1000}\sum_{j=1}^{1000}\left|\frac{T'^{,FFT}_{simu,j}(\xi_k)}{N_p^{FFT}}\right|^2$ for the averaged FFT spectrum where $T'^{,FFT}_{simu,j}(\xi_k)$ are the complex numbers we obtain directly from the

FFT algorithm in the case of the transformation of $T'_{simu,j}(t_i)$. The purpose here is to illustrate the preservation of energy from the time domain to the frequency domain (i.e. the summation of the discrete energy values on the FFT spectrum is equal to the summation of the squared time series in the time domain).

Hence, similarly to equation (5.17), (5.18), and (5.21), from this 1000 simulations, we have

$$\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{simu,j}(t_i)\right|^2\right)=\frac{1}{N_p^{FFT}}\sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{simu,j}(\xi_k)\right|^2\right) \tag{5.23}$$

$$\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{GW,j}(t_i)\right|^2\right)=\frac{1}{N_p^{FFT}}\sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{GW,j}(\xi_k)\right|^2\right) \tag{5.24}$$

$$\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left[2\bullet T'_{GW,j}(t_i)\bullet T'_{noise,j}(t_i)\right]\right)=\frac{1}{N_p^{FFT}}\sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left[2\bullet T'^{,FFT}_{GW,j}(\xi_k)\bullet T'^{,FFT}_{noise,j}(\xi_k)\right]\right) \tag{5.25}$$

where $T'^{,FFT}_{simu,j}(\xi_k), T'^{,FFT}_{GW,j}(\xi_k), T'^{,FFT}_{noise,j}(\xi_k)$ are the corresponding FFT of $T'_{simu,j}(t_i), T'_{GW,j}(t_i), T'_{noise,j}(t_i)$. We consider $T'_{GW,j}(t_i)$ and $T'_{noise,j}(t_i)$ to be uncorrelated, and $\frac{1}{1000}\sum_{j=1}^{1000}\left[T'_{noise,j}(t_i)\right]=\vec{0}$ since $T'_{noise,j}(t_i)$ has a mean of 0 at different t_i among different simulations. According to the general correlation coefficient calculation equation $\rho_{X,Y}=\frac{E[XY]-E[X]E[Y]}{\sqrt{E[X^2]-[E[X]]^2}\sqrt{E[Y^2]-[E[Y]]^2}}=0$, where we regard X as $T'_{GW,j}(t_i)$, Y as $T'_{noise,j}(t_i)$, equation (5.25) is equal to 0. Hence, according to equation (5.20),

$$\begin{aligned}\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{simu,j}(t_i)\right|^2\right)&=\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{GW,j}(t_i)\right|^2\right)+\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{noise,j}(t_i)\right|^2\right)\\&=\frac{1}{N_p^{FFT}}\sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{simu,j}(\xi_k)\right|^2\right)\\&=\frac{1}{N_p^{FFT}}\sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{GW,j}(\xi_k)\right|^2\right)+\frac{1}{N_p^{FFT}}\sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{noise,j}(\xi_k)\right|^2\right)\end{aligned} \tag{5.26}$$

We now take the ratio of equation (5.24) over (5.23), we have

$$P_f = \sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{GW,j}(\xi_k)}{N_p}\right|^2\right)\Bigg/\sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{simu,j}(\xi_k)}{N_p}\right|^2\right)$$
$$= \sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{GW,j}(t_i)\right|^2\right)\Bigg/\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{simu,j}(t_i)\right|^2\right) = P \quad (5.27)$$

The term $\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{GW,j}(t_i)\right|^2\right)\Big/\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{simu,j}(t_i)\right|^2\right)$ in equation (5.27) denotes in the time domain, on average, how much gravity wave perturbation energy should occupy the derived total temperature perturbation energy (which is the definition of P). Hence, we confirm $P = P_f$. An important note here is that we plan to submit the work in this chapter to Journal of Atmospheric and Oceanic Technology, and the mathematical derivations presented here still need more considerations to rigorously prove the spectral proportion method.

In the algorithm, this gravity wave area portion P_f is calculated via the first equality in equation (5.27). We first determine the spectral noise floor level of the obtained averaged FFT spectrum in the frequency domain as described before. By doing that, we select the gravity wave area (the area below the mean FFT curve and above the noise floor level within the obtained frequency range which corresponds to the term $\sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|\frac{T'^{,FFT}_{GW,j}(\xi_k)}{N_p^{FFT}}\right|^2\right)$) and the white noise area (the area below the noise floor level and above 0 within the obtained frequency range which corresponds to the term $\sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|\frac{T'^{,FFT}_{simu,j}(\xi_k)}{N_p^{FFT}}\right|^2\right) - \sum_{k=0}^{N_p^{FFT}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|\frac{T'^{,FFT}_{GW,j}(\xi_k)}{N_p^{FFT}}\right|^2\right)$). According to equation (5.27), when we calculate the proportions using the decided gravity wave and white noise areas, we should take the summation or average of the discrete energy values in the selected areas in the averaged FFT spectrum. Thus, gravity wave proportion can be calculated from the sums within the selected areas and $E_{pm,GW}$ can be solved for according to equation (5.10).

As introduced in Section 5.1, the spectral proportion method can also be applied to the calculation of other atmospheric parameters such as $E_{km,GW}$. We first derive the total zonal ($u'(z,t)$) and meridional ($v'(z,t)$) wind perturbations utilizing background subtraction and wave filtering process from the observed $u(z,t), v(z,t)$. When extracting the purely gravity waves induced perturbations $u'_{GW}(z,t)$ and $v'_{GW}(z,t)$ from $u'(z,t), v'(z,t)$ and $\delta u(z,t), \delta v(z,t)$, the gravity wave area proportions P_u, P_v should be estimated in the frequency domain utilizing the spectral proportion method, so that

$$E_{km,GW}(z) = \frac{1}{2}(\overline{u'^2_{GW}(z,t)} + \overline{v'^2_{GW}(z,t)}) = \frac{1}{2}(\overline{u'^2(z,t)} \bullet P_u + \overline{v'^2(z,t)} \bullet P_v) \tag{5.28}$$

In a more general sense, assume we have an observation of quantity $Q(t)$ and its corresponding uncertainties $\delta Q(t)$ with respect to time t. We extract certain perturbations $Q'(t)$ from $Q(t)$ depending on what part of the signal spectrum we are interested in. We wish to calculate the quantity $\overline{(Q'_{signal}(t))^2}$ from $\overline{(Q'(t))^2}$ by excluding the influence of $\delta Q(t)$ on $Q'(t)$, therefore the spectral proportion method applies.

5.3. Validation with Forward Modeling

We now validate this method with forward modeling tests. 1000 time series of simulated temperature observations are constructed with a constant background, a planetary wave, two monochromatic gravity waves and randomly generated Gaussian white noise, following the format of

$$y = 240 + 4\cos(\frac{2\pi}{36}t) + 2\cos(\frac{2\pi}{8}t) + 1\cos(\frac{2\pi}{5}t) + Noise \tag{5.29}$$

where 240 K is the constant temperature background. The amplitude and period of the planetary wave are set to 4 K and 36 h, respectively. The two gravity waves have periods of 8 and 5 h, and amplitudes of 2 and 1 K, respectively. The $Noise$ term is the Gaussian white noise generated by drawing random numbers from a normal distribution with a mean of 0 and a standard deviation of 4 K ($\sigma_{Noise} = 4K$), which is twice the preset amplitude of the strongest gravity wave. An example of the generated

temperature time series is shown in Figure 5.4a. The solid red line indicates the combination of the background, the planetary wave, and the gravity waves. The black line denotes the randomly drawn Gaussian white noise (artificially shifted 250 K to fit the white noise in the same y-axis range in this plot). The overall simulated temperature time series is shown as the blue line in Figure 5.4a. Because the two monochromatic gravity waves are manually implemented, the exact gravity wave energy put into the time series is known, which is indicated by $Energy_{GW}$. Now we let the 1000 simulated temperature observations from equation (5.29) go through background subtraction, wave filtering, and 1D FFT to obtain the 1000 corresponding spectra. An average was taken over these 1000 spectra to obtain a mean spectrum on which we will locate the noise floor. This averaged spectrum is illustrated in Figure 5.4b. Here we can immediately locate the two peaks corresponding to the preset two gravity waves with frequencies of ~ 0.125 1/h and ~ 0.2 1/h. The planetary wave with a period of 36 h is filtered out during the wave filtering process when we apply a Butterworth high pass filter with cutoff frequency 1/11 (1/h). The small ripples beside the main peaks are caused by the implied square window imposed to the time series before performing FFT. We realized that applying other windows (e.g., Hanning, Hamming, Blackman, etc.) to the time series causes dramatic changes to the time series itself. Since the energy still conserves from the time domain to the frequency domain even with the square window induced ripples, we decided not to impose other windows to the time series before FFT. The determination of the noise floor is similar to what we have described previously. We first locate all the regional minima in the averaged FFT spectrum among frequency ranges larger than 0.15 (1/h) as indicated by the green dots in Figure 5.4b. Then we find the vicinity around every minimum and solve for their corresponding vicinity means. Finally, we take the mean of the smaller half vicinity means as the level of noise floor for frequency ranges larger than 0.15 (1/h). For frequency ranges smaller than 0.15 (1/h), we use the averaged FFT

spectrum as the noise floor. We take the summation of the discrete energy values below the spectral noise floor and below the mean FFT spectrum to obtain the white noise energy ($Energy_{Noise}$) and the total temperature perturbation energy ($Energy_{Total}$), respectively. The gravity wave energy ($Energy_{GW,SPM}$) obtained via the spectral proportion method is determined by $Energy_{GW,SPM} = Energy_{Total} - Energy_{Noise}$.

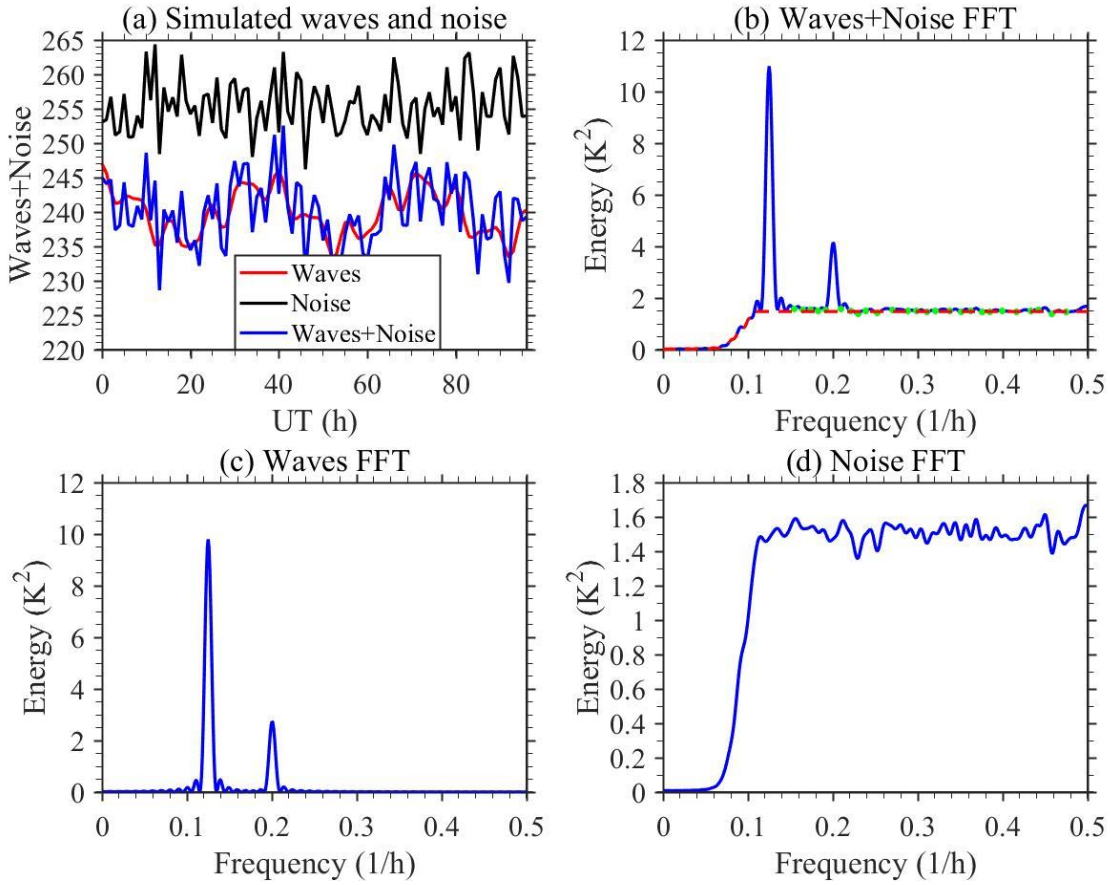

Figure 5.4. (a) Simulated waves and Gaussian white noise in time domain. Red curve denotes the combination of background, planetary wave, and gravity waves. Black curve denotes the randomly generated Gaussian white noise (Note in order to show the white noise time series in the same plot, we artificially added a mean of 250 to shift the time series in the y-axis range, this is done only when plotting this panel). Blue curve denotes the overall simulated total temperature time series. (b) Mean FFT spectrum of the 1000 simulated (gravity waves + Gaussian white noise) total temperature perturbations with all the regional minimums indicated by the green

dots. The determined noise floor is indicated by the red dashed line. (c) Mean FFT spectrum of the 1000 simulated gravity wave time series. (d) Mean FFT spectrum of the 1000 simulated Gaussian white noise time series.

We realize there are two significant processes that might influence the determination of the gravity wave energy in the spectral proportion method. The first one is the modulation of gravity wave energy from the background subtraction and wave filtering process. The second one is the algorithms for the noise floor determination. Hence in the following forward modeling validation, we compare $Energy_{GW}$ with $Energy_{GW,SPM}$ to investigate the overall effects on the determination of gravity wave energy from the background subtraction, wave filtering, and the noise floor determination algorithms. We also perform a second comparison to assess the effects only from noise floor determination algorithms. To achieve this, we run the 1000 wave time series (the combination of the background, the planetary wave, and the gravity waves) through the background subtraction, wave filtering, and 1DFFT. The mean FFT spectrum is obtained by taking an average of the 1000 gravity wave spectra, which is illustrated in Figure 5.4c. The two preset gravity waves can be immediately identified again. By taking the summation of the discrete energy values in this mean FFT spectrum, we get a theoretical value of how much gravity wave energy remains after the two preset gravity waves are imposed with a planetary wave and a background and go through the background subtraction and wave filtering process. We represent this energy with $Energy_{GW,Filter}$. $Energy_{GW,Filter}$ is compared with $Energy_{GW,SPM}$ to isolate the effects from the background subtraction and wave filtering process, hence, only investigate the effects from the noise floor determination algorithms when extracting gravity wave energy in the spectral proportion method. Similarly, we can also run the 1000 simulated Gaussian white noise time series through background subtraction, wave filtering, 1D FFT, and obtain the simulated Gaussian white noise spectrum in frequency domain which is shown in Figure 5.4d.

We can justify this is a white noise spectrum due to its near constant value across the spectral range from 0.1 to 0.5 (1/h). The small values at frequency ranges from 0 to ~0.08 (1/h) results from the high pass filter.

We define two gravity wave energy relative differences as,

$$\Delta Enegy1 = (Energy_{GW,SPM} - Energy_{GW}) / Energy_{GW} \tag{5.30}$$

$$\Delta Enegy2 = (Energy_{GW,SPM} - Energy_{GW,Filter}) / Energy_{GW,Filter} \tag{5.31}$$

Here we explain the reason we compare two quantities $\Delta Energy1$ and $\Delta Energy2$. We set up the simulated temperature observations in a very specific way with specific monochromatic planetary wave, gravity waves, and background. In order to extract gravity waves from such observations, 6th order highpass Butterworth filter with cutoff frequency 1/11 (1/h) is chosen. All these criteria can be varied in different situaitons and thus affect the following forward modeling validation. For example, we could have set up the simulated temperature observatinos in more complicated manners, or chosen a slightly different filter. Hence, we compare the two quantities to isolate two major steps in the spectral proportion method to investigate the influences separately. As an example, with the above stated parameters, temporal resolution $\triangle t = 1h$, simulated time length $t_{total} = 96h$, and 1000 instances of simulations, $\Delta Energy1$ and $\Delta Energy2$ are 0.068 and 0.038, respectively, which are regarded as small. There are many variables which might affect the performance of the proposed E_{pm} calculation method. These parameters include: 1) the number of simulations, 2) the relative amplitude of noise to waves, 3) the length of zero padding upon FFT, 4) the temporal length of a single simulated time series, 5) the temporal resolution of a single simulated time series. We will discuss these parameters individually in terms of their influence on the algorithm's performance. For all simulations below, we change $\sigma_{Noise} = 2K$, which enables good simulation conditions versus when σ_{Noise} is large. All other parameters remain the same as above unless stated otherwise.

1) The number of simulations

The number of simulations affect the accuracy of the results essentially because this number represents whether we have generated sufficient samples for the Monte Carlo process. Once this number approaches a critical value that is large enough, the gravity wave energy determined from the spectral proportion method $Energy_{GW,SPM}$ will converge to the theoretical gravity wave area energies $Energy_{GW}$ and $Energy_{GW,Filter}$. As the test results indicate in Figure 5.5a and 5.6a, when the number of simulations increases from 100 to 2000, the gravity wave energy relative differences $\Delta Energy1$ and $\Delta Energy2$ become smaller from 10^{-1} to 10^{-2}. For most of the simulation conditions, they show similar trend and are on the order of 10^{-2}.

2) The relative amplitude of noise to waves

We vary the standard deviation of the randomly generated Gaussian white noise σ_{Noise} from 0.01 K to 4 K to investigate how the results change corresponding to this variation. The results in Figure 5.5b and 5.6b show that $\Delta Energy1$ and $\Delta Energy2$ generally becomes larger once the σ_{Noise} is comparable to or larger than the amplitudes of the preset gravity waves. $\Delta Energy1$ and $\Delta Energy2$ reach ~0.04–0.06 when σ_{Noise} is set to 4 K. It is important to note that the preset two gravity waves have amplitudes of 2 K and 1 K, respectively, which are smaller than σ_{Noise} = 4 K in this case. When the noise amplitude is smaller (less than 1 K), $\Delta Energy1$ and $\Delta Energy2$ tend to stay small and share similar trend.

3) The length of zero padding on FFT

The length of zero padding here refers to the final length of the time series after we add zeros onto the end of the original time series before performing the 1D FFT. We choose the power of 2 to be the final length of the time series to be put into FFT. As Figure 5.5c and 5.6c shows, we use 2^7 to 2^{12} as the final length of the time series after zero padding. Note that the original time length is 96 in this case. $\Delta Energy1$ and $\Delta Energy2$ decrease once the zero padding goes to the next 4^{th} power of

2 of the original time length (i.e., 96 is the original time length, 128 is the next 1st power of 2 of the original time length, 256 is the next 2nd power of 2 of the original time length, …, 1024 is the next 4th power of 2 of the original time length). $\Delta Energy1$ and $\Delta Energy2$ do not change significantly (on the order of 10^{-2}) within this test range, possibly due to zero padding only changing the resolution of the obtained mean FFT spectrum, but does not significantly influence the simulation errors. However, the absolute values of $\Delta Energy1$ and $\Delta Energy2$ tend to converge with more zero paddings, which represents that zero paddings have reached a long enough value such that it does not affect algorithm performance dramatically. Obviously, $\Delta Energy1$ is always larger than $\Delta Energy2$, thus, we judge $Energy_{GW}$ is smaller than $Energy_{GW,Filter}$. This results from the addition of the planeraty wave in the temperature observation simulation and the applied 6th order Butterworth highpass filter spectal leakage in the wave filtering process. Hence, we can extrapolate that in most such forward modeling tests, $Energy_{GW}$ is likely not equal to $Energy_{GW,Filter}$ due to technical issues in reality.

4) The temporal length of a single simulated time series

The temporal length of a single simulated time series t_{total} is varied from 12 to 768. Attentions need to be paid that when changing this parameter, the length of zero padding will also change. This is because the length of the time series is zero padded to the next 4th power of 2 before performing FFT. With longer temporal length, $\Delta Energy1$ and $\Delta Energy2$ varies around 0 on the order of 10^{-2}, which means both $\Delta Energy1$ and $\Delta Energy2$ also converge in this case. Within the test range in Figure 5.5d and 5.6d, significant differences between $\Delta Energy1$ and $\Delta Energy2$ show up when t_{total} is short. Such situations occur due to that the manner of temperature observation reconstruction (what waves are included), the background subtraction, and the wave filtering process play more vital roles when t_{total} is short. However, even in such conditions, $\Delta Energy1$ and $\Delta Energy2$ are still regarded small considering the

forward modeling is performed at σ_{Noise} =2 K, which is the same as the amplitude of the strongest gravity wave.

5) The temporal resolution of a single simulated time series

The temporal resolution of a single simulated time series $\triangle t$ was varied from 0.2 to 2. The length of zero padding will also change when we alter this parameter. That is due to the fact that the number length of a single simulated time series t_{total} will consequently be different when we vary $\triangle t$. The results in Figure 5.5e and 5.6e show $\Delta Energy1$ and $\Delta Energy2$ usually stay on the order of 10^{-2} and share the same trend within the test range. Although there is a decrease trend (also a zero crossing) of both $\Delta Energy1$ and $\Delta Energy2$ when the resolutions become courser, the results generally indicate the performance of the spectral proportion method does not strongly depend on the temporal resolution of a single simulated time series. Note that if we continue to increase $\triangle t$ and test with $\triangle t$ equal to 4, then the Nyquist frequency will become 0.125 (1/h), which will hinder the determination of the noise floor since we have a preset 0.2 (1/h) gravity wave. Hence, we chose to stop at $\triangle t = 2$.

