PHD

Satellite Observations and Spectral Analysis of Stratospheric Gravity Wave Dynamics

Hindley, Neil

Award date:
2016

Awarding institution:
University of Bath

Satellite Observations and Spectral Analysis of Stratospheric Gravity Wave Dynamics

Neil Hindley

A thesis submitted for the degree of Doctor of Philosophy

University of Bath
Department of Electronic and Electrical Engineering
January 2016

Signature of Author . . . .

Abstract

Satellite Observations and Spectral Analysis of Stratospheric Gravity Wave Dynamics

Neil Hindley, University of Bath

January 2016

Gravity waves play a crucial role in the dynamics of the middle atmosphere through the vertical transport of tropospheric energy and momentum. Despite the importance of gravity wave effects, nearly all general circulation models significantly underestimate gravity wave drag around the southern wintertime polar vortex, leading to large discrepancies from observed behaviour and a strong impediment to model progress. Here we use GPS radio occultation data from the COSMIC satellite constellation to investigate key properties of gravity waves in and around the southern polar vortex and over the gravity wave hot spot of the southern Andes/Drake Passage/Antarctic Peninsula. We also develop spectral analysis tools in order to further the capabilities of our data for this purpose.

By analysing vertical profiles of atmospheric temperature from COSMIC, we find evidence of the meridional propagation of waves into the polar vortex from sources far to the north. We develop a new wavelet-based analysis technique for the quantitative identification of gravity waves in COSMIC profiles, and use it to investigate gravity wave intermittency over the hot spot region and around the edge of the polar vortex. We then estimate gravity wave momentum flux over the hot spot from closely-spaced pairs of COSMIC profiles.

Finally, we develop a new two-dimensional spectral analysis method for the measurement of gravity wave amplitudes, horizontal wavelengths and directions of propagation from AIRS measurements, based upon the two-dimensional Stockwell transform. We show that, by using an alternative elliptical spectral window, we can dramatically improve the measurement of wave amplitude. We apply our two-dimensional Stockwell transform to AIRS measurements over the known gravity wave hot spots of South Georgia and the Antarctic Peninsula, measuring gravity properties and momentum fluxes with improved confidence, accuracy and rigour over current methods.

Acknowledgements

Firstly, I would like to thank my supervisor **Nick Mitchell** for all of his time, help, patience and guidance over the course of this PhD. Special mention must also go to my second supervisor **Nathan Smith**, for tirelessly giving up so much of his time to provide the unscrupulous mathematical wizardry and attention to detail that made so much of this work possible.

I'd also like to thank my colleague **Corwin Wright** for all his help and patience with my blithering incompetence, and especially my fellow weary traveller, **Andrew Moss**. We two have had quite the adventure over the last few years.

Thanks must also go to my funding body **NERC** and the **University of Bath** for providing me with food, shelter and WiFi for the last four years, not to mention funds to go off gallivanting around the world on various courses and conferences. I'd also like to thank everyone at **NCAS** and **ERCA** for their brilliant courses and all the kind and clever faces I've met on them.

Finally I'd like to thank **my parents** for pretty much everything, my brothers **Owen** and **Sid** for being bros, **James Lebbon** for keeping my hands dirty, **OSRFC** for keeping my bones broken, **BUST**, **Jon Chouler** and **Giovanni L'Immigrato** for showing me that a PhD student could be so much more than a brain on a stick. And last but not least, **Sarah**, for all of your love, support and ducks.

Contents

List of Figures

Publications

Journal papers authored or co-authored by the author during the course of this PhD are listed here:

4. **A two-dimensional Stockwell transform for gravity wave analysis of AIRS measurements**
   N. P. Hindley, N. D. Smith, C. J. Wright, and N. J. Mitchell
   Atmospheric Measurement Techniques (accepted, in press)
   DOI: 10.5194/amt-2015-383, **2016**

3. **Combining AIRS and MLS observations for three-dimensional gravity wave measurement**
   C. J. Wright, N. P. Hindley and N. J. Mitchell
   Geophysical Research Letters
   DOI: 10.1002/2015GL067233, **2016**

2. **Multi-instrument gravity-wave measurements over Tierra del Fuego and the Drake Passage, Part 1: potential energies and vertical wavelengths from AIRS, COSMIC, HIRDLS, MLS-Aura, SAAMER, SABER and radiosondes**
   C. J. Wright, N. P. Hindley, A. C. Moss, D. C. Fritts, D. Janches and N. J. Mitchell
   Atmospheric Measurement Techniques
   DOI: 10.5194/amt-9-877-2016, **2015**

1. **The southern stratospheric gravity-wave hot spot: individual waves and their momentum fluxes measured by COSMIC GPS-RO**
   N. P. Hindley, C. J. Wright, N. D. Smith and N. J. Mitchell
   Atmospheric Chemistry and Physics
   DOI: 10.5194/acp-15-7797-2015, **2015**

a. [in preparation] **Multi-instrument gravity-wave measurements over Tierra del Fuego and the Drake Passage, Part 2: Gravity wave momentum fluxes, accelerations, intermittency and spectral subsetting**
   C. J. Wright, M. Ern, D. Janches, D. C. Fritts, N. P. Hindley, N. J. Mitchell, and co-authors.
   In preparation for Atmospheric Measurement Techniques

Chapter 1

Introduction

1.1 The structure of the atmosphere

The atmosphere of the earth can be broadly divided into horizontal layers or “spheres”, bounded above and below by a change in the sign of the rate of temperature decrease with respect to height (Fig. 1.1).

In the lowest layer, the troposphere, temperature decreases uniformly with increasing altitude until the rate of decrease “pauses” as temperature reaches a minimum value, called the tropopause. The height of the tropopause ranges from $\sim$8 km in polar regions to $\sim$16 km in the tropics. The temperature decrease with increasing altitude in the troposphere is due to incoming solar radiation heating up the earth’s surface faster than the surrounding air, making air near the ground hotter than air at altitude.

The dynamics of the troposphere are dominated by convective processes. This layer contains more than 75% of the atmosphere’s molecular mass and is home to all meteorological phenomena and day-to-day weather. The temperature minimum at the tropopause acts as a “cold trap”, forcing water vapour to condense out. Atmospheric layers above the tropopause are thus very dry, with water vapour concentrations around 3-4 orders of magnitude less than the troposphere.

Above the troposphere is the stratosphere, so-called due its relatively stable stratification compared to the more tumultuous troposphere. Here, temperature increases with increasing height due to subtle changes in the radiative forcing of key chemical constituents, illustrated in Fig. 1.1(b). Increased concentrations of ozone in the stratosphere mean that, during the day, ultra-violet radiation from the sun is absorbed faster than it is radiated away by carbon dioxide (CO_2), water vapour and the ozone itself, resulting in a net heat-