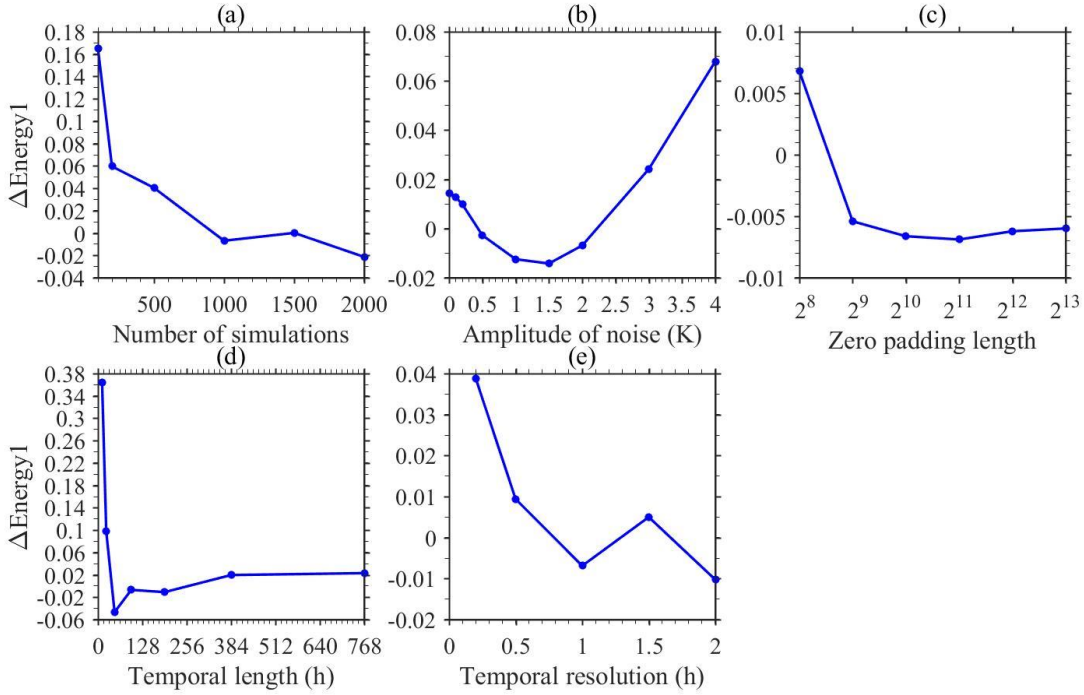

Figure 5.5. Variations of gravity wave energy relative difference $\Delta Energy1$ as the number of simulations (a), the relative amplitude of noise to waves (b), the length of zero padding (c), the temporal length of a single simulated time series (d), and the temporal resolution of a single simulated time series (e) increases.

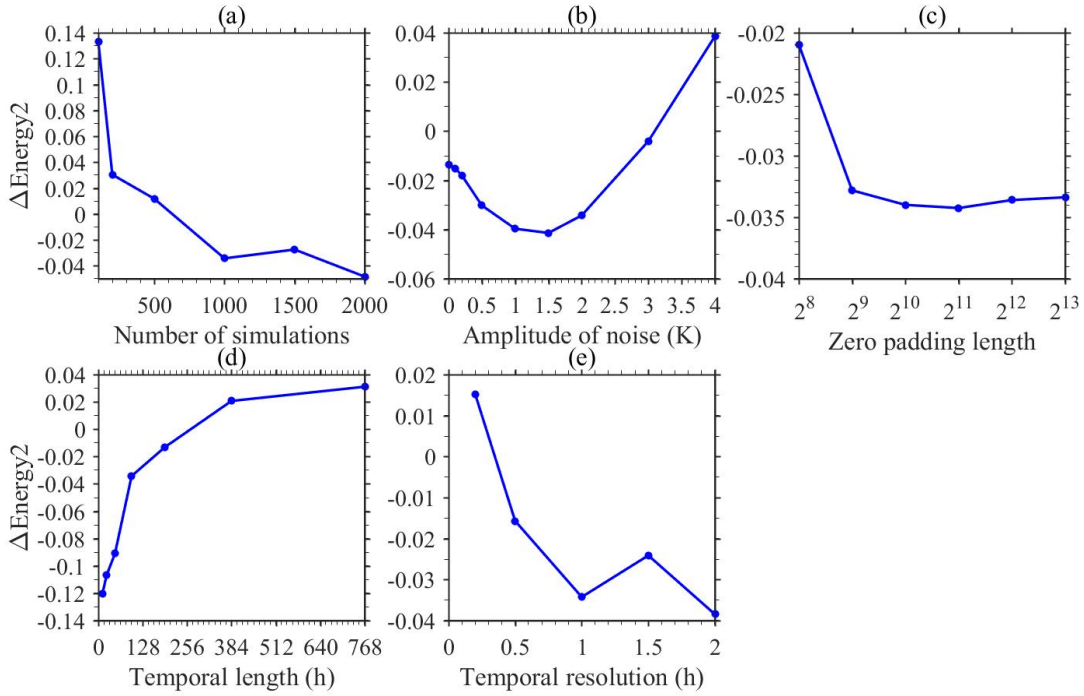

Figure 5.6. Variations of gravity wave energy relative difference $\Delta Energy2$ as the number of simulations (a), the relative amplitude of noise to waves (b), the length of zero padding (c), the temporal length of a single simulated time series (d), and the temporal resolution of a single simulated time series (e) increases.

5.4. Application to Lidar Observations from McMurdo

We now compare the spectral proportion method with the traditional method using lidar temperature measurements with small temperature uncertainties in the case of calculating $E_{pm,GW}(z)$. Figure 5.7 illustrates the $E_{pm,GW}(z)$ calculation using spectral proportion method for the case on 24 June 2011. The original ($T'_{total}(z,t)$) and three simulated ($T'_{simu}(z,t)$) total temperature perturbation time series at 48.64 km

are illustrated in Figure 5.7a via blue and red curves, respectively. The curves show better agreement compared to the case on 01 December 2015 due to the smaller temperature uncertainties in wintertime when McMurdo is under 24 h darkness. The calculated mean FFT spectrum after the 1000 simulations for 48.64 km is shown in Figure 5.7b with the regional minima indicated by green dots and the determined noise floor shown by a red horizontal dashed line. Figure 5.7c illustrates the derived $P(z)$ after locating the noise floor levels. $P(z)$ still generally decreases with increasing altitude due to the increase of temperature uncertainty. However, the decrease is much slower than the case on 01 December 2015 due to overall lower temperature uncertainties. In Figure 5.7d, we show the calculated $E_{pm,GW}(z)$ along with its error bars via a blue curve. We also show $E_{pm,GW}(z)$ calculated through the traditional method by a red line. We can immediately conclude that for the observations with small temperature uncertainties, the two $E_{pm,GW}(z)$ calculation methods give rise to similar results. This comparison validates our spectral proportion calculation method within the working range of the traditional method (i.e., observations with small observational uncertainties). Hence, we demonstrated that the spectral proportion method can give rise to valid $E_{pm,GW}$ in scenarios where the temperature uncertainties are either large or small.

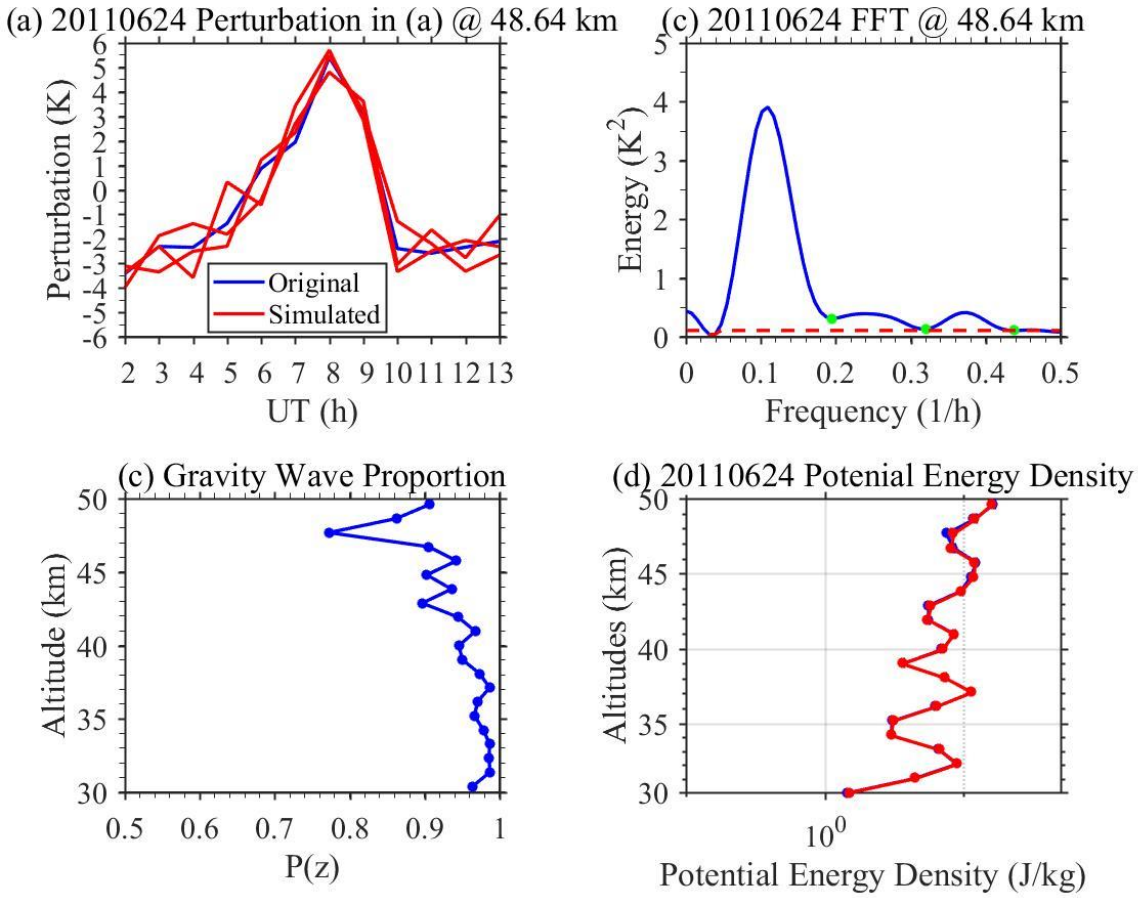

Figure 5.7. (a) Original (blue, $T^{'}(z,t)$) and 3 of the 1000 simulated (red, $T^{'}_{Simu}(z,t)$) total temperature perturbations at 48.64 km on 24 June 2011. (b) The obtained mean FFT spectrum after 1000 simulations for 48.64 km. Green dots are the regional minimums. Red dashed line indicates the determined noise floor. (c) The derived gravity wave proportion $P(z)$. (d) The calculated $E_{pm}(z)$, $\delta E_{pm}(z)$ (blue) using spectral proportion method. The red line indicates the $E_{pm}(z)$ derived via traditional method.

Utilizing the spectral proportion method, we are capable of deriving $E_{pm,GW}$ time series from the year 2011 to 2015 at McMurdo, Antarctica (altitude mean from 30 km to 50 km). Figure 5.8 illustrates the 5-year variations of $E_{pm,GW}$ (note all the observations have 12 h data length). The red dots are the $E_{pm,GW}$ produced by the traditional method. In summer months, the traditional method was not able to yield

a valid $E_{pm,GW}$ due to the large temperature uncertainty induced by the contamination from high solar background to the lidar signal. However, with the spectral proportion method, we are able to obtain the $E_{pm,GW}$ time series over the 5 years continuously (blue dots). We can identify that for observations with small temperature uncertainties (mostly southern hemisphere wintertime), the two methods give rise to mutually agreeable $E_{pm,GW}$. The utility of the spectral proportion method is that it is capable of producing reasonable $E_{pm,GW}$ when the temperature uncertainty is relatively large (in this case due to southern hemispheric summertime during McMurdo lidar observation). The values of $E_{pm,GW}$ ranges from ~0.20 (J/kg) to ~20 (J/kg) with maxima in winter and minima in summer. The peak values of $E_{pm,GW}$ are comparable from year to year with a smaller peak in 2014. The summertime $E_{pm,GW}$ (on the order of 10^{-1} (J/kg)) calculated via the spectral proportion method shows agreement with the results from previous studies from Rothera (67.5°S, 68.0°W), Antarctica and Haute-Provence Observatory (43.9°N, 5.7°E), in addition to other studies [e.g., Yamashita et al., 2009; Mze et al., 2014]. In the case of McMurdo lidar measurements, the spectral proportion method enables the calculation of a continuous $E_{pm,GW}$ all year via utilizing otherwise unusable datasets regardless of the relatively large amplitudes of the observational uncertainties. From our observation campaign from 2011 to 2015, the traditional method yields 107 $E_{pm,GW}$ results for observations. By employing the spectral proportion method, $E_{pm,GW}$ are produced from 216 observations. The detailed characteristic study of this 5-year $E_{pm,GW}$ was presented in Chu et al., [2018].

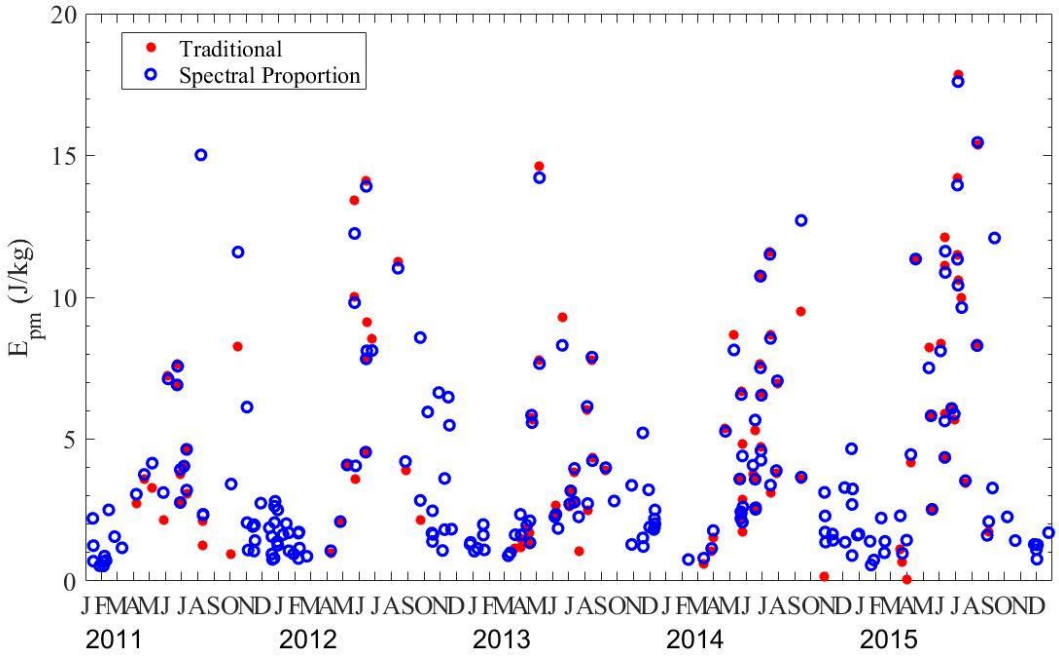

Figure 5.8. Five years of E_{pm} time series derived with spectral proportion method (blue) and traditional method (red). E_{pm} is averaged among 30 to 50 km. X-axis denotes years and months.

5.5. Summary

In Chapter 3, to seek the wave sources for the stratospheric gravity waves, we further investigated the strength variations of the stratospheric gravity waves at McMurdo. Potential energy density (E_{pm}) is regarded as the indication of the strength level of gravity wave activities. For the purpose of characterizing E_{pm} from McMurdo lidar campaign continuously through 5 years and making the most of the viable datasets, we have proposed and developed the spectral proportion method. Rigorous mathematical derivations and forward modeling validations of the spectral proportion method are provided here. Such a method to accurately calculate wave energy is especially suitable when observational uncertainties are large.

Utilizing McMurdo lidar wintertime observations, we confirm that when the observational uncertainties are small, the spectral proportion method shows

excellent agreement with the traditional method. However, when the observational uncertainties become relatively large, the traditional method will yield unrealistic results where negative $E_{pm,GW}$ are obtained even though the gravity wave signatures are apparent in the total temperature perturbations. In such cases, the spectral proportion method still yields reasonable $E_{pm,GW}$ results based on the derived gravity wave induced perturbations. The spectral proportion method is also validated via various forward modeling tests. Several parameters that might affect the algorithm's performance are discussed. Throughout the tests, the relative energy differences usually stay on the order of 10^{-2}, which is sufficiently small. We further present a 5-year $E_{pm,GW}$ time series calculated via the spectral proportion method from McMurdo lidar observations to illustrate the advantages of the new calculation approach. $E_{pm,GW}$ shows clear seasonal variations with winter maximum (~20 (J/Kg)) and summer minimum (~0.20 (J/Kg)).

Although we use the calculation of gravity wave potential energy density as the example to illustrate the spectral proportion method, this new methodology certainly applies to the calculation of many other atmospheric parameters such as wave kinetic energy density. When we detect a certain wave or a superposition of waves with parameters such as wind, temperature, and atmospheric density, we usually derive wave kinetic energy density with wind observations and wave potential energy density with either temperature or atmospheric density observations (refer to Appendix I). It is obvious to realize that the calculation of wave potential energy density comes from the mean square of the wave induced temperature or atmosphere density perturbations, and the calculation of wave kinetic energy density comes from the mean square of the wave induced wind perturbations. Employing spectral proportion method when performing such calculations can potentially save huge amount of observations with relatively large observational uncertainties. Generally speaking, if we have an observation of quantity $Q(t)$ and its

corresponding uncertainties $\delta Q(t)$ with respect to time t. Depending on what spectral portion of the signal perturbations we are interested in, we extract certain perturbations $Q^{'}(t)$ from $Q(t)$ utilizing background subtraction and spectrum analysis. We wish to calculate the quantity $\overline{(Q^{'}_{signal}(t))^2}$ from $\overline{(Q^{'}(t))^2}$ by excluding the influence of $\delta Q(t)$ on $Q^{'}(t)$, the spectral proportion method applies.

This chapter complements Chapter 3 as the spectral proportion method utilized in Chapter 3 is confirmed rigorously via mathematical derivations and forward modeling. The spectral proportion method applies to situations where we wish to extract accurate wave energy from observations of various qualities. Even with over-estimated observational errors, the spectral proportion method can still yield accurate wave energy according to the derived wave energy proportions with respect to all the energy detected. With increasing observational methods, the datasets available to scientists gradually begin to accumulate with various qualities, it is important that the spectral proportion method can perform remarkably to extract wave energy regardless of the quality of the data as long as the demanded signal is still recognizable in the observations.

CHAPTER VI

FAST GRAVITY WAVES (~1.5 HR) FROM THE STRATOSPHERE TO THE THERMOSPHERE OVER MCMURDO: CHARACTERIZATION AND POSSIBLE SOURCES

6.1. Introduction

An important discovery from McMurdo lidar observations is the fast gravity waves with ~1.5 h periods and long vertical wavelengths that clearly show up in the thermosphere-ionosphere Fe (TIFe) neutral layers during the mid-winter in Antarctica [Chu et al., 2011b]. The morphology of such fast gravity waves is shown in Figure 6.1a. Utilizing Fe layer observations, neutral temperatures can be retrieved (Figure 6.1b). As a matter of fact, temperature profiles form ~30 km to ~150 km can be derived from the Fe Boltzmann lidar observations with the occurrence of TIFe layers (Figure 6.1c). The periods, vertical wavelengths, and vertical phase speeds of gravity waves in the MLT are shown in Figure 6.1e and 6.1f.

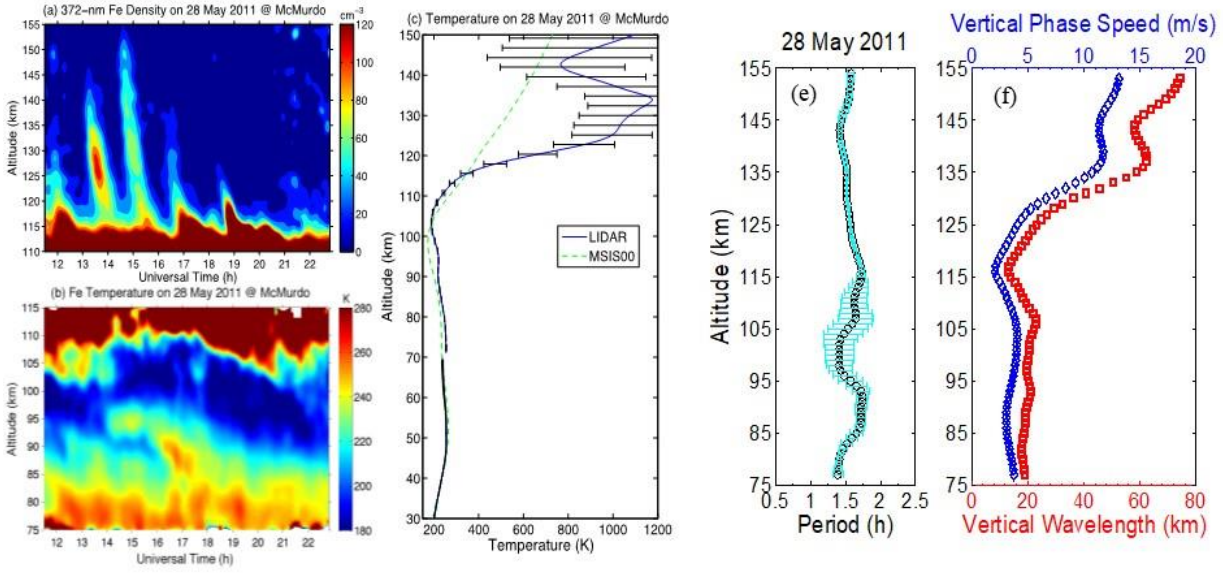

Figure 6.1. (a) Fast gravity waves with periods of ~1.5 h observed by Fe Boltzmann lidar on 28 May 2011 at McMurdo, Antarctica [Chu et al., 2011b]. (b) Temperatures

retrieved in the MLT. (c) Temperature profile from 30 km to 150 km. (e) ~1.5 hr gravity wave dominant periods in the MLT. (f) ~1.5 hr gravity wave vertical wavelengths and vertical phase speeds in the MLT.

The TIFe modeling of the event on 28 May 2011 [Yu, 2014; Chu and Yu, 2017] demonstrates the important roles played by such a fast gravity wave (refer to Figure 6.2). First, the gravity-wave-induced vertical shears of horizontal and vertical winds turn out to be a key factor in converging Fe^+ ions via Lorentzian force under polar magnetic and electric fields, and the converged Fe^+ ions are then neutralized to form neutral TIFe layers via recombination with electrons. Second, the gravity-wave-induced vertical winds are responsible for the observed TIFe layer shapes with repeated gravity wave patterns. Furthermore, the dissipation of (various) gravity waves may contribute to the observed temperature inversion layer in the thermosphere, because an initial investigation indicates that the effects induced by aurora particle precipitation alone may not be sufficient to account for all the temperature enhancement [Wu et al., AGU abstract, 2018]. To further develop the TIFe model for simulation of numerous TIFe cases observed in the last 8 years at McMurdo, it is crucial to characterize the fast gravity waves more thoroughly and implement more accurate wave parameters into the TIFe model.

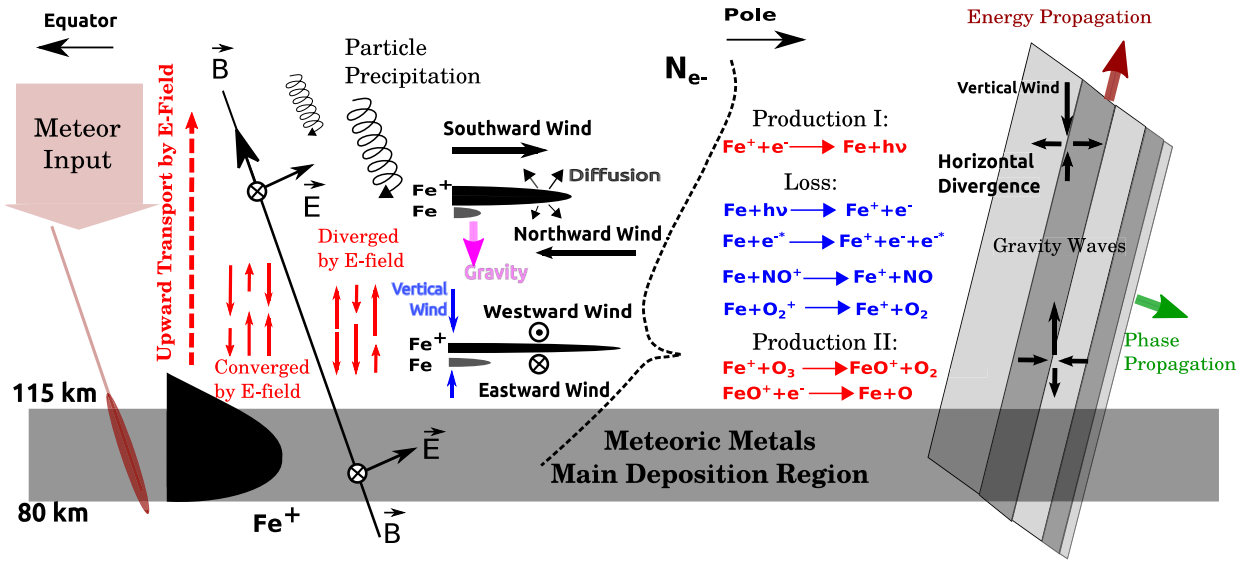

Figure 6.2. The roles played by the fast gravity waves in the formation of TIFe layers. This illustration was taken from Figure 1 of Chu and Yu [2017].

However, it is very challenging to characterize such gravity waves from near the surface all the way to the thermosphere, because 1) the observational data must have high resolutions (at least about half-hour resolution) and cover a very large altitude range, and 2) the wave analysis must be sufficiently sophisticated to extract the fast gravity waves out of many dominant waves (e.g., persistent gravity waves dominate the MLT region, while both planetary waves and long-period gravity waves dominate the stratosphere). We are fortunate to have accumulated abundant nights of high-quality Fe Boltzmann lidar data during winter, which not only show active TIFe layers but also enable the large altitude coverage (30 to ~170 km) with sufficient temporal and vertical resolutions. A focus of this chapter is to develop sophisticated wave analysis methods to properly extract the fast gravity waves out of complex dominant waves.

Chu et al. [2011b] characterized for the first time the fast gravity waves from 75 to 155 km altitude in the case of 28 May 2011, and then Chen [2016] further characterized this case down to 30 km and studied such fast gravity waves in more

cases. Nevertheless, none of the earlier studies touched on the issues of wave sources, and the methods used in Chen [2016] to extract the fast gravity waves were relatively primitive, i.e., through harmonic fittings to extract waves, rather than sophisticated wave spectral analysis and recognition [Chen and Chu, 2017]. These studies were done before the formulation of secondary gravity wave generation [Vadas et al., 2018]. Without the knowledge of secondary wave generation, waves with upward phase progression were removed, and the remaining downward phase progression waves were not examined in a way to detect this generation process [i.e., the perturbations were not multiplied by the square root of atmospheric density]. Now equipped with a suite of sophisticated wave analysis methods, we are in a good position to revisit the fast gravity waves for better characterization. More importantly, the new understandings developed in the last several years via journal papers [Zhao et al., 2017; Chu et al., 2018; Vadas et al., 2018; Becker ad Vadas, 2018; Vadas and Becker, 2018] provide new clues to search for the wave sources. For example, it is possible that the wide spectral range of the secondary gravity wave generation enables the excitation of 1.5 h waves from the primary wave dissipation in the stratosphere and lower mesosphere. Similar analysis methods and criteria developed in Vadas et al. [2018] can now be applied to study the fast gravity waves. Therefore, the last part of this dissertation is devoted to an investigation of these fast gravity waves, utilizing multiple data sources (Rayleigh temperatures, Fe temperatures, and TIFe layers). Besides characterizing these waves with better methods, we aim to search for their possible sources. These long-vertical-wavelength fast gravity waves represent a unique coupling process from the lower atmosphere to the thermosphere—an important aspect of SAIR.

6.2. Observational Data and Methodology

To study the fast gravity waves, various parameters measured by the McMurdo Fe Boltzmann lidar in different altitude ranges will be combined to cover a

large altitude range from ~30 km to ~150 km or higher. 1) The TIFe densities provide the highest resolution in the thermosphere (above ~110 or 115 km) as Fe temperatures derived from the tentative TIFe layers do not have sufficient resolutions. 2) In the MLT region, the Fe temperature provides the best data for characterizing gravity waves. 3) Below the MLT, both the Rayleigh temperatures and the atmospheric densities will be used to characterize gravity waves in the mesosphere and stratosphere. In this study we will handle all four types of lidar data. For the Rayleigh temperature, the temporal and vertical integration resolutions were chosen to be 0.5 h and 0.96 km, respectively. The temporal and vertical step resolutions were chosen to be 0.1 h and 0.96 km, respectively. The atmospheric density in the same region are retrieved with the same resolutions. For the Fe temperature, the step temporal and vertical resolutions were chosen to be 0.1 h and 0.1 km, respectively. For the Fe density, the step temporal and vertical resolutions were chosen to be 0.1 h and 0.1 km, respectively.

A major challenge in characterizing the fast gravity waves is that except in the TIFe layer altitudes, the 1.5-hr fast gravity waves are not the dominant waves in the MLT and in the stratosphere. They are even less dominant in the stratosphere because the planetary waves and the inertial-period gravity waves have much larger amplitudes than the fast gravity waves. As an example, in Figure 6.3, the raw Fe temperatures are obviously dominated by the 3 to 10 h persistent gravity waves in the MLT, although the higher-frequency ~1.5 h waves are certainly visible.

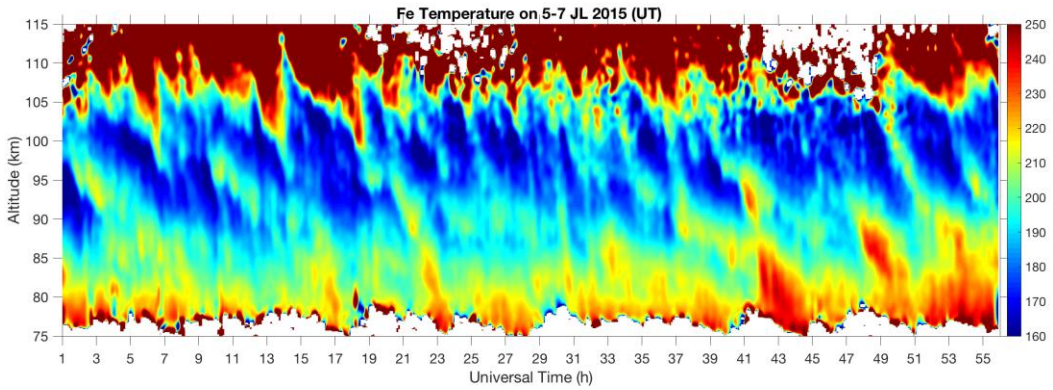

Figure 6.3. Fe temperatures in the MLT observed by the Fe Boltzmann lidar on 5-7 July 2015 at McMurdo, Antarctica.

Consequently, the methodology starts with removing the dominant planetary or inertial gravity waves from the temperature observation while preserving the fast gravity waves as much as possible. Here we use the observation in Figure 6.3 to illustrate the procedures employed to extract the fast gravity waves. The first step is to obtain the relative temperature perturbations by removing a temporal mean across the observational time span, and then dividing by the temporal mean at each altitude. The fast gravity waves become clearer in the relative temperature perturbations (Figure 6.4a). Then a 6th order Butterworth high-pass filter is used to remove the low frequency persistent gravity waves while keeping the shorter-period fast gravity waves. This process is demonstrated in Figure 6.4b. Figure 6.4c indicates the remaining low frequency gravity waves obtained via subtracting Figure 6.4b from 6.4a. Generally, the extraction of the ~1.5 hr fast gravity waves in the next sections follow the same procedures as above except for some details in the setup of filters. Furthermore, the consistent existence of the ~1.5 hr gravity waves in the MLT (Figure 6.4b) becomes another motivation for the analysis of such waves.

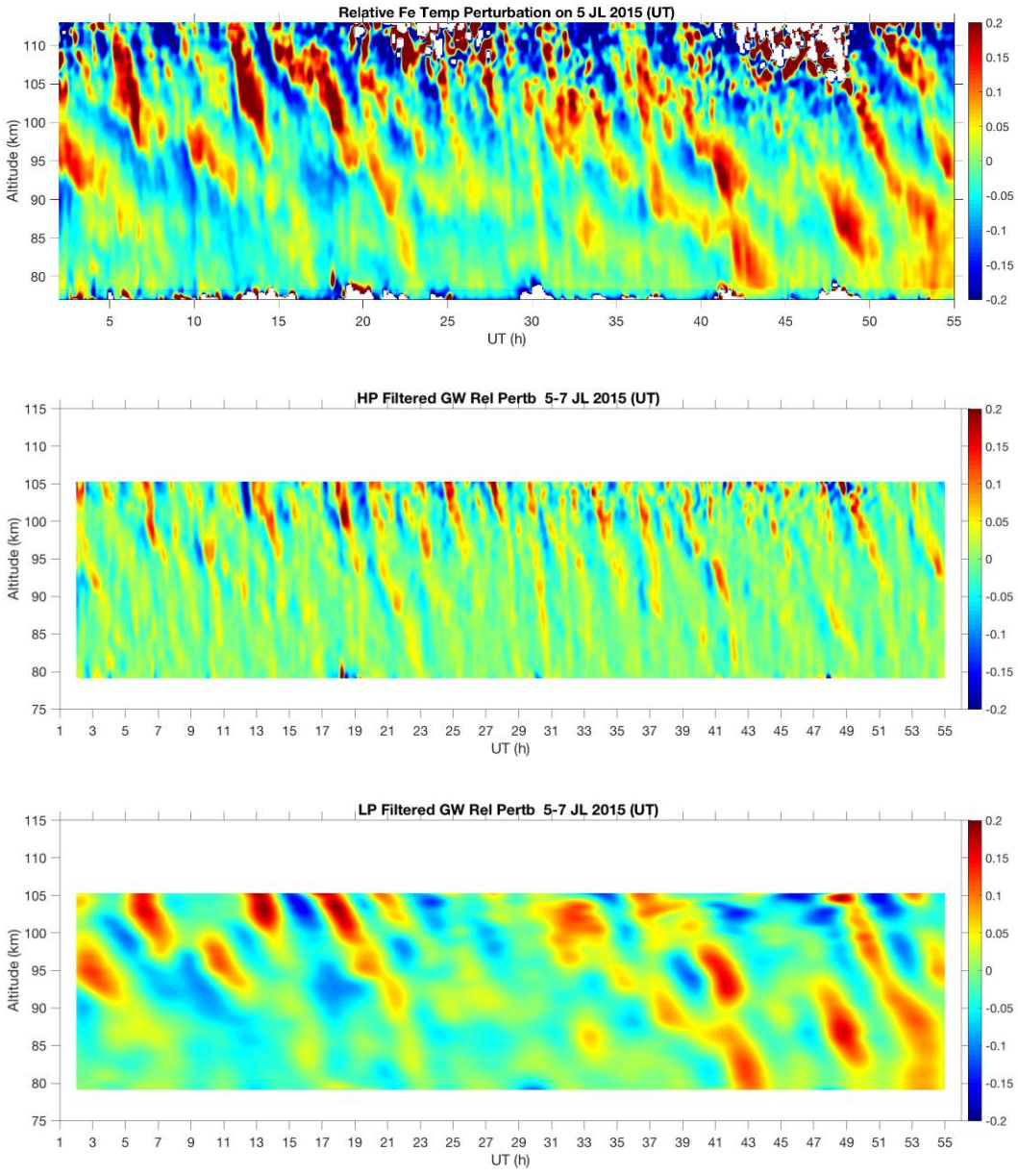

Figure 6.4. (a) Top panel, relative perturbations of Fe temperature derived from the raw data on 5-7 July 2015 at McMurdo. (b) Middle panel, high-pass filtered gravity wave induced relative temperature perturbations. (c) Bottom panel, remaining low frequency gravity wave induced relative temperature perturbations on the same days.

6.3. Case Study on 28 May 2011

6.3.1. Characterization of ~1.5 hr gravity waves from the stratosphere to the thermosphere

Here we analyze the ~1.5 hr gravity waves from lidar observed temperature, atmospheric density, and Fe density on 28 May 2011 from the McMurdo Fe Boltzmann lidar campaign. It is worth to note that the famous TIFe layer event occurred on this day, overlapping with a very strong auroral event and geomagnetic storm [Chu et al., 2011b; Chu and Yu, 2017].

In the Rayleigh region, we utilize both atmospheric density and temperature observations to derive the gravity waves of interest. Following the procedures described above, we apply the background subtraction and wave filtering process to the corresponding observations to extract the ~1.5 hr gravity waves. The background subtraction involves subtracting the temporal mean and dividing the perturbations by the temporal mean at every altitude. In the wave filtering process, we utilize a 6^{th} order high-pass filter with a cutoff frequency of 1/2 (1/h) in the time domain, and a band pass filter with a cutoff frequency of 1/10 (1/ km) and 1/50 (1/km) in the altitude domain. The extracted gravity waves are shown in Figure 6.5.

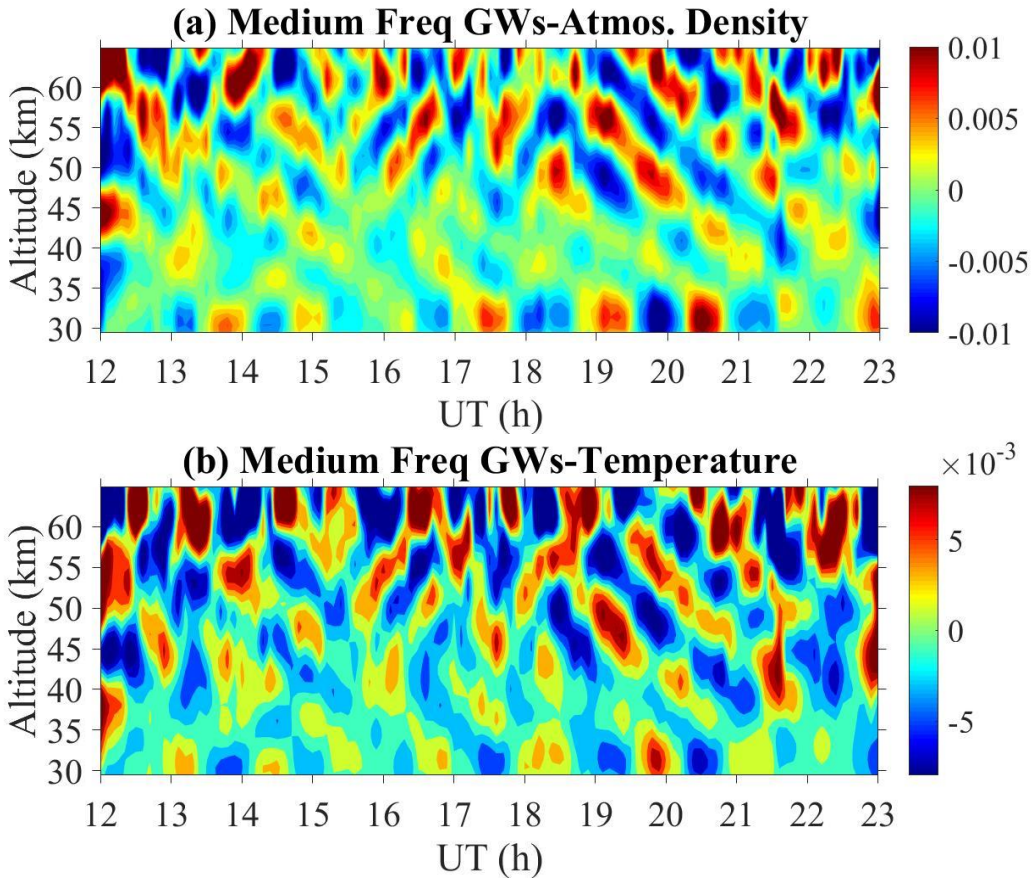

Figure 6.5. The extracted ~1.5 hr gravity wave relative perturbations from atmospheric density (a) and from temperature (b) observations in the Rayleigh region.

We notice the derived gravity wave perturbations from the temperature observation and from the atmospheric density observation have a phase difference of ~180° in the Rayleigh region. Intuitively, this can be predicted since wherever the temperature is higher, the air density should be lower. Now we confirm this relationship from the gravity wave polarization relations. When gravity wave vertical wavelength $\lambda_z \le 4\pi H$ ($\lambda_z \le 90$ km in the middle atmosphere), the Boussinesq limit occurs. In this limit, the gravity wave polarization relations are derived in Vadas [2013] as the following (Equation B13 and B14 in the paper),

$$\rho_p = \frac{-imN_B^2}{g(N_B^2 - \omega_I^{\,2})} p_p \tag{6.1}$$

$$T_p = \frac{imN_B^2}{g(N_B^2 - \omega_I^{\,2})} p_p \tag{6.2}$$

Thus, $T_p = -\rho_p$. As a result, we confirm that gravity wave perturbations derived from temperature perturbations and atmospheric density perturbations should be ~180° out of phase.

In order to investigate the potential connections of such ~1.5 hr gravity waves from the stratosphere to the MLT, we combine temperature observations in the Rayleigh region (the stratosphere) with temperature observations in the Fe temperature region (the MLT) and Fe density in the TIFe layer region (the thermosphere) on 28 May 2011. For Rayleigh temperatures and Fe temperatures, we follow the above procedures to extract the gravity wave of interest. After subtracting and dividing the temporal mean at every altitude, a 6th order Butterworth filter is applied to extract gravity waves with periods shorter than 2 hr and vertical wavelength longer than 10 km but shorter than 50 km. The ~1.5 hr gravity waves are finally illustrated in Figure 6.6 from Fe density and temperature observations. Note here the Fe temperature perturbations are scaled up by 10 times, and the Rayleigh and Fe temperature perturbations are scaled by the square root of the corresponding background atmospheric density to cancel out the exponential growth of the wave amplitude in altitude. To match the color bar in the range of gravity wave perturbations, the Fe density here (from ~113 km to ~155 km) is plotted as $\left[\log_{10}\left(\textit{absolute Fe density}\left(cm^{-3}\right)\right)-1\right]/2500$.

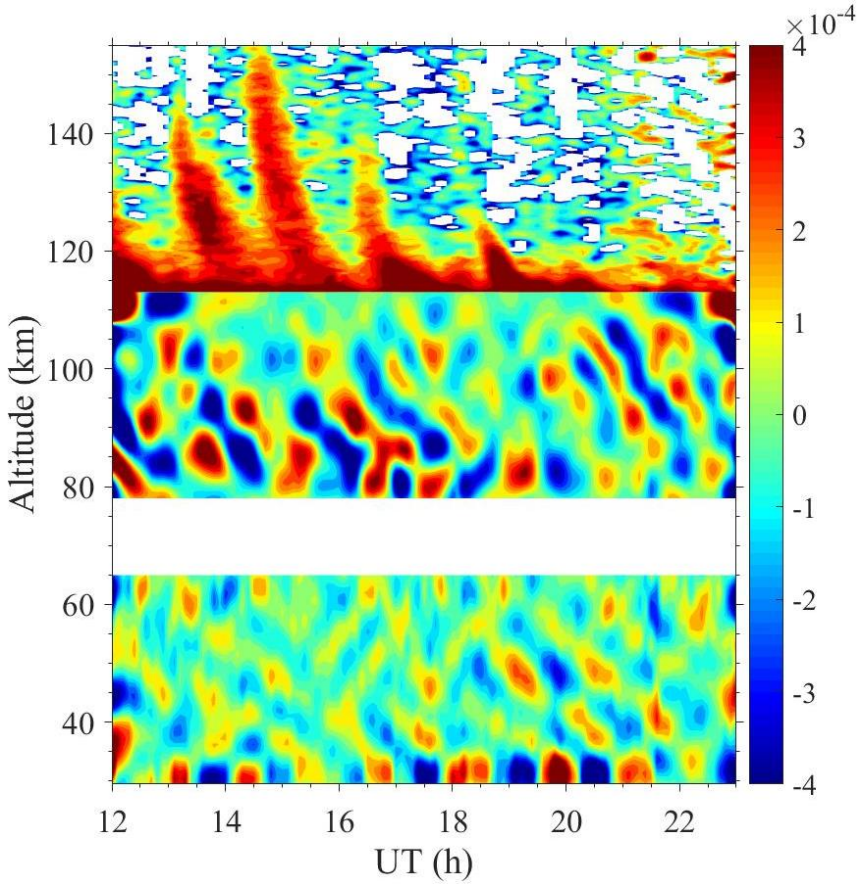

Figure 6.6. The derived ~1.5 hr gravity waves from the stratosphere to the thermosphere from Fe density (plotted as $\left[\log_{10}\left(\textit{absolute Fe density}\left(cm^{-3}\right)\right)-1\right]/2500$) and temperature observations.

We characterize the dominant gravity wave period at every altitude by applying 1DFFT on the gravity wave induced temperature relative perturbations in the Rayleigh region and Fe temperature region. Figure 6.7 shows the corresponding FFT results. The white dots indicate the dominant periods determined for every altitude. As is illustrated in both the time domain (Figure 6.6) and the frequency domain (Figure 6.7), such ~1.5 hr gravity waves exist from the stratosphere to the thermosphere consistently. If we exclude the upward-phase-progression gravity waves in the stratosphere and in the MLT, the obtained gravity wave perturbations are shown in Figure 6.8. We can visually identify certain consistent gravity wave phase lines from the stratosphere all the way up to the thermosphere (illustrated by

the black diamonds). Hence, there is a possibility that the ~1.5 hr fast gravity waves in the TIFe layer come directly from the lower atmosphere such as the troposphere.

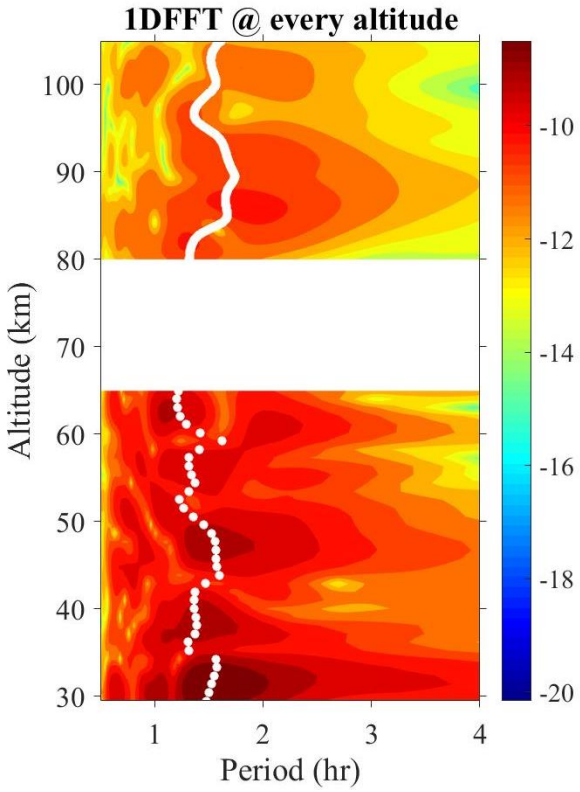

Figure 6.7. 1DFFT of gravity wave induced relative perturbations in the stratosphere and MLT. The white dots indicate the determined dominant periods.

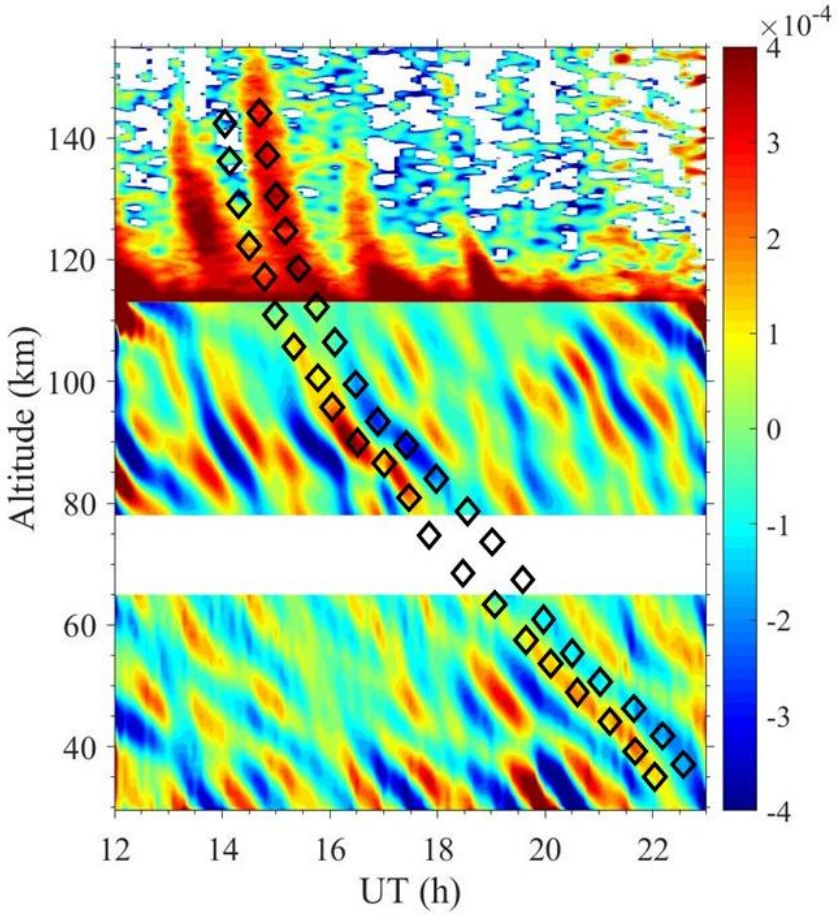

Figure 6.8. The derived ~1.5 hr downward-phase-progression gravity waves from the stratosphere to the thermosphere from Fe density and temperature observations (The plotting convention is the same as Figure 6.6).

6.3.2. Investigation of secondary gravity wave generation providing wave origin for the ~1.5 hr fast gravity waves in the TIFe layer

Currently, we cannot eliminate the possibility that the ~1.5 hr fast gravity waves propagate from the troposphere all the way to the thermosphere. However, based on the visually identified fishbone structure in the stratosphere, we discuss the possibility of the secondary gravity wave generation mechanism acting as the source of the ~1.5 hr fast gravity wave in the TIFe layer. The procedures in Chapter 4 were

followed in the analysis. Again, the gravity wave perturbations are scaled by the square root of the atmospheric density along the altitude. We estimate that the knee altitude is at around 44.80 km following criterion 1) in Chapter 4. Based on the procedures described previously in Chapter 4, we separate the removed gravity waves and the secondary gravity waves, as are shown in Figure 6.9. In Figure 6.9b, the gravity waves are severely dissipated around 40 to 45 km, thus satisfying criterion 2). Additionally, the phase lines do not become vertical approaching the knee altitude, thereby satisfying criterion 3). Only small amplitude gravity waves with upward phase progression occur above the knee altitude, therefore satisfying criterion 4). Based on Figure 6.9a and 6.6, the fishbone region is chosen to be from 14 to 17 h, and 33 to 60 km. The 2DFFT of the removed and secondary gravity waves are shown in Figure 6.10. The dominant secondary gravity wave parameters above the knee altitude are period =1.71±0.12 h, vertical wavelength=13.65±1.03 km, below the knee altitude are period=1.83±0.12 h, vertical wavelength=14.46±1.28 km. The dominant removed gravity wave parameters above the knee altitude are period =0.90±0.10 h, vertical wavelength=16.38±1.10 km, below the knee altitude are period=1.31±0.11 h, vertical wavelength=11.70±1.11 km. Since the periods of the secondary gravity waves below the knee altitude is very different from that of the removed gravity waves above the knee altitude, the secondary gravity waves below the knee altitude is not the continuation of the removed gravity waves above the knee altitude, satisfying criterion 5). And since the period of the secondary gravity wave above the knee altitude is very different from that of the removed gravity waves below the knee altitude, the secondary gravity wave above the knee altitude is not the continuation of the removed gravity wave below the knee altitude, satisfying criterion 6). More importantly, the gravity wave parameters are similar above and below the knee altitude, and their amplitude are within a factor of 2.5 of each other (1.58×10^{-9} above knee and 2.03×10^{-9} below knee), satisfying criterion 7). To confirm criterion 8), we

plotted the zonal and meridional winds from MERRA-2 in Figure 6.11. From the vertical wavelength of the secondary gravity waves, we calculate the intrinsic phase speed of the secondary gravity waves is ~ 46.85 m/s according to the approximation in Chapter 4. Since the intrinsic phase speed of the gravity waves are mostly larger than the wind speeds within the secondary gravity wave generation region, we conclude that for this case, the upward (downward) phase progression indicates downward (upward) propagation of gravity waves.

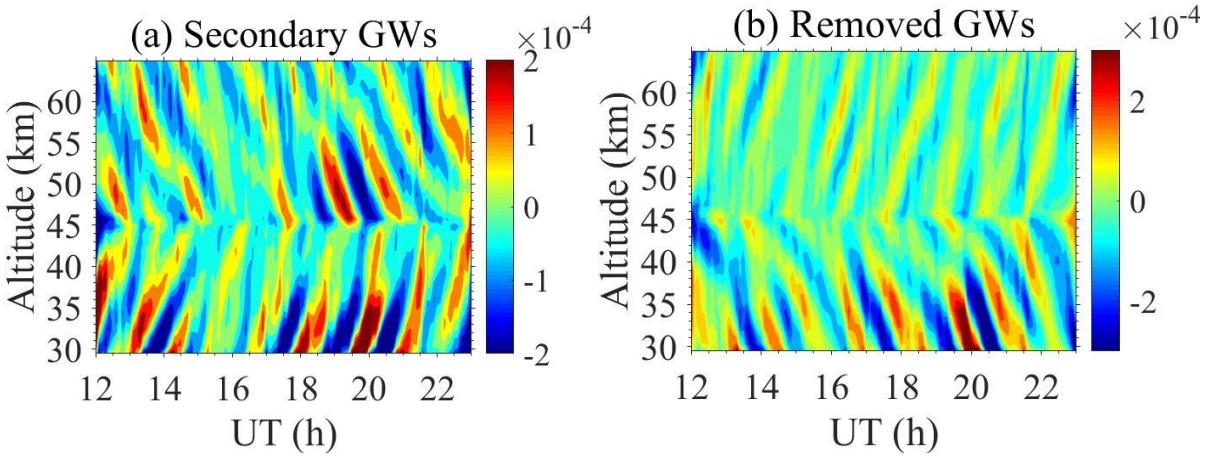

Figure 6.9. Derived secondary gravity waves (a) and removed gravity waves (b) in the stratosphere.

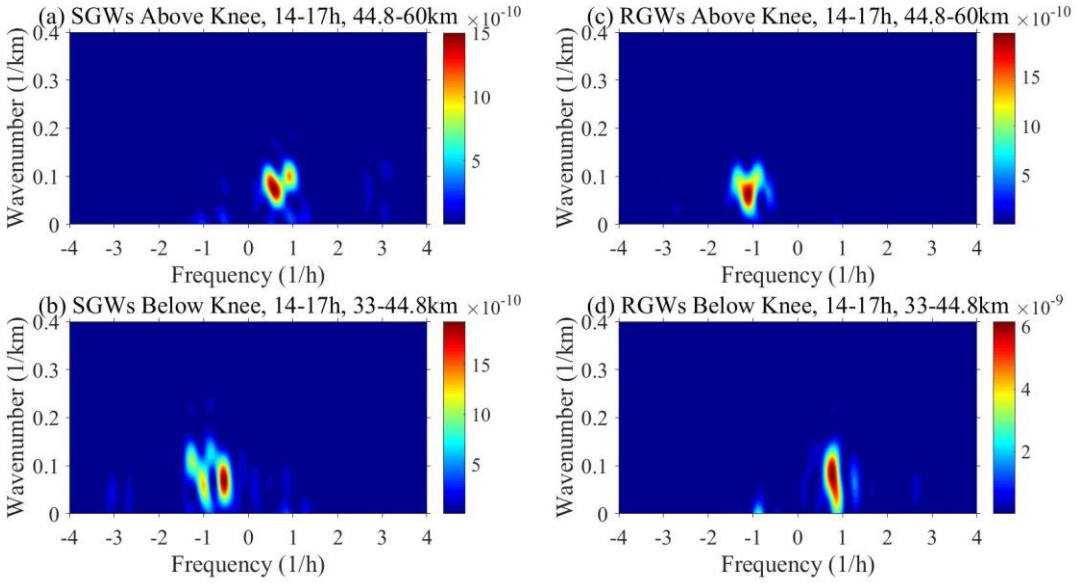

Figure 6.10. 2DFFT of secondary gravity waves (a, b) and removed gravity waves (c,d) above and below the knee altitude in the selected area.

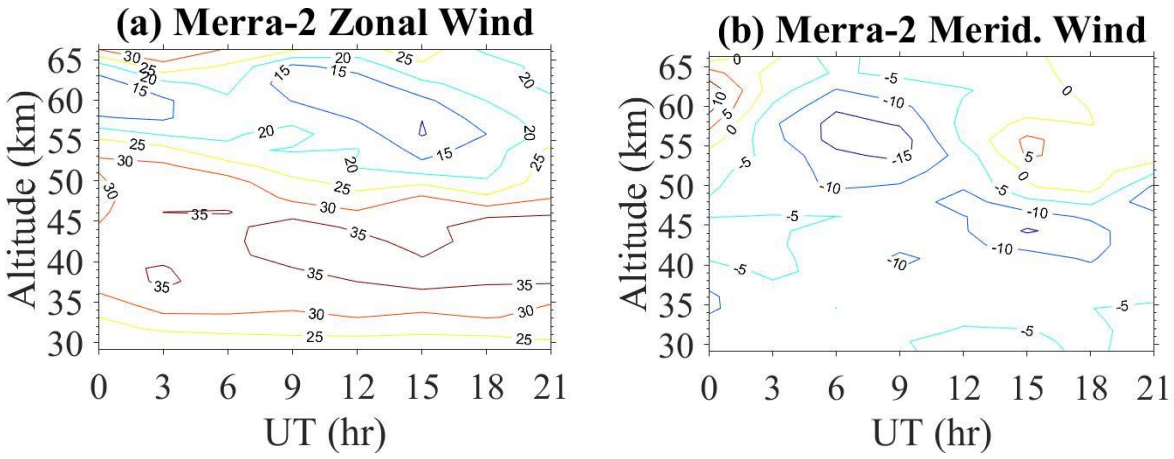

Figure 6.11. Zonal (a) and meridional (b) wind from Merra-2 in the stratosphere on 28 May 2011.

Thus, since criteria 1) to 8) are all satisfied, we conclude that the ~1.5 hr gravity waves may have been created by the secondary gravity wave generation

mechanism. Therefore, this process may provide the source for the 1.5 hr gravity waves in the TIFe layer.

However, there are still some potential issues for the secondary gravity wave generation theory to provide the source of the 1.5 hr fast gravity waves in the TIFe layer. We illustrate these problems in Figure 6.12. Even if in this case there were secondary gravity waves generated in the upper stratosphere, these waves should not be the waves observed in the TIFe layer at the same time because the newly generated waves need time to propagate into the MLT. And by the time they propagate into the MLT, it is also very likely that they have already left the field of view of McMurdo lidar. The 1.5 hr gravity waves we observed in the TIFe layer might come from secondary gravity waves from other locations, or simply from the lower atmosphere. We also note that secondary gravity waves usually have a broad wave spectrum [Vadas et al., 2018]. This means that even though these waves may be from the secondary gravity wave generation mechanism, the corresponding primary gravity waves might have considerably different periods. Therefore, it is clear that there is still work to do before reaching any conclusions. To further discuss the source of the 1.5 hr fast gravity waves in the TIFe layer, we may need to utilize results from gravity wave resolving models to fully trace the wave generation process and propagation into the thermosphere.

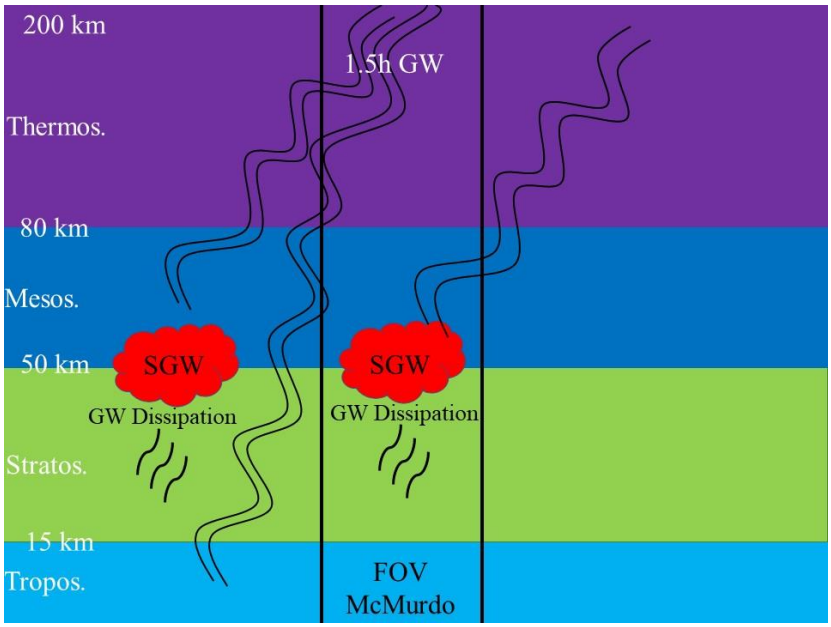

Figure 6.12. The potential issues of secondary gravity wave generation providing the source of the 1.5 hr gravity waves in the TIFe layer.

6.4. Summary

In the last part of this dissertation, we attempt to characterize the ~1.5 hr gravity waves from the stratosphere to the thermosphere due to their significant roles in the formation of Thermosphere-Ionosphere Fe layer. In the stratosphere, gravity waves derived from the atmospheric density and temperature have a ~180° phase difference, which can be confirmed via the gravity wave polarization relations. Although we cannot rule out the possibility that the ~1.5 hr gravity waves simply propagated from the troposphere to the thermosphere, due to the visually identified fishbone structures in the stratosphere, it is possible that the secondary gravity wave generation mechanism may account for their source. Indeed, we presented a case study to shed light on the possibility of the secondary gravity wave generation mechanism around the stratopause providing the source for the 1.5 hr fast gravity waves in the TIFe layer. However, more effort is needed to pinpoint the exact source of these waves.

CHAPTER VII

CONCLUSIONS AND OUTLOOK

The SAIR actively shapes the Earth's atmosphere to form a habitable environment for the life on Earth. Studying the dynamics and chemistry in the SAIR helps us understand the fundamental processes for a planet to harbor life. Within the SAIR, gravity waves are of vital importance to the transportation and distribution of momentum and energy in the middle and upper atmosphere. They are capable of affecting the atmospheric circulation, variability, composition, and thermal structure, etc. primarily via vertical coupling through different altitude regions in the atmosphere.

In this dissertation, via analyzing the invaluable observations at McMurdo, Antarctica from the Fe Boltzmann lidar, we attempt to establish an overall vertical picture of gravity wave coupling from near the surface to the thermosphere, contributing to people's understanding of the dynamics and chemistry in the SAIR. The dissertation starts from the investigation of the source of the recently discovered MLT persistent gravity waves. In order to testify if such waves come directly from the stratospheric dominant gravity waves, we perform a statistical study on the characterization of the dominant gravity waves in the stratosphere. Via comparing various gravity wave properties between the stratospheric gravity waves and the MLT persistent gravity waves, we conclude that the persistent gravity waves in the MLT do not come directly from the dominant gravity waves in the stratosphere. We continue to characterize the gravity wave potential energy density for the stratospheric gravity waves and seek their possible sources in terms of orographic wave source, critical level filtering, *in-situ* wave source, and Doppler shift effect. In exploring the source of the MLT persistent gravity waves, we invoke the secondary

gravity wave generation theory. Observational bases for such mechanism are provided. In the last part of this dissertation, we attempt to characterize the ~1.5 hr gravity waves from the stratosphere to the thermosphere and search the source for the 1.5 hr fast gravity waves in the TIFe layer. Secondary gravity wave generation is invoked again to account for the origin, in parallel with the possible lower atmosphere sources. Along the process, we have developed the spectral proportion method to calculate gravity wave potential energy density, which is also validated mathematically and numerically in this dissertation. In the following section, we provide the major conclusions and outlook along the track of the dissertation.

7.1. Major Conclusions and Outlook Throughout the Dissertation

Persistent gravity waves with periods of 3 to 10 h and vertical wavelengths of 20 to 30 km were discovered in the McMurdo MLT with Fe lidar observations [Chen et al., 2013, 2016; Chen and Chu, 2017]. In Chapter 2, to investigate whether the source of the MLT persistent gravity waves lies in the stratosphere, five years of stratospheric temperature data of the Fe lidar are used to characterize the gravity wave vertical wavelengths, periods, vertical phase speeds, frequency spectra, and vertical wavenumber spectra from 30–50 km altitudes at McMurdo, Antarctica. The seasonal distributions of the gravity wave vertical wavelengths, periods, and vertical phase speeds in summer, winter, and spring/fall are found obeying lognormal distributions. The monthly-mean vertical wavelengths and periods exhibit clear seasonal cycles with vertical wavelengths growing from summer ~5.5 km to winter ~8.5 km, and periods increasing from summer ~4.5 h to winter ~6 h. Statistically significant linear correlations are found between the monthly-mean vertical wavelengths/periods and the monthly-mean zonal wind velocities from 30–50 km. Such linear correlations are explained via the gravity wave dispersion relation and Doppler shift effect.

Assuming monthly-mean horizontal phase speeds are independent of month, the monthly-mean gravity wave horizontal wavelengths, intrinsic periods, horizontal group velocities, vertical group velocities, and propagation azimuth and elevation angles are estimated for stratospheric gravity waves. The monthly-mean horizontal wavelengths of the stratospheric gravity waves range from 350 to 450 km, which are substantially small compared to those of the MLT persistent gravity waves. Based primarily on this finding, we conclude that these dominant gravity waves in the stratosphere are not the waves that propagate into the MLT and become the observed persistent gravity waves directly.

In seeking the wave sources of the stratospheric gravity waves, 5 years of gravity wave potential energy per unit mass (E_{pm}) are characterized from McMurdo Fe Boltzmann lidar temperature observations in the stratosphere (30–50 km) in Chapter 3. The results indicate that stronger stratospheric gravity wave dissipation occurs in the winter than in the summer. Altitude-mean $\overline{E}_{pm}$ are found to obey lognormal distributions among different seasons. $\overline{E}_{pm}$ can vary significantly from observation to observation, but follow a repeated seasonal pattern with winter maxima and summer minima.

Utilizing the background wind information from MERRA-2, large values of $\overline{E}_{pm}$ are found to occur when McMurdo is well inside the polar vortex. Monthly-mean $\overline{E}_{pm}$ are found to be anti-correlated with wind rotation angles but positively correlated with the tropospheric and stratospheric winds. Overall, we conclude that the most probable sources for the stratospheric gravity waves are mainly the orographic gravity waves generated in the lower atmosphere (the troposphere, near the surface). When such waves propagate upward from the troposphere to the stratosphere, by the effect of critical level filtering due to the variation of background wind structure over seasons, the winter and summer asymmetry is observed in E_{pm}

in the stratosphere. *In-situ* wave generation and the Doppler shift effect in the stratosphere also play roles in explaining the winter E_{pm} variations.

Continuing to explore the origin of the MLT persistent gravity waves, secondary gravity wave generation is regarded as the most possible wave source in Chapter 4. This theory, established by Vadas et al. [2003, 2018], refers to the situation that the primary gravity waves (such as the orographic gravity waves generated by a downslope, eastward wind from the Trans-Antarctic Mountains to the west coast of the Ross Sea) propagate upward and then break, thereby generating local body forces (localized acceleration of the background mean flow). These generated body forces then excite larger scale secondary gravity waves, which possibly account for the source of the MLT persistent gravity waves with horizontal wavelengths of thousands of kilometers. From the McMurdo lidar campaign, we have identified two cases of secondary gravity wave generation which satisfy all 8 criteria required by the secondary gravity wave generation theory. Such scenarios are regarded to account for the source of the discovered large-scale MLT persistent gravity waves. Although we only enumerate 2 case studies in this dissertation, numerous secondary wave generation patterns can be identified from the long-lasting McMurdo campaign.

Along the process, we have developed the reliable spectral proportion method to accurately estimate wave energies such as atmospheric gravity wave kinetic and potential energy densities via rigorous mathematical formalism in Chapter 5. This new approach is especially suitable when the uncertainty of the observed parameter is considerably large. In the dissertation, temperature measurements from lidar observations are utilized to illustrate the calculation of gravity wave potential energy per unit mass (E_{pm}) via the spectral proportion method and the traditional method. The two methods give rise to nearly the same E_{pm} when the observed temperature uncertainties are small. However, when the observed temperature uncertainties become relatively large compared to the gravity-wave-induced perturbations, the

traditional method cannot yield realistic E_{pm} even though the gravity wave signatures are obvious in the temperature perturbations. In such cases, the spectral proportion method can robustly give reasonable E_{pm} based on the derived gravity wave induced perturbations. Rigorous mathematical development of this new methodology is provided and validated with intensive forward modeling. Several variables in the implementation that might affect the algorithm performance are studied in detail.

In Chapter 6, we attempt to characterize the ~1.5 hr gravity waves from the stratosphere all the way up to the thermosphere. The ~1.5 hr fast gravity waves play a significant role in the formation of TIFe layers. The vertical shears generated by the gravity wave induced horizontal and vertical winds effectively converge Fe^+ to form high density Fe^+ layer, where they recombine with electrons directly to form neutral Fe, hence generating the recently discovered TIFe layers. It is fair to say that the ~1.5 hr fast gravity wave directly shapes the morphology of the TIFe layer as the model simulations indicate [Chu and Yu, 2017]. In handling different datasets, we identified the ~180° phase difference between the gravity waves derived from atmospheric density and from temperature in the stratosphere. This relation can be confirmed via gravity wave polarization relations. In characterizing the ~1.5 hr gravity waves from the stratosphere to the thermosphere, both the lower atmosphere sources and secondary gravity wave generation are invoked to explain their consistent existence across the middle and upper atmosphere.

With all the previously stated conclusions, to move forward, there are still quite a few possible research directions. To be specific,

1) How does the stratospheric gravity wave potential energy change with respect to the positions of the polar vortex? We identify that in the McMurdo stratosphere, the gravity wave potential energy (E_{pm}) reach maximum when McMurdo is inside polar vortex edge. However, the fact that the polar vortex itself is gaining strength

during the winter should also be taken into consideration. If there is a lidar site sitting directly under the polar vortex edge all the time, we should detect the gravity waves there act even more fiercely as the polar vortex gaining strength. We are also not clear about the year to year variability of the formation and migration of the polar vortex. How does it affect the gravity wave potential energy in the McMurdo stratosphere?

2) We have performed case studies in order to provide the observational basis for the secondary gravity wave generation theory. To continue, statistical studies can be done to such events to summarize the characteristics of the gravity waves generated via such mechanism.
3) We identified lognormal distributions of the stratospheric gravity wave vertical wavelengths, ground-relative periods, vertical phase speeds, and $\overline{E}_{pm}$. What do they imply in terms of wave sources and dissipation? One possibility is that this could indicate secondary gravity wave generation in the stratosphere since Vadas et al., [2018] showed that secondary gravity waves follow lognormal distributions.
4) The source of the stratospheric gravity waves in the summer. Is it solely gravity waves from the lower atmosphere with significant eastward phase speeds?
5) The inter-annual variations of $\overline{E}_{pm}$ in the stratosphere and the causes.
6) Characterize periods, vertical wavelengths, $\overline{E}_{pm}$, etc. for the MLT gravity waves all year around via temperature and Fe density observations.
7) With the Na Doppler lidar measuring vertical winds at Arrival Heights, McMurdo, Antarctica, high frequency gravity waves as well as gravity wave momentum and energy flux may be inferred.
8) With advanced spectral analysis method such as 2D wavelet, gravity wave intermittency can be characterized.
9) Comprehensively explain the persistency of the MLT gravity waves and determine the source of the 1.5 hr fast gravity wave in the TIFe layer.

7.2. Gravity Wave Coupling at McMurdo, Antarctica: A Clearer Vertical Picture from Near the Surface to the Thermosphere

We have summarized the general picture of gravity wave source, propagation, dissipation, and regeneration in the Introduction (primarily in section 1.1.2). With the research in this dissertation, we have gradually developed an overall vertical picture for gravity waves above McMurdo, Antarctica. The schematics are depicted in Figure 7.1.

To be more illustrative, we describe the gravity wave vertical pictures for winter and summer separately. In the winter (Figure 7.1a), strong orographic gravity waves are generated in the troposphere owing to the elevated tropospheric winds. More orographic gravity waves (mostly westward propagating) can reach the stratospheric height because of less critical level filtering resulted from the strong and steady eastward background wind. With the formation of the polar vortex, the induced unbalanced flow possibly provides *in-situ* gravity wave generation in the winter stratosphere. Additionally, the Doppler shift effect can take place due to the enhanced winds in the winter stratosphere. All these mechanisms lead to the observed stronger gravity wave activity in the winter stratosphere. For the source of the MLT persistent gravity waves, we cannot rule out that certain gravity waves could propagate from the lower atmosphere into the MLT. But primarily due to the huge differences of the horizontal wavelengths of the dominant stratospheric gravity waves and the MLT persistent gravity waves, we conclude that the dominant gravity waves in the stratosphere are not the waves that propagate into the MLT and become the observed persistent gravity waves there. However, we illustrate that when strong primary gravity waves (such as orographic gravity waves) reach large amplitudes and dissipate or break in the upper stratosphere or lower mesosphere, the resulted secondary gravity wave generation can be regarded as a major potential source for the MLT persistent gravity waves. The ~1.5 hr fast gravity waves occurring in the

TIFe layer may have sources from both the lower atmosphere and secondary gravity wave generation.

In the summer (Figure 7.1b), the effect of critical level filtering is enhanced due to the background wind structures. Orographic gravity waves with near zero or slow phase speeds usually cannot propagate well into the stratosphere. Combined with the breakdown of the polar vortex which makes the *in-situ* wave generation unlikely, the stratospheric gravity wave activity is much weaker. Consequently, the explanation of secondary gravity wave generation providing source for the MLT persistent gravity waves also encounters difficulty. Strong primary gravity waves in the summer stratopause height are needed in order to generate intermittent body forces to excite secondary gravity waves. However, according to our analysis, the gravity wave activity in the summer stratosphere is much lower than that in the winter, which might be interpreted as unfavorable. The persistent gravity waves observed in the summer MLT over McMurdo, however, can be related to the primary non-orographic gravity waves. These waves could be generated in the troposphere by spontaneous emission during the breakdown of synoptic-scale Rossby waves and filtered by the prevailing wind system in the stratosphere. Following this explanation, only gravity waves with significant eastward phase speeds can propagate into the summer mesosphere.

In this dissertation, we aim to gradually assemble a clear vertical picture of gravity wave source, propagation, dissipation, and regeneration from the troposphere to the thermosphere. We have generally established such a blueprint via analyzing the measurements from the McMurdo lidar campaign. However, a comprehensive gravity wave coupling investigation that declares all the possible wave mechanisms is truly difficult and beyond the scope of this work. In this research, through characterizing possible wave connections across the whole atmosphere column, hopefully, we can humbly pave the way to human's ultimate understanding of the

sophisticated dynamics, physics, and chemistry in the Earth's space-atmosphere interaction region.

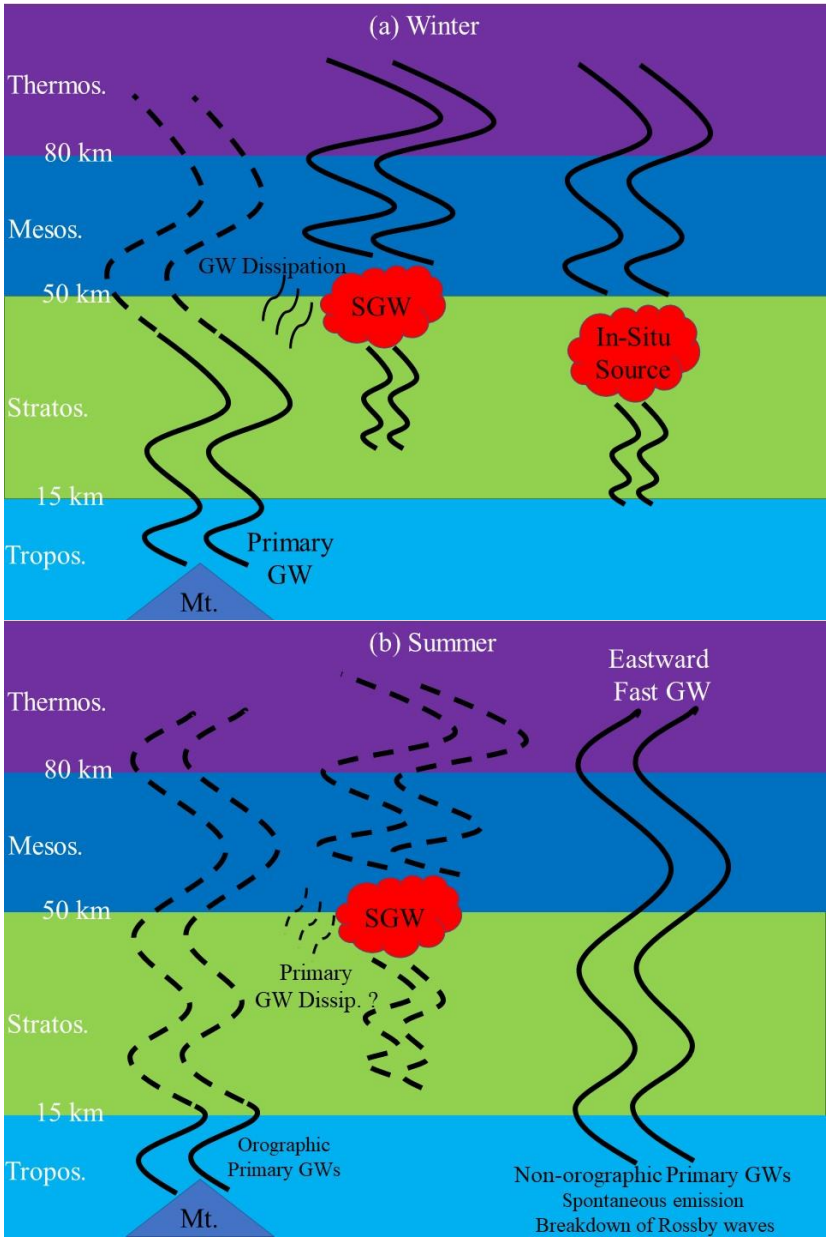

Figure 7.1. A cartoon representing the gravity wave vertical pictures we have summarized in this dissertation for winter (a) and for summer (b).

BIBLIOGRAPHY

Alexander, M. J., and L. Pfister (1995), Gravity wave momentum flux in the lower stratosphere over convection, Geophys. Res. Lett., 22(15), 2029–2032, doi:10.1029/95GL01984.

Alexander, M. J., and J. R. Holton (2004), On the spectrum of vertically propagating gravity waves generated by a transient heat source, Atmos. Chem. Phys., 4(4), 923–932, doi:10.5194/acp-4-923-2004.

Alexander, M. J., and C. Barnet (2007), Using Satellite Observations to Constrain Parameterizations of Gravity Wave Effects for Global Models, J. Atmos. Sci., 64, 1652–1665, doi: 10.1175/JAS3897.1.

Alexander, M. J., and H. Teitelbaum (2007), Observation and analysis of a large amplitude mountain wave event over the Antarctic peninsula, J. Geophys. Res., 112, D21103, doi:10.1029/2006JD008368.

Alexander, M. J., et al. (2008), Global estimates of gravity wave momentum flux from High Resolution Dynamics Limb Sounder observations, J. Geophys. Res., 113, D15S18, doi:10.1029/2007JD008807.

Alexander, S. P., A. R. Klekociuk, and T. Tsuda (2009), Gravity wave and orographic wave activity observed around the Antarctic and Arctic stratospheric vortices by the COSMIC GPS-RO satellite constellation, J. Geophys. Res., 114, D17103, doi:10.1029/2009JD011851.

Alexander, M. J., Geller, M., McLandress, C., Polavarapu, S., Preusse, P., Sassi, F., Sato, K., Eckermann, S., Ern, M., Hertzog, A., Kawatani, Y., Pulido, M., Shaw, T. A., Sigmond, M., Vincent, R. and Watanabe, S. (2010), Recent developments in gravity-wave effects in climate models and the global distribution of gravity-

wave momentum flux from observations and models. Q.J.R. Meteorol. Soc., 136: 1103–1124. doi:10.1002/qj.637

Alexander, S. P., A. R. Klekociuk, and D. J. Murphy (2011), Rayleigh lidar observations of gravity wave activity in the winter upper stratosphere and lower mesosphere above Davis, Antarctica (69°S, 78°E), J. Geophys. Res., 116, D13109, doi:10.1029/2010JD015164.

Alexander, M. J., and A. W. Grimsdell (2013), Seasonal cycle of orographic gravity wave occurrence above small islands in the Southern Hemisphere: Implications for effects on the general circulation, J. Geophys. Res. Atmos., 118, 11,589–11,599, doi:10.1002/2013JD020526.

Alexander, S. P., Murphy, D. J., and Klekociuk, A. R. (2013), High resolution VHF radar measurements of tropopause structure and variability at Davis, Antarctica (69° S, 78° E), Atmos. Chem. Phys., 13, 3121-3132, https://doi.org/10.5194/acp-13-3121-2013.

Allen, S. J., and R. A. Vincent (1995), Gravity wave activity in the lower atmosphere: Seasonal and latitudinal variations, J. Geophys. Res., 100(D1), 1327–1350, doi:10.1029/94JD02688.

Baumgaertner, A. J. G., and A. J. McDonald (2007), A gravity wave climatology for Antarctica compiled from Challenging Minisatellite Payload/Global Positioning System (CHAMP/GPS) radio occultations, J. Geophys. Res., 112, D05103, doi:10.1029/2006JD007504.

Becker, E. (2012), Dynamical control of the middle atmosphere, Space Sci. Rev., doi: 10.1007/s11214-011-9841-5.

Becker, E. (2017). Mean-Flow Effects of Thermal Tides in the Mesosphere and Lower Thermosphere. Journal of the Atmospheric Sciences, 74(6), 2043–2063. doi:10.1175/jas-d-16-0194.1.

Becker, E., and S. L. Vadas (2018), Secondary gravity waves in the winter mesosphere: Results from a high-resolution global circulation model, J. Geophys. Res. Atmos., 123, https://doi.org/10.1002/2017JD027460.

Beldon, C. L., & Mitchell, N. J. (2009). Gravity waves in the mesopause region observed by meteor radar, 2: Climatologies of gravity waves in the Antarctic and Arctic. Journal of Atmospheric and Solar-Terrestrial Physics, 71(8-9), 875-884. https://doi.org/10.1016/j.jastp.2009.03.009.

Bosilovich, M. et al. (2015), MERRA-2: Initial Evaluation of the Climate, NASA Tech. Rep. Series on Global Modeling and Data Assimilation, NASA/TM–2015-104606, Vol. 43.

Chanin, M.-L., and A. Hauchecorne (1981), Lidar observation of gravity and tidal waves in the stratosphere and mesosphere, J. Geophys. Res., 86(C10), 9715–9721, doi:10.1029/JC086iC10p09715.

Chen, C., X. Chu, A. J. McDonald, S. L. Vadas, Z. Yu, W. Fong and X. Lu (2013), Inertia-gravity waves in Antarctica: A case study using simultaneous lidar and radar measurements at McMurdo/Scott Base (77.8°S, 166.7°E), J. Geophys. Res. Atmos., 118, 2794–2808, doi:10.1002/jgrd.50318.

Chen, C., X. Chu, J. Zhao, B. R. Roberts, Z. Yu, W. Fong, X. Lu, and J. A. Smith (2016a), Lidar observations of persistent gravity waves with periods of 3–10 h in the Antarctic middle and upper atmosphere at McMurdo (77.83°S, 166.67°E), J. Geophys. Res. Space Physics, 121, 1483–1502, doi:10.1002/2015JA022127.

Chen, C., Lima, A., Smith, J., Huang, W., and Chu, X. (2016b). Investigation of vertical wavenumber and frequency spectra of gravity-wave-induced vertical wind and temperature perturbations. 18th Coherent Laser Radar Conference.

Chen, C. (2016c), Exploration of the Mystery of Polar Wave Dynamics with Lidar/Radar Observations and General Circulation Models & Development of

New Wave Analysis Methods. Aerospace Engineering Sciences Graduate Theses & Dissertations. 146. https://scholar.colorado.edu/asen_gradetds/146.

Chen, C., and X. Chu (2017), Two-dimensional Morlet wavelet transform and its application to wave recognition methodology of automatically extracting two-dimensional wave packets from lidar observations in Antarctica, J. Atmos. Solar-Terr. Phys., http://dx.doi.org/10.1016/j.jastp.2016.10.016.

Chu, X., W. Pan, G. C. Papen, C. S. Gardner, and J. A. Gelbwachs (2002), Fe Boltzmann temperature lidar: Design, error analysis, and initial results at the North and South Poles, Appl. Opt., 41(21), 4400–4410, doi:10.1364/AO.41.004400.

Chu, X., and G. C. Papen (2005), Resonance Fluorescence Lidar for Measurements of the Middle and Upper Atmosphere, in the book of Laser Remote Sensing, edited by T. Fujii and T. Fukuchi, published by CRC Press, Taylor & Francis Group, ISBN: 0-8247-4256-7, page 179-432.

Chu, X., P. J. Espy, G. J. Nott, J. C. Diettrich, and C. S. Gardner (2006), Polar mesospheric clouds observed by an iron Boltzmann lidar at Rothera (67.5°S, 68.0°W), Antarctica from 2002 to 2005: Properties and implications, J. Geophys. Res., 111, D20213, doi:10.1029/2006JD007086.

Chu, X., C. Yamashita, P. J. Espy, G. J. Nott, E. J. Jensen, H.-L. Liu, W. Huang, and J. P. Thayer (2009), Responses of polar mesospheric cloud brightness to stratospheric gravity waves at the South Pole and Rothera, Antarctica, J. Atmos. Sol. Terr. Phys., 71, 434–445, doi:10.1016/j.jastp.2008.10.002.

Chu, X., W. Huang, W. Fong, Z. Yu, Z. Wang, J. A. Smith, and C. S. Gardner (2011a), First lidar observations of polar mesospheric clouds and Fe temperatures at McMurdo (77.8°S, 166.7°E), Antarctica, Geophys. Res. Lett., 38, L16810, doi:10.1029/2011GL048373.

Chu, X., Z. Yu, C. S. Gardner, C. Chen, and W. Fong (2011b), Lidar observations of neutral Fe layers and fast gravity waves in the thermosphere (110–155 km) at

McMurdo (77.8°S, 166.7°E), Antarctica, Geophys. Res. Lett., 38, L23807, doi:10.1029/2011GL050016.

Chu, X., Z. Yu, W. Fong, C. Chen, J. Zhao, I. F. Barry, J. A. Smith, X. Lu, W. Huang, and C. S. Gardner (2016), From Antarctica lidar discoveries to OASIS exploration, EPJ Web of Conferences, 119, 12001, doi: http://dx.doi.org/10.1051/epjconf/201611912001, Proceedings of the 27th International Laser Radar Conference, New York, July 2015.

Chu, X., and Z., Yu (2017), Formation mechanisms of neutral Fe layers in the thermosphere at Antarctica studied with a thermosphere-ionosphere Fe/Fe+ (TIFe) model, J. Geophys. Res. Space Physics, 122, doi:10.1002/2016JA023773.

Chu, X., J. Zhao, X. Lu, V.L. Harvey, R.M. Jones, E. Becker, C. Chen, W. Fong, Z. Yu, B.R. Roberts, and A. Dörnbrack (2018), Lidar observations of stratospheric gravity waves from 2011 to 2015 at McMurdo (77.84° S, 166.69° E), Antarctica: Part II. Potential energy densities, lognormal distributions, and seasonal variations, J. Geophys. Res., under revision.

Collins, R. L., A. Nomura, and C. S. Gardner (1994), Gravity waves in the upper mesosphere over Antarctica: Lidar observations at the South Pole and Syowa, J. Geophys. Res., 99(D3), 5475–5485, doi:10.1029/93JD03276.

Collins, R. L., and C. S. Gardner (1995), Gravity wave activity in the stratosphere and mesosphere at the South Pole, Adv. Space Res., 16(5), 81–90, doi:10.1016/0273-1177(95)00175-E.

Collins, R. L., T. J. Hallinan, R. W. Smith, and G. Hernandez (1996), Lidar observations of a large high-altitude sporadic Na layer during active aurora, Geophys. Res. Lett., 23(24), 3655–3658, doi:10.1029/96GL03337.

Dee, D. P., et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc., 137, 553–597, doi:10.1002/qj.828.

Dewan, E. M., and N. Grossbard (2000), Power spectral artifacts in published balloon data and implications regarding saturated gravity wave theories, J. Geophys. Res., 105(D4), 4667–4683, doi:10.1029/1999JD901108.

Duck, T. J., J. A. Whiteway, and A. I. Carswell (1998), Lidar observations of gravity wave activity and arctic stratospheric vortex core warming, Geophys. Res. Lett., 25, 2813–2816, doi:10.1029/98GL02113.

Duck, T. J., J. A. Whiteway, and A. I. Carswell (2001), The gravity wave–Arctic stratospheric vortex interaction. J. Atmos. Sci., 58, 3581–3596, doi: 10.1175/1520-0469(2001)058<3581:TGWASV>2.0.CO;2.

Dörnbrack, A., T. Gerz, and U. Schumann (1995), Turbulent breaking of overturning gravity waves below a critical level, Appl. Sci. Res., 54(3), 163-176, doi: 10.1007/BF00849114.

Dörnbrack, A., S. Gisinger and B. Kaifler (2017), On the interpretation of gravity wave measurements by ground-based lidars, Atmosphere, 8, 49, doi:10.3390/atmos8030049.

Dowdy, A. J., R. A. Vincent, M. Tsutsumi, K. Igarashi, Y. Murayama, W. Singer, and D. J. Murphy (2007), Polar mesosphere and lower thermosphere dynamics: 1. Mean wind and gravity wave climatologies, J. Geophys. Res., 112, D17104, doi:10.1029/2006JD008126.

Eckermann, S. D., Hirota, I., & Hocking, W. K. (1995). Gravity wave and equatorial wave morphology of the stratosphere derived from long-term rocket soundings. Quarterly Journal of the Royal Meteorological Society, 121(521), 149–186. doi:10.1002/qj.49712152108.

Eckermann, S., and P. Preusse (1999), Global Measurements of Stratospheric Mountain Waves from Space, Science, 286(5444), 1534-1537, doi:10.1126/science.286.5444.1534.

Eckermann S. D., D. Broutman, J. Ma, J. D. Doyle, P-D. Pautet, M. J. Taylor, K. Bossert, B. P. Williams, D. C. Fritts, and R. B. Smith, (2016), Dynamics of Orographic Gravity Waves Observed in the Mesosphere over the Auckland Islands during the Deep Propagating Gravity Wave Experiment (DEEPWAVE), J. Atmos. Sci., 73, 3855–3876, doi: 10.1175/JAS-D-16-0059.1.

Eliassen, A., and E. Palm, 1961: On the transfer of energy in stationary mountain waves. Geofys. Publ., 22, 1–23.

Espy, P. J., G. O. L. Jones, G. R. Swenson, J. Tang, and M. J. Taylor (2004), Seasonal variations of the gravity wave momentum flux in the Antarctic mesosphere and lower thermosphere, J. Geophys. Res., 109, D23109, doi:10.1029/2003JD004446.

Espy, P. J., R. E. Hibbins, G. R. Swenson, J. Tang, M. J. Taylor, D. M. Riggin, and D. C. Fritts (2006), Regional variations of mesospheric gravity-wave momentum flux over Antarctica, Ann. Geophys., 24, 81–88, doi:10.5194/angeo-24-81-2006.

Fong, W., X. Lu, X. Chu, T. J. Fuller-Rowell, Z. Yu, B. R. Roberts, C. Chen, C. S. Gardner, and A. J. McDonald (2014), Winter temperature tides from 30 to 110 km at McMurdo (77.8°S, 166.7°E), Antarctica: Lidar observations and comparisons with WAM, J. Geophys. Res. Atmos., 119, 2846–2863, doi:10.1002/2013JD020784.

Fong, W., Chu, X., Lu, X., Chen, C., Fuller-Rowell, T. J., Codrescu, M. and Richmond, A. D. (2015), Lidar and CTIPe model studies of the fast amplitude growth with altitude of the diurnal temperature "tides" in the Antarctic winter lower thermosphere and dependence on geomagnetic activity. Geophys. Res. Lett., 42: 697–704. doi: 10.1002/2014GL062784.

Forbes, J. M. (2013). Tidal and Planetary Waves. Geophysical Monograph Series, 67–87. doi:10.1029/gm087p0067.

Friedman, J. S., and X. Chu (2007), Nocturnal temperature structure in the mesopause region over the Arecibo Observatory (18.35°N, 66.75°W): Seasonal variations, J. Geophys. Res., 112, D14107, doi: 10.1029/2006JD008220.

Friedman, J. S., X. Chu, C. Brum and X. Lu (2013), Observation of a thermospheric descending layer of neutral K over Arecibo, Journal of Atmospheric and Solar-Terrestrial Physics, doi: 10.1016/j.jastp.2013.03.002.

Fritts, D. C., and Z. Luo (1992), Gravity wave excitation by geostrophic adjustment of the jet stream. Part 1: Two-dimensional forcing, J. Atmos. Sci., 49, 681–697.

Fritts, D. C., and M. J. Alexander (2003), Gravity wave dynamics and effects in the middle atmosphere, Rev. Geophys., 41, 1003, doi:10.1029/2001RG000106, 1.

Fritts, D. C., S. L. Vadas, K. Wan, and J. A. Werne (2006), Mean and variable forcing of the middle atmosphere by gravity waves, J. Atmos. Sol.-Terr. Phys., 68, 247265, doi: 10.1016/j.jastp.2005.04.010.

Fritts, D. C., and M. J. Alexander (2012), Correction to "Gravity wave dynamics and effects in the middle atmosphere," Rev. Geophys., 50, RG3004, doi:10.1029/2012RG000409.

Gao, Q., X. Chu, X. Xue, X. Dou, T. Chen, and J. Chen (2015), Lidar observations of thermospheric Na layers up to 170 km with a descending tidal phase at Lijiang (26.7N, 100.0E), China, J. Geophys. Res. Space Physics, 120, doi: 10.1002/2015JA021808.

Garcia, R. R., and S. Solomon (1985), The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere, J. Geophys. Res., 90(D2), 3850–3868, doi: 10.1029/JD090iD02p03850.

Garcia, R. R. (1989), Dynamics, radiation and photochemistry in the mesosphere — implications for the formation of noctilucent clouds, J. Geophys. Res., 94 (D12), 14,605 – 14,615, doi: 10.1029/JD094iD12p14605.

Garcia, R.R. and B.A. Boville (1994), "Downward control" of the mean meridional circulation and temperature distribution of the polar winter stratosphere. J. Atmos. Sci., 51, 2238–2245, doi:10.1175/1520-0469(1994)051<2238:COTMMC>2.0.CO;2.

Garcia, R. R., M. López-Puertas, B. Funke, D. R. Marsh, D. E. Kinnison, A. K. Smith, and F. González-Galindo (2014), On the distribution of CO2 and CO in the mesosphere and lower thermosphere, J. Geophys. Res. Atmos., 119, 5700–5718, doi:10.1002/2013JD021208.

Gardner, C. S., D. C. Senfit, T. J. Beatty, R. E. Bills, and C. A. Hostetler (1989), Rayleigh and sodium lidar techniques for measuring middleatmospheric density, temperature and wind perturbations and their spectra, in World Ionosphere/Thermosphere Study Handbook, vol. 2,edited by C. H. Liu and B. Edwards, pp. 148–187, Int. Congr. of Sci. Unions, Urbana, Ill.

Gardner, C. S., X. Tao, and G. C. Papen (1995), Simultaneous lidar observations of vertical wind, temperature, and density profiles in the upper mesosphere: Evidence for nonseparability of atmospheric perturbation spectra, Geophys. Res. Lett., 22, 2877-2880.

Gardner, C.S. (1996), Testing theories of atmospheric gravity wave saturation and dissipation. J. Atmos. Terr. Phys. 58: 1575–1589.

Gardner, C.S., Y. Zhao, and A. Z. Liu (2002), Atmospheric stability and gravity wave dissipation in the mesopause region, Journal of Atmospheric and Solar-Terrestrial Physics, Volume 64, Issues 8–11, Pages 923-929, https://doi.org/10.1016/S1364-6826(02)00047-0.

Gelbwachs J. A. (1994), Iron Boltzmann factor LIDAR: proposed new remote-sensing technique for mesospheric temperature, Appl. Opt. 33, 7151-7156.

Geller, M.A., M.J. Alexander, P.T. Love, J. Bacmeister, M. Ern, A. Hertzog, E. Manzini, P. Preusse, K. Sato, A.A. Scaife, and T. Zhou (2013), A comparison between gravity wave momentum fluxes in observations and climate models. J. Climate, 26, 6383–6405, doi: 10.1175/JCLI-D-12-00545.1.

Godin, O. A., and N. A. Zabotin (2016), Resonance vibrations of the Ross Ice Shelf and observations of persistent atmospheric waves, J. Geophys. Res. Space Physics, 121, 10,157–10,171, doi:10.1002/2016JA023226.

Gong, J., D. L. Wu, and S. D. Eckermann, (2012), Gravity wave variances and propagation derived from AIRS radiances, Atmos. Chem. Phys., 12, 1701-1720, doi:10.5194/acp-12-1701-2012.

Hamilton, K. (1991). Climatological statistics of stratospheric inertia-gravity waves deduced from historical rocketsonde wind and temperature data. Journal of Geophysical Research, 96(D11), 20831. doi:10.1029/91jd02188.

Hamilton, K., R. J. Wilson, and R. S. Hemler (1999), Middle atmosphere simulated with high vertical and horizontal resolution versions of a GCM: Improvements in the cold pole bias and generation of a QBO-like oscillation in the tropics. J. Atmos. Sci., 56, 3829–3846, doi:10.1175/1520-0469(1999)056,3829: MASWHV.2.0.CO;2.

Harvey, V. L., R. B. Pierce, T. D. Fairlie, and M. H. Hitchman (2002), A climatology of stratospheric polar vortices and anticyclones, J. Geophys. Res., 107(D20), 4442, doi:10.1029/2001JD001471.

Hauchecorne, A., M. L. Chanin, P. Keckhut, and D. Nedeljkovic (1992), Lidar monitoring of the temperature in the middle and lower atmosphere, Appl. Phys. B, 55(1), 29–34, doi:10.1007/BF00348609

Haynes, P. (2005), Stratosphere dynamics, Annu. Rev. Fluid Mech. 2005, 37(1), 263–293, doi: 10.1146/annurev.fluid.37.061903.175710.

Hertzog, A., P. Cocquerez, R. Guilbon, J. Valdivia, S. Venel, C. Basdevant, G. Boccara, J. Bordereau, B. Brioit, F. Vial, A. Cardonne, A. Ravissot, and É. Schmitt (2007), Stratéole/Vorcore—long-duration, superpressure balloons to study the Antarctic lower stratosphere during the 2005 winter. J. Atmos. Oceanic Technol., 24, 2048–2061, doi:10.1175/2007JTECHA948.1.

Hertzog, A., G. Boccara, R. A. Vincent, F. Vial, and P. Cocquerez (2008), Estimation of gravity wave momentum flux and phase speeds from quasi-Lagrangian stratospheric balloon flights. Part II: Results from the VORCORE campaign in Antarctica, J. Atmos. Sci., 65, doi:10.1175/2008JAS2710.1.

Hertzog, A., M.J. Alexander, and R. Plougonven (2012), On the intermittency of gravity wave momentum flux in the stratosphere. J. Atmos. Sci., 69, 3433–3448, doi: 10.1175/JAS-D-12-09.1.

Hibbins, R. E., P. J. Espy, M. J. Jarvis, D. M. Riggin, and D. C. Fritts (2007), A climatology of tides and gravity wave variance in the MLT above Rothera, Antarctica obtained by MF radar, J. Atmos. Sol.-Terr. Phys., 69(4–5), 578–588, doi:10.1016/j.jastp.2006.10.009.

Hines, C. O. (1960), Internal atmospheric gravity waves at ionospheric heights, Can. J. Phys., 38, 1441, doi:10.1139/p60-150.

Hines, C. O. (1964), Minimum vertical scale sizes in the wind structure above 100 kilometers, J. Geophys. Res., 69(13), 2847–2848, doi:10.1029/JZ069i013p02847.

Hines, C. O. (1974), The Upper Atmosphere in Motion, pp. 13–344, AGU, Washington, D. C., doi:10.1029/GM018p0014.

Hirota, I. (1984). Climatology of gravity waves in the middle atmosphere. Journal of Atmospheric and Terrestrial Physics, 46(9), 767–773. doi:10.1016/0021-9169(84)90057-6.

Hirota, I., & Niki, T. (1985). A Statistical Study of Inertia-Gravity Waves in the Middle Atmosphere. Journal of the Meteorological Society of Japan. Ser. II, 63(6), 1055–1066. doi:10.2151/jmsj1965.63.6_1055.

Hitchman, M. H., J. C. Gille, C. D. Rodgers, and G. Brasseur (1989), The separated polar winter stratopause: A gravity wave driven climatological feature. J. Atmos. Sci., 46, 410–422, doi: 10.1175/1520-0469(1989)046<0410:TSPWSA>2.0.CO;2.

Hocke, K., and K. Schlegel (1996), Areview of atmospheric gravity waves and traveling ionospheric disturbances, Ann. Geophys., 14, 917–940, doi: 10.1007/s00585-996-0917-6.

Hoffmann, L., and M. J. Alexander (2009), Retrieval of stratospheric temperatures from Atmospheric Infrared Sounder radiance measurements for gravity wave studies, J. Geophys. Res., 114, D07105, doi:10.1029/2008JD011241.

Hoffmann, P., E. Becker, W. Singer, and M. Placke (2010), Seasonal variation of mesospheric waves at northern middle and high latitudes, J. Atmos. Sol.-Terr. Phys., 72, 1068-1079.

Hoffmann, L., X. Xue, and M. J. Alexander (2013), A global view of stratospheric gravity wave hotspots located with Atmospheric Infrared Sounder observations, J. Geophys. Res. Atmos., 118, 416–434, doi:10.1029/2012JD018658.

Hoffmann L., M. J. Alexander, C. Clerbaux, A. W. Grimsdell, C. I. Meyer, T. Rößler, and B. Tournier (2014), Intercomparison of stratospheric gravity wave observations with AIRS and IASI, Atmospheric Measurement Techniques, 7(12), 4517—4537, doi:10.5194/amt-7-4517-2014.

Hoffmann, L., A. W. Grimsdell, and M. J. Alexander (2016), Stratospheric gravity waves at Southern Hemisphere orographic hotspots: 2003–2014 AIRS/Aqua observations, Atmos. Chem. Phys., 16, 9381-9397, doi:10.5194/acp-16-9381-2016.

Holton, J. R. (1982), The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere. J. Atmos. Sci., 39, 791–799, https://doi.org/10.1175/1520-0469(1982)039<0791:TROGWI>2.0.CO;2.

Holton, J. R. (1983), The influence of gravity wave breaking on the general circulation of the middle atmosphere. J. Atmos. Sci., 40, 2497–2507, doi:10.1175/1520-0469(1983)040<2497:TIOGWB>2.0.CO;2.

Holton, J. R. (2004), An Introduction to Dynamic Meteorology, Academic Press, Volume 88 of International Geophysics.

Hostetler, C. A., and C. S. Gardner (1994), Observations of horizontal and vertical wave number spectra of gravity wave motions in the stratosphere and mesosphere over the mid-Pacific, J. Geophys. Res., 99(D1), 1283–1302, doi:10.1029/93JD02927.

Kaifler, B., F.-J. Lübken, J. Höffner, R. J. Morris, and T. P. Viehl (2015), Lidar observations of gravity wave activity in the middle atmosphere over Davis (69°S, 78°E), Antarctica, J. Geophys. Res. Atmos., 120, 4506–4521. doi:10.1002/2014JD022879.

Kim, Y.-J., S. D. Eckermann, and H.-Y. Chun (2003), An overview of the past, present and future of gravity wave drag parameterization for numerical climate and weather prediction models, Atmos. Ocean, 41, 65–98, doi:10.3137/ao.410105.

Kitamura, Y., and I. Hirota (1989), Small-scale disturbances in thelower stratosphere revealed by daily rawin sonde observations, J. Meteorol. Soc. Jpn., 67, 817–830, doi: 10.2151/jmsj1965.67.5_817.

Klostermeyer, J. (1972), Numerical calculation of gravity wave propagation in a realistic thermosphere, J. Atmos. Terr. Phys., 34, 765–774, doi: 10.1016/0021-9169(72)90109-2.

Kogure, M., T. Nakamura, M. K. Ejiri, T. Nishiyama, Y. Tomikawa, M. Tsutsumi, H. Suzuki, T. T. Tsuda, T. D. Kawahara, and M. Abo (2017), Rayleigh/Raman lidar observations of gravity wave activity from 15 to 70 km altitude over Syowa (69°S, 40°E), the Antarctic, J. Geophys. Res. Atmos., 122, 7869–7880, doi:10.1002/2016JD026360.

Lee, Y., Shindell, D., Faluvegi, G., Wenig, M., Lam, Y., Ning, Z., Hao, S., & Lai, C. (2014). Increase of ozone concentrations, its temperature sensitivity and the precursor factor in South China. Tellus B, 66, doi:http://dx.doi.org/10.3402/tellusb.v66.23455

Lindzen, R. S. (1981), Turbulence and stress owing to gravity wave and tidal breakdown, J. Geophys. Res., 86(C10), 9707–9714, doi:10.1029/JC086iC10p09707.

Liu H.-L., and S. L. Vadas (2013), Large-scale ionospheric disturbances due to the dissipation of convectively-generated gravity waves over Brazil, J. Geophys. Res. Space Physics, 118, 2419–2427, doi: 10.1002/jgra.50244.

Liu, H.-L., J. M. McInerney, S. Santos, P. H. Lauritzen, M. A. Taylor, and N. M. Pedatella (2014), Gravity waves simulated by high-resolution Whole Atmosphere Community Climate Model, Geophys. Res. Lett., 41, 9106–9112, doi:10.1002/2014GL062468.

Liu, X., J. Yue, J. Xu, L. Wang, W. Yuan, J. M. Russell III, and M. E. Hervig (2014), Gravity wave variations in the polar stratosphere and mesosphere from SOFIE/AIM temperature observations, J. Geophys. Res. Atmos., 119, 7368–7381, doi:10.1002/2013JD021439.

Lu, X., X. Chu, T. Fuller-Rowell, L. Chang, W. Fong, and Z. Yu (2013), Eastward propagating planetary waves with periods of 1–5 days in the winter Antarctic stratosphere as revealed by MERRA and lidar, J. Geophys. Res. Atmos., 118, 9565–9578, doi:10.1002/jgrd.50717.

Lu, X., X. Chu, W. Fong, C. Chen, Z. Yu, B. R. Roberts, and A. J. McDonald (2015a), Vertical evolution of potential energy density and vertical wave number spectrum of Antarctic gravity waves from 35 to 105 km at McMurdo (77.8°S, 166.7°E). J. Geophys. Res. Atmos., 120, 2719–2737. doi: 10.1002/2014JD022751.

Lu, X., C. Chen, W. Huang, J. A. Smith, X. Chu, T. Yuan, P.-D. Pautet, M. J. Taylor, J. Gong, and C. Y. Cullens (2015b), A coordinated study of 1 h mesoscale gravity waves propagating from Logan to Boulder with CRRL Na Doppler lidars and temperature mapper, J. Geophys. Res. Atmos., 120, doi:10.1002/2015JD023604.

Lu, X., X. Chu, H. Li, C. Chen, J. Smith, S. Vadas (2016), Statistical characterization of high-to-medium frequency mesoscale gravity waves by lidar-measured vertical winds and temperatures in the MLT, J. Atmos. Solar- Terr. Phys., doi:10.1016/j.jastp.2016.10.009.

Lu, X., X. Chu, C. Chen, V. Nguyen, and A. K. Smith (2017). First observations of short-period eastward propagating planetary waves from the stratosphere to the lower thermosphere (110 km) in winter Antarctica, Geophys. Res. Lett., 44, 10,744–10,753. https://doi.org/10.1002/2017GL075641.

Lübken, F.-J., J. Höffner, T. P. Viehl, E. Becker, R. Latteck, B. Kaifler, D. J. Murphy, and R. J. Morris (2015), Winter/summer transition in the Antarctic mesopause region, J. Geophys. Res. Atmos., 120, 12,394–12,409, doi:10.1002/2015JD023928.

Lue, H.Y., F.S. Kuo, S. Fukao, and T. Nakamura (2013), Studies of gravity wave propagation in the mesosphere observed by MU radar, Annales Geophysicae, 31(5), 845-858, doi:10.5194/angeo-31-845-2013.

Madden, R. A. (2007). Large-scale, free Rossby waves in the atmosphere—an update. Tellus A: Dynamic Meteorology and Oceanography, 59(5), 571–590. doi:10.1111/j.1600-0870.2007.00257.x.

Malvern, L. E. (1969), Introduction to the Mechanics of a Continuous Medium, pp. 120–135, Prentice-Hall, Englewood Cliffs, New Jersey.

Manson, A. H., & Meek, C. E. (1993). Characteristics of gravity waves (10 min-6 hours) at Saskatoon (52°N, 107°W): Observations by the phase coherent medium frequency radar. Journal of Geophysical Research, 98(D11), 20357. doi:10.1029/93jd02369.

Marks, C. J., and S. D. Eckermann (1995), A three-dimensional non-hydrostatic ray-tracing model for gravity waves: Formulation and preliminary results for the middle atmosphere, J. Atmos. Sci., 52, 1959–1984, doi: 10.1175/1520-0469(1995)052<1959:ATDNRT>2.0.CO;2.

Martinis C, Hickey D, Oliver W, Aponte N, Brum CGM, Akmaev R, Wright A and Miller C (2013), "The midnight temperature maximum from Arecibo incoherent scatter radar ion temperature measurements", Journal of Atmospheric and Solar-Terrestrial Physics, October, 2013. Vol. 103, pp. 129-137, doi: 10.1016/j.jastp.2013.04.014

Matsuda, T. S., T. Nakamura, M. K. Ejiri, M. Tsutsumi, and K. Shiokawa (2014), New statistical analysis of the horizontal phase velocity distribution of gravity waves observed by airglow imaging, J. Geophys. Res. Atmos., 119, 9707–9718, doi: 10.1002/2014JD021543.

Matsuda, T. S., T. Nakamura, M. K. Ejiri, M. Tsutsumi, Y. Tomikawa, M. J. Taylor, Y. Zhao, P. - D. Pautet, D. J. Murphy, and T. Moffat - Griffin (2017), Characteristics of mesospheric gravity waves over Antarctica observed by Antarctic Gravity Wave Instrument Network imagers using 3 - D spectral analyses, J. Geophys. Res. Atmos., 122, 8969–8981, doi: 10.1002/2016JD026217.

McDonald, A. J., S. E. George, and R. M. Woollands (2009), Can gravity waves significantly impact PSC occurrence in the Antarctic?, Atmos. Chem. Phys., 9, 8825-8840, doi:10.5194/acp-9-8825-2009.

McLandress, C., W. E. Ward, V. I. Fomichev, K. Semeniuk, S. R. Beagley, N. A. McFarlane, and T. G. Shepherd (2006), Large-scale dynamics of the mesosphere and lower thermosphere: An analysis using the extended Canadian Middle Atmosphere Model, J. Geophys. Res., 111, D17111, doi:10.1029/2005JD006776.

McLandress, C., T. G. Shepherd, S. Polavarapu, and S. R. Beagley (2012), Is Missing Orographic Gravity Wave Drag near 60°S the Cause of the Stratospheric Zonal Wind Biases in Chemistry–Climate Models?, J. Atmos. Sci., 69(3), 802–818, doi:10.1175/JAS-D-11-0159.1.

Mitchell, N. J., and V. S. C. Howells (1998), Vertical velocities asso-ciated with gravity waves measured in the mesosphere andlower thermosphere with the EISCAT VHF radar, Ann.Geophys., 16, 1367–1379.

Mihalikova, M., K. Sato, M. Tsutsumi, and T. Sato (2016), Properties of inertia-gravity waves in the lowermost stratosphere as observed by the PANSY radar over Syowa Station in the Antarctic, Annales Geophysicae, 34(5), 543–555, doi:10.5194/angeo-34-543-2016.

Moffat-Griffin, T., R. E. Hibbins, M. J. Jarvis, and S. R. Colwell (2011), Seasonal variations of gravity wave activity in the lower stratosphere over an Antarctic Peninsula station, J. Geophys. Res., 116, D14111, doi:10.1029/2010JD015349.

Moffat-Griffin, T., M. J. Jarvis, S. R. Colwell, A. J. Kavanagh, G. L. Manney, and W. H. Daffer (2013), Seasonal variations in lower stratospheric gravity wave energy above the Falkland Islands, J. Geophys. Res. Atmos., 118, 10,861–10,869, doi:10.1002/jgrd.50859.

Moreels, G. and M. Herse (1977), Photographic evidence of waves around the 85 km level, Planet. Space ScL, 25, 265, doi: 10.1016/0032-0633(77)90137-4.

Murphy, D. J., S. P. Alexander, A. R. Klekociuk, P. T. Love, and R. A. Vincent (2014), Radiosonde observations of gravity waves in the lower stratosphere over Davis, Antarctica, J. Geophys. Res. Atmos., 119, 11,973–11,996, doi:10.1002/2014JD022448.

Mzé, N., A. Hauchecorne, P. Keckhut, and M. Thétis (2014), Vertical distribution of gravity wave potential energy from long-term Rayleigh lidardata at a northern middle-latitude site, J. Geophys. Res. Atmos., 119, 12,069–12,083, doi:10.1002/2014JD022035.

Nakamura, T., Tsuda, T., Yamamoto, M., Fukao, S., & Kato, S. (1993). Characteristics of gravity waves in the mesosphere observed with the middle and upper

atmosphere radar: 1. Momentum flux. Journal of Geophysical Research: Atmospheres, 98(D5), 8899–8910. doi:10.1029/92jd02978.

Nakamura, T., Tsuda, T., Fukao, S., Manson, A. H., Meek, C. E., Vincent, R. A., & Reid, I. M. (1996). Mesospheric gravity waves at Saskatoon (52°N), Kyoto (35°N), and Adelaide (35°S). Journal of Geophysical Research: Atmospheres, 101(D3), 7005–7012. doi:10.1029/95jd03826.

National Research Council (2013), Solar and Space Physics, National Academies Press, Washington, D.C., doi: 10.17226/13060.

Nicolls, M. J., R. H. Varney, S. L. Vadas, P. A. Stamus, C. J. Heinselman, R. B. Cosgrove, and M. C. Kelley (2010), Influence of an inertia-gravity wave on mesospheric dynamics: A case study with the Poker Flat Incoherent Scatter Radar, J. Geophys. Res., 115, D00N02, doi:10.1029/ 2010JD014042.

OASIS (2014), Exploring the interaction of Earth's atmosphere with space, Report to National Science Foundation.

Peterson, A. W. and Kieffaber, L. M. (1973) Infrared photography of OH airglow structures. Nature 244, 92, doi: 10.1038/242321a0.

Pfenninger, M., A. Z. Liu, G. C. Papen, and C. S. Gardner (1999), Gravity wave characteristics in the lower atmosphere at south pole, J. Geophys. Res., 104(D6), 5963–5984, doi:10.1029/98JD02705.

Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, J. Geophys. Res., 107(A12), 1468, doi:10.1029/2002JA009430.

Plougonven, R., A. Hertzog, and H. Teitelbaum (2008), Observations and simulations of a large-amplitude mountain wave breaking over the Antarctic Peninsula, J. Geophys. Res., 113, D16113, doi:10.1029/2007JD009739.

Plougonven, R., and F. Zhang (2014), Internal gravity waves from atmospheric jets and fronts, Rev. Geophys., 52, 33–76, doi:10.1002/2012RG000419.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1986), Numerical Recipes: The Art of Scientific Computing, Cambridge Univ. Press, New York.

Preusse, P., S. D. Eckermann, and M. Ern (2008), Transparency of the atmosphere to short horizontal wavelength gravity waves, J. Geophys. Res., 113, D24104, doi:10.1029/2007JD009682.

Rabier, F. et al. (2012), The Concordiasi Field Experiment over Antarctica: First Results from Innovative Atmospheric Measurements, Bull. Amer. Meteor. Soc., 94, ES17–ES20, doi: 10.1175/BAMS-D-12-00005.1.

Rastogi, P. K. (1981). Radar studies of gravity waves and tides in the middle atmosphere: a review. Journal of Atmospheric and Terrestrial Physics, 43(5-6), 511–524. doi:10.1016/0021-9169(81)90113-6

Richter, J. H., F. Sassi, and R. R. Garcia (2010), Towards a physically based gravity wave source parameterization in a general circulation model, J. Atmos. Sci., 67, 136–156, doi: http://dx.doi.org/10.1175/2009JAS3112.1

Roble, R. G. (1995), Energetics of the mesosphere and thermosphere, the upper mesosphere and lower thermosphere: A review of experiment and theory, Geophys. Monogr., 87, 1–22, doi: 10.1029/GM087

Santos, P. T., C. G. M. Brum, C. A. Tepley, N. Aponte, S. A. González, and E. Robles (2011), Using incoherent scatter radar to investigate the neutral wind long-term trend over Arecibo, J. Geophys. Res., 116, A00H13, doi: 10.1029/2011JA016514.

Sato, K., T. Kumakura, and M. Takahashi (1999), Gravity waves appearing in a high-resolution GCM simulation. J. Atmos. Sci., 56, 1005–1018, doi: 10.1175/1520-0469(1999)056<1005:GWAIAH>2.0.CO;2.

Sato, K., and M. Yoshiki (2008), Gravity wave generation around the polar vortex in the stratosphere revealed by 3 hourly radiosonde observations at Syowa Station, J. Atmos. Sci., 65, 3719–3735, doi:10.1175/2008JAS2539.1.

Sato, K., M. Kohma, M. Tsutsumi, and T. Sato (2017), Frequency spectra and vertical profiles of wind fluctuations in the summer Antarctic mesosphere revealed by MST radar observations, J. Geophys. Res. Atmos., 122, 3–19, doi:10.1002/2016JD025834.

Senft, D. C., and C. S. Gardner (1991), Seasonal variability of gravity wave activity and spectra in the mesopause region at Urbana, J. Geophys. Res., 96, 17,229–17,264, doi:10.1029/91JD01662.

Shibuya, R., K. Sato, M. Tsutsumi, T. Sato, Y. Tomikawa, K. Nishimura, and M. Kohma (2017), Quasi-12 h inertia–gravity waves in the lower mesosphere observed by the PANSY radar at Syowa Station (39.6° E, 69.0° S), Atmos. Chem. Phys., 17, 6455-6476, doi:10.5194/acp-17-6455-2017.

Smith, A. K. (2012), Global dynamics of the MLT, Surv. Geophys., 33, 1177–1230, doi:10.1007/s10712-012-9196-9.

Smith R.B., A. D. Nugent, C. G. Kruse, D. C. Fritts, J. D. Doyle, S. D. Eckermann, M.l J. Taylor, A. Dörnbrack, M. Uddstrom, W. Cooper, P. Romashkin, J. Jensen, and S. Beaton (2016), Stratospheric Gravity Wave Fluxes and Scales during DEEPWAVE, J. Atmos. Sci., 73, 2851–2869, doi: 10.1175/JAS-D-15-0324.1.

Steele, H. M., P. Hamill, M. P. McCormick, and T. J. Swissler (1983), The formation of polar stratospheric clouds, J. Atmos. Sci., 40, 2055-2068, doi: 10.1175/1520-0469(1983)040<2055:TFOPSC>2.0.CO;2.

Tan, B. (2012), Observational and modeling study of polar middle atmosphere dynamics and thermal structures, PhD dissertation, University of Colorado Boulder.

Thurairajah, B., R. L. Collins, V. L. Harvey, R. S. Lieberman, and K. Mizutani (2010), Rayleigh lidar observations of reduced gravity wave activity during the formation of an elevated stratopause in 2004 at Chatanika, Alaska (65°N, 147°W), J. Geophys. Res., 115, D13109, doi:10.1029/2009JD013036.

Tsuda, T., S. Kato, T. Yokoi, T. Inoue, M. Yamamoto, T. E. VanZandt, S. Fukao, and T. Sato (1990a), Gravity waves in the mesosphere observed with the middle and upper atmosphere radar, Radio Sci., 25(5), 1005–1018, doi: 10.1029/RS025i005p01005.

Tsuda, T., Murayama, Y., Yamamoto, M., Kato, S., & Fukao, S. (1990b). Seasonal variation of momentum flux in the mesosphere observed with the MU radar. Geophysical Research Letters, 17(6), 725–728. doi:10.1029/gl017i006p00725.

Tsuda, T., Murayama, Y., Nakamura, T., Vincent, R. A., Manson, A. H., Meek, C. E., & Wilson, R. L. (1994). Variations of the gravity wave characteristics with height, season and latitude revealed by comparative observations. Journal of Atmospheric and Terrestrial Physics, 56(5), 555–568. doi:10.1016/0021-9169(94)90097-3.

Tsuda, T., M. Nishida, C. Rocken, and R. H. Ware (2000), A global morphology of gravity wave activity in the stratosphere revealed by the GPS occultation data (GPS/MET), J. Geophys. Res., 105(D6), 7257–7273, doi:10.1029/1999JD901005.

Taylor, M.J., M.A. Hapgood, and P. Rothwell (1987), Observations of gravity wave propagation in the 01(557.7nm), Na(589.2nm) and the near infrared OH nightglow emissions, Planer. Space Sci., 35,413, doi: 10.1016/0032-0633(87)90098-5.

Taylor, M. J., & Edwards, R. (1991). Observations of short period mesospheric wave patterns: In situ or tropospheric wave generation? Geophysical Research Letters, 18(7), 1337–1340. doi:10.1029/91gl01516.

Taylor, M. J., Ryan, E. H., Tuan, T. F., & Edwards, R. (1993). Evidence of preferential directions for gravity wave propagation due to wind filtering in the middle atmosphere. Journal of Geophysical Research: Space Physics, 98(A4), 6047–6057. doi:10.1029/92ja02604.

Taylor, M.J., V. Taylor, and R Edwards (1995), An investigation of thunderstorms as a source of short period mesospheric gravity waves, The Upper Mesosphere and Lower Thermosphere: A Review of Experiment and Theory, AGU Geophys. Monograph,177, doi: 10.1029/GM087p0177

Vadas, S. L. and D. C. Fritts (2001), Gravity wave radiation and mean responses to local body forces in the atmosphere, J. Atmos. Sci., 58, 2249–2279.

Vadas, S. L., D. C. Fritts, and M. J. Alexander (2003), Mechanism for the generation of secondary waves in wave breaking regions. J. Atmos. Sci., 60, 194–214, doi: 10.1175/1520-0469(2003)060<0194:MFTGOS>2.0.CO;2.

Vadas, S. L., and D. C. Fritts (2005), Thermospheric responses to gravity waves: Influences of increasing viscosity and thermal diffusivity, J. Geophys. Res., 110, D15103, doi:10.1029/2004JD005574.

Vadas, S. L., M. J. Taylor, P.-D. Pautet, P. A. Stamus, D. C. Fritts, H.-L. Liu, F. T. São Sabbas, V. T. Rampinelli, P. Batista and H. Takahashi (2009), Convection: the likely source of medium-scale gravity waves observed in the OH airglow layer near Basilia, Brazil, during the SpreadFEx Campaign, Ann. Geophys., 27, 231–259.

Vadas, S. L. (2013), Compressible f-plane solutions to body forces, heatings, and coolings, and application to the primary and secondary gravity waves generated by a deep convective plume, J. Geophys. Res. Space Physics, 118, 2377-2397, doi:10.1002/jgra.50163.

Vadas, S. L., and Becker, E.(2018). Numerical modeling of the excitation, propagation, and dissipation of primary and secondary gravity waves during wintertime at McMurdo Station in the Antarctic. Journal of Geophysical Research: Atmospheres, 123, 9326–9369. https://doi.org/10.1029/2017JD027974.

Vadas, S. L., J. Zhao, X. Chu, and E. Becker (2018), The excitation of secondary gravity waves from body forces: Theory and observation, J. Geophys. Res. Atmos., in press (paper number 2017JD027970R).

Vincent, R. A., & Fritts, D. C. (1987). A Climatology of Gravity Wave Motions in the Mesopause Region at Adelaide, Australia. Journal of the Atmospheric Sciences, 44(4), 748–760. doi:10.1175/1520-0469(1987)044<0748:acogwm>2.0.co;2.

Vincent, R. A., A. Hertzog, G. Boccara, and F. Vial (2007), Quasi-Lagrangian superpressure balloon measurements of gravity-wave momentum fluxes in the polar stratosphere of both hemispheres, Geophys. Res. Lett., 34, L19804, doi:10.1029/2007GL031072.

Walterscheid, R. L., L. J. Gelinas, C. R. Mechoso, and G. Schubert (2016), Spectral distribution of gravity wave momentum fluxes over the Antarctic Peninsula from Concordiasi superpressure balloon data, J. Geophys. Res. Atmos., 121, 7509–7527, doi:10.1002/2015JD024253.

Wang, Z., X. Chu, W. Huang, W. Fong, J. A. Smith, and B. Roberts (2012), Refurbishment and upgrade of Fe Boltzmann/Rayleigh temperature lidar at Boulder for McMurdo lidar campaign in Antarctica, in Proceeding of the 26th International Laser Radar Conference, pp. 207–210, Porto Heli, Greece.

Watanabe, S., K. Sato, and M. Takahashi (2006), A general circulation model study of the orographic gravity waves over Antarctica excited by katabatic winds, J. Geophys. Res., 111 (D18104), doi:10.1029/2005JD006851.

Whiteway, J.A. and A.I. Carswell (1994), Rayleigh lidar observations of thermal structure and gravity wave activity in the high Arctic during a stratospheric warming. J. Atmos. Sci., 51, 3122–3136, doi:10.1175/1520-0469(1994)051<3122:RLOOTS>2.0.CO;2.

Whiteway, J. A., and A. I. Carswell (1995), Lidar observations of gravity wave activity in the upper stratosphere over Toronto, J. Geophys. Res., 100(D7), 14113–14124, doi:10.1029/95JD00511.

Whiteway, J. A., T. J. Duck, D. P. Donovan, J. C. Bird, S. R. Pal, and A. I. Carswell (1997), Measurements of gravity wave activity within and around the Arctic stratospheric vortex, Geophys. Res. Lett., 24, 1387–1390, doi:10.1029/97GL01322.

Wilson, R., M. L. Chanin, and A. Hauchecorne (1991), Gravity waves in the middle atmosphere observed by Rayleigh lidar: 1. Case studies, J. Geophys. Res., 96(D3), 5153–5167, doi:10.1029/90JD02231.

Wright, C. J., and J. C. Gille (2013), Detecting overlapping gravity waves using the S-Transform, Geophys. Res. Lett., 40, 1850–1855, doi:10.1002/grl.50378.

Yamashita, C., X. Chu, H.-L. Liu, P. J. Espy, G. J. Nott, and W. Huang (2009), Stratospheric gravity wave characteristics and seasonal variations observed by lidar at the South Pole and Rothera, Antarctica, J. Geophys. Res., 114, D12101, doi:10.1029/2008JD011472.

Yasuda, Y., K. Sato, and N. Sugimoto (2015), A theoretical study on the spontaneous radiation of inertia-gravity waves using the renormalization group method. Part I: Derivation of the renormalization group equations, J. Atmos. Sci., 72, 957–983.

York, D., Evensen, N., Martinez, M., Delgado, J., (2004), Unified equations for the slope, intercept, and standard errors of the best straight line. Am. J. Phys. 72, 367–375, doi: 10.1119/1.1632486

Yoshiki, M., and K. Sato (2000), A statistical study of gravity waves in the polar regions based on operational radiosonde data, J. Geophys. Res., 105(D14), 17995–18011, doi:10.1029/2000JD900204.

Yoshiki, M., N. Kizu, and K. Sato (2004), Energy enhancements of gravity waves in the Antarctic lower stratosphere associated with variations in the polar vortex

and tropospheric disturbances, J. Geophys. Res., 109, D23104, doi:10.1029/2004JD004870.

Yu, Z., X. Chu, W. Huang, W. Fong, and B. R. Roberts (2012), Diurnal variations of the Fe layer in the mesosphere and lower thermosphere: Four season variability and solar effects on the layer bottomside at McMurdo (77.8°S, 166.7°E), Antarctica, J. Geophys. Res., 117, D22303, doi:10.1029/2012JD018079.

Yu, Z. (2014), Lidar Observations and Numerical Modeling Studies of Thermospheric Metal Layers and Solar Effects on Mesospheric Fe Layers, Aerospace Engineering Sciences Graduate Theses & Dissertations. 90. https://scholar.colorado.edu/asen_gradetds/90.

Zhang, S. D., Huang, C. M., Huang, K. M., Zhang, Y. H., Gong, Y., & Gan, Q. (2017). Vertical wavenumber spectra of three dimensional windsrevealed by radiosonde observations at midlatitude. Annales de Geophysique, 35(1), 107–116. https://doi.org/10.5194/angeo-35-107-2017.

Zhao, J., X. Chu, C. Chen, X. Lu, W. Fong, Z. Yu, R. M. Jones, B. R. Roberts, and A. Dörnbrack (2017), Lidar observations of stratospheric gravity waves from 2011 to 2015 at McMurdo (77.84° S, 166.69° E), Antarctica: Part I. Vertical wavelengths, periods, and frequency and vertical wavenumber spectra, J. Geophys. Res. Atmos., 122, doi:10.1002/2016JD026368.

Zhu, X., and J. R. Holton (1987), Mean fields induced by local gravity-wave forcing in the middle atmosphere, J. Atmos. Sci., 44, 620–630, doi: 10.1175/1520-0469(1987)044<0620:MFIBLG>2.0.CO;2.

Zink, F., and R. A. Vincent (2001a), Wavelet analysis of stratospheric gravity wave packets over Macquarie Island: 1. Wave parameters, J. Geophys. Res., 106(D10), 10275–10288, doi:10.1029/2000JD900847.

Zink, F., and R. A. Vincent (2001b), Wavelet analysis of stratosphericgravity wave packets over Macquarie Island, 2, Intermittency and mean-flow accelerations, J. Geophys. Res., 106, 10,289–10,297, doi: 10.1029/2000JD900846.

Appendix I. Mathematical Derivation of the Spectral Proportion Method

We develop the mathematical derivations behind this methodology. Conventionally, Fourier transform refers to the decomposition of a time series (time domain) into the frequencies that make it up (frequency domain). The definition is,

$$f^{FFT}(\xi) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i \xi t} dt \qquad \text{(AI.1)}$$

where t represents time, ξ represents frequency, $f(t)$ is a series in time domain, $f^{FFT}(\xi)$ is the corresponding Fourier transform in frequency domain. In modern computers, we employ Discrete Fourier Transform (DFT) since numerical algorithms in computers can only deal with a finite amount of data. The Definition is,

$$X_k = \sum_{n=0}^{N-1} x_n \bullet e^{-2\pi i k n / N} \qquad \text{(AI.2)}$$

where $\{x_n\} = x_0, x_1, ..., x_{N-1}$ is a series of discrete numbers in time domain, $\{X_k\} = X_0, X_1, ..., X_{N-1}$ is the corresponding DFT in frequency domain. The algorithms that developed in modern computers to handle DFT is called Fast Fourier Transform (FFT). According to Parseval's theorem, the sum (or integration) of the square of a function in time domain is equal to the sum (or integration) of the square of its transformation in frequency domain. In the case of DFT, we have

$$\sum_{n=0}^{N-1} |x_n|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X_k|^2 \qquad \text{(AI.3)}$$

In order to be more illustrative, we first come back to the example of temperature perturbations we used in this paper. Here we replace x_n with our derived filtered temperature perturbation $T'(z, t_i)$. For simplicity, we now only consider one single altitude, so we ignore the altitude parameter z and define $T'(t_i) = T'(z, t_i)$. Let $T'(t_i)$ replace x_n in equation (AI.3), we get

$$\sum_{i=0}^{N_p-1} |T'(t_i)|^2 = \frac{1}{N_p} \sum_{k=0}^{N_p-1} |T'^{,FFT}(\xi_k)|^2 \qquad \text{(AI.4)}$$

where N_p is the number length of the observation in time domain, $T^{',FFT}(\xi_k)$ is the corresponding DFT of $T^{'}(t_i)$. We have the equation solving for E_{pm} (here we ignore the altitude parameter z again for simplicity),

$$E_{pm} = \frac{1}{2}\frac{g^2}{\overline{N^2}}\overline{\left(\frac{T_{GW}^{'}(t)}{T_{Bkg}}\right)^2} = \frac{1}{2}\frac{g^2}{\overline{N^2}\bullet T_{Bkg}^2}\frac{1}{N_p}\sum_{i=0}^{N_p-1}\left|T_{GW}^{'}(t_i)\right|^2 = \frac{1}{2}\frac{g^2}{\overline{N^2}\bullet T_{Bkg}^2}\frac{1}{N_p}\sum_{i=0}^{N_p-1}\left|T^{'}(t_i)\right|^2 \bullet P \quad \text{(AI.5)}$$

In this equation, we define gravity wave proportion P as on average, how much the energy of the gravity wave induced perturbations $\sum_{i=0}^{N_p-1}\left|T_{GW}^{'}(t_i)\right|^2$ occupy the total energy of the filtered temperature perturbations $\sum_{i=0}^{N_p-1}\left|T^{'}(t_i)\right|^2$ in time domain.

The key is to estimate the gravity wave proportion P so that we know everything on the right side of equation (AI.5). It is difficult to estimate P in time domain since the perturbations induced by gravity waves and temperature uncertainties are mixed all together. We attempt to solve for P in frequency domain according to equation (AI.4). Here we confirm that if we obtain gravity wave proportion P_f in frequency domain, then the gravity wave proportion P in time domain should be exactly equal to P_f in frequency domain via mathematical derivations. We decompose the filtered temperature perturbation $T^{'}(t_i)$ into gravity wave induced perturbation $T_{GW}^{'}(t_i)$ and noise (observed temperature uncertainty) induced perturbation $T_{Noise}^{'}(t_i)$,

$$T^{'}(t_i) = T_{GW}^{'}(t_i) + T_{Noise}^{'}(t_i) \quad \text{(AI.6)}$$

According to the linearity property of FFT,

$$T^{',FFT}(\xi_k) = T_{GW}^{',FFT}(\xi_k) + T_{Noise}^{',FFT}(\xi_k) \quad \text{(AI.7)}$$

where $T^{',FFT}(\xi_k), T_{GW}^{',FFT}(\xi_k), T_{Noise}^{',FFT}(\xi_k)$ are the corresponding FFT of $T^{'}(t_i), T_{GW}^{'}(t_i), T_{Noise}^{'}(t_i)$. Here we invoke Parseval's theorem (i.e., equation (AI.3)) so that,

$$\sum_{i=0}^{N_p-1}\left|T^{'}(t_i)\right|^2 = \frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left|T^{',FFT}(\xi_k)\right|^2 \quad \text{(AI.8)}$$

$$\sum_{i=0}^{N_p-1}\left|T_{GW}^{'}(t_i)\right|^2 = \frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left|T_{GW}^{',FFT}(\xi_k)\right|^2 \quad \text{(AI.9)}$$

$$\sum_{i=0}^{N_p-1}\left|T'_{Noise}(t_i)\right|^2=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left|T'^{,FFT}_{Noise}(\xi_k)\right|^2 \qquad \text{(AI.10)}$$

Since we employed zero padding before applying FFT on the original time series, here N_p represents the number length of the original time series, N_p^{Final} represents the number length of the time series after zero padding. Plug equation (AI.6) and (AI.7) into the left and right side of equation (AI.8), respectively, we have

$$\sum_{i=0}^{N_p-1}\left|T'(t_i)\right|^2=\sum_{i=0}^{N_p-1}\left|T'_{GW}(t_i)\right|^2+\sum_{i=0}^{N_p-1}\left|T'_{Noise}(t_i)\right|^2+\sum_{i=0}^{N_p-1}\left[2\bullet T'_{GW}(t_i)\bullet T'_{Noise}(t_i)\right]$$

$$=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left|T'^{,FFT}(\xi_k)\right|^2$$

$$=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left|T'^{,FFT}_{GW}(\xi_k)\right|^2+\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left|T'^{,FFT}_{Noise}(\xi_k)\right|^2+\frac{1}{N_p^{Final}}\sum_{i=0}^{N_p^{Final}-1}\left[2\bullet T'^{,FFT}_{GW}(\xi_k)\bullet T'^{,FFT}_{Noise}(\xi_k)\right]$$

(AI.11)

Considering equation (AI.9), (AI.10), and (AI.11), we have

$$\sum_{i=0}^{N_p-1}\left[2\bullet T'_{GW}(t_i)\bullet T'_{Noise}(t_i)\right]=\frac{1}{N_p^{Final}}\sum_{i=0}^{N_p^{Final}-1}\left[2\bullet T'^{,FFT}_{GW}(\xi_k)\bullet T'^{,FFT}_{Noise}(\xi_k)\right] \qquad \text{(AI.12)}$$

We recall that we perform a Monte Carlo simulation as described before where we simulate the temperature observation field for 1000 times using the combination of the observed temperature field and observed temperature uncertainties. Hence, after background subtraction and wave filtering process, we have 1000 simulated temperature perturbation field $T'_{Simu,j}(t_i)$. Here $j=1,2,...,1000$ denotes the 1000 times of simulations. In this case, we have

$$T'_{Simu,j}(t_i)=T'_{GW,j}(t_i)+T'_{Noise,j}(t_i) \qquad \text{(AI.13)}$$

Since $T'_{GW,j}(t_i)$ is the perturbations induced purely by gravity waves, theoretically, it should be the same during every simulation (i.e., for every j). $T'_{Noise,j}(t_i)$ is the perturbations induced by noise (temperature uncertainty). We apply FFT on the 1000 simulated perturbation time series $T'_{Simu,j}(t_i)$ individually, and get a mean FFT spectrum from the obtained 1000 FFT spectra. Once we transform the perturbations

into frequency domain (i.e., the mean FFT spectrum), it is relatively a lot easier to estimate gravity wave proportion since we can intuitively determine the noise floor level. In order to make the plotted averaged FFT spectrum more illustrative, we only showed the positive portion of the FFT previously since FFT gives symmetric results for positive and negative frequencies with respect to frequency 0. Hence, similarly to equation (AI.8), (AI.9), (AI.10), (AI.11), and (AI.12), from this 1000 simulations, we have

$$\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{Simu,j}(t_i)\right|^2\right)=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{Simu,j}(\xi_k)\right|^2\right)$$
$$=\frac{N_p^2}{4\bullet N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{Simu,j}(\xi_k)}{N_p}\right|^2\right) \quad \text{(AI.14)}$$

$$\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{GW,j}(t_i)\right|^2\right)=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{GW,j}(\xi_k)\right|^2\right)$$
$$=\frac{N_p^2}{4\bullet N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{GW,j}(\xi_k)}{N_p}\right|^2\right) \quad \text{(AI.15)}$$

$$\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{Noise,j}(t_i)\right|^2\right)=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{Noise,j}(\xi_k)\right|^2\right)$$
$$=\frac{N_p^2}{4\bullet N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{Noise,j}(\xi_k)}{N_p}\right|^2\right) \quad \text{(AI.16)}$$

$$
\begin{aligned}
&\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{Simu,j}(t_i)\right|^2\right)\\
&=\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{GW,j}(t_i)\right|^2\right)+\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{Noise,j}(t_i)\right|^2\right)\\
&+\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left[2\bullet T'_{GW,j}(t_i)\bullet T'_{Noise,j}(t_i)\right]\right)\\
&=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{\mathrm{Si}mu,j}(\xi_k)\right|^2\right)\\
&=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{GW,j}(\xi_k)\right|^2\right)+\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{Noise,j}(\xi_k)\right|^2\right)\\
&+\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left[2\bullet T'^{,FFT}_{GW,j}(\xi_k)\bullet T'^{,FFT}_{Noise,j}(\xi_k)\right]\right)
\end{aligned}
\tag{AI.17}
$$

$$
\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left[2\bullet T'_{GW,j}(t_i)\bullet T'_{Noise,j}(t_i)\right]\right)=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left[2\bullet T'^{,FFT}_{GW,j}(\xi_k)\bullet T'^{,FFT}_{Noise,j}(\xi_k)\right]\right) \tag{AI.18}
$$

where $T'^{,FFT}_{Simu,j}(\xi_k), T'^{,FFT}_{GW,j}(\xi_k), T'^{,FFT}_{Noise,j}(\xi_k)$ are the corresponding FFT of $T'_{Simu,j}(t_i), T'_{GW,j}(t_i), T'_{Noise,j}(t_i)$. We consider $T'_{GW,j}(t_i)$ and $T'_{Noise,j}(t_i)$ are uncorrelated during different simulations, and $\frac{1}{1000}\sum_{j=1}^{1000}\left[T'_{Noise,j}(t_i)\right]=\vec{0}$ since $T'_{Noise,j}(t_i)$ should have a mean of 0 at different t_i among different simulations, according to a general correlation coefficient calculation equation $\rho_{X,Y}=\frac{E[XY]-E[X]E[Y]}{\sqrt{E[X^2]-[E[X]]^2}\sqrt{E[Y^2]-[E[Y]]^2}}$

Equation (AI.18) is equal to 0,

$$
\begin{aligned}
&\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left[2\bullet T'_{GW,j}(t_i)\bullet T'_{Noise,j}(t_i)\right]\right)\\
&=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left[2\bullet T'^{,FFT}_{GW,j}(\xi_k)\bullet T'^{,FFT}_{Noise,j}(\xi_k)\right]\right)=0
\end{aligned}
\tag{AI.19}
$$

So from equation (AI.17) and (AI.19), we have,

$$\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{Simu,j}(t_i)\right|^2\right)=\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{GW,j}(t_i)\right|^2\right)+\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{Noise,j}(t_i)\right|^2\right)$$
$$=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{Simu,j}(\xi_k)\right|^2\right) \quad (AI.20)$$
$$=\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{GW,j}(\xi_k)\right|^2\right)+\frac{1}{N_p^{Final}}\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'^{,FFT}_{Noise,j}(\xi_k)\right|^2\right)$$

We now take the ratio of (15) over (14), we have

$$P_f=\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{GW,j}(\xi_k)}{N_p}\right|^2\right)\Bigg/\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{Simu,j}(\xi_k)}{N_p}\right|^2\right)$$
$$=\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{GW,j}(t_i)\right|^2\right)\Bigg/\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{Simu,j}(t_i)\right|^2\right)=P \quad (AI.21)$$

The term $\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{GW,j}(t_i)\right|^2\right)\Big/\sum_{i=0}^{N_p-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|T'_{Simu,j}(t_i)\right|^2\right)$ in equation (AI.21) denotes in time domain, on average, how much the gravity wave perturbation energy should occupy the derived filtered temperature perturbation energy, which is the definition of P. Hence, we confirm $P=P_f$.

In the algorithm, this gravity wave area portion P_f is calculated via the first equality in equation (AI.21). We first determine the spectral noise floor level of the obtained averaged FFT spectrum in the frequency domain as described before. By doing that, we select the gravity wave area (the area below the mean FFT curve and above the noise floor level within the obtained frequency range which corresponds to the term $\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{GW,j}(\xi_k)}{N_p}\right|^2\right)$) and the white noise area (the area below the noise floor level and above 0 within the obtained frequency range which corresponds to the term $\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{Simu,j}(\xi_k)}{N_p}\right|^2\right)-\sum_{k=0}^{N_p^{Final}-1}\left(\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T'^{,FFT}_{GW,j}(\xi_k)}{N_p}\right|^2\right)$). We need to pay attention that since we are only deriving proportions of areas in frequency domain, all the scaling factors such as $\frac{N_p^2}{4\bullet N_p^{Final}}$ in the last term in equation (AI.14) does not

matter in the calculation since they will get canceled out when solving for proportions as is shown in deriving equation (AI.21). Also according to equation (AI.21), when we calculate the proportions using our decided gravity wave area and white noise area, we should take sums of discrete values such as $\frac{1}{1000}\sum_{j=1}^{1000}\left|2\bullet\frac{T_{GW,j}^{',FFT}(\xi_k)}{N_p}\right|^2$ in the selected areas. However, the method of taking integrations in the selected areas gives the same result of gravity wave proportion P_f as the method of taking sums of discrete values. This is because when calculating areas by the method of taking integration and taking sums, the results only differ by a multiplication factor. This multiplication factor will eventually get canceled in the process of solving for proportions in the same sense as the above-mentioned scaling factors (e.g., $\frac{N_p^2}{4\bullet N_p^{Final}}$ in the last term in equation (AI.14)). Anyway, gravity wave proportion P can be calculated from the summation of the selected areas as described before and E_{pm} can be solved for according to equation (AI.5).

The method we described in this paper can also be applied to the calculation of many other atmospheric parameters such as the situation when we have horizontal wind measurements (u : zonal wind, v : meridional wind) along with their observed uncertainties ($\delta u, \delta v$) and try to solve for wave kinetic energy density. The process should be similar. We derive the demanded perturbations u' and v' from u and v via background subtraction and wave filtering process depending on what kinds of waves we are interested in. In order to calculate wave kinetic energy per unit mass according to

$$E_k = \frac{1}{2}(\overline{u_{wave}'^2} + \overline{v_{wave}'^2}) \qquad \text{(AI.22)}$$

we still have to extract the perturbations induced purely by waves u_{wave}' and v_{wave}' from the derived u', v' and the observed $\delta u, \delta v$. Hence, we should calculate the wave proportions P_u, P_v in frequency domain utilizing the above method, so that

$$E_k = \frac{1}{2}(\overline{u'^2_{wave}} + \overline{v'^2_{wave}}) = \frac{1}{2}(\overline{u'^2} \bullet P_u + \overline{v'^2} \bullet P_v) \qquad \text{(AI.23)}$$

In a more general sense, say we have an observation of quantity $Q(t)$ and its corresponding uncertainties $\delta Q(t)$ with respect to time t. We extract certain perturbations $Q'(t)$ from $Q(t)$ depending on what part of the signal spectrum we are interested in. We wish to calculate the quantity $\overline{(Q'_{signal}(t))^2}$ from $\overline{(Q'(t))^2}$ by excluding the influence of $\delta Q(t)$ on $Q'(t)$, the above methodology applies.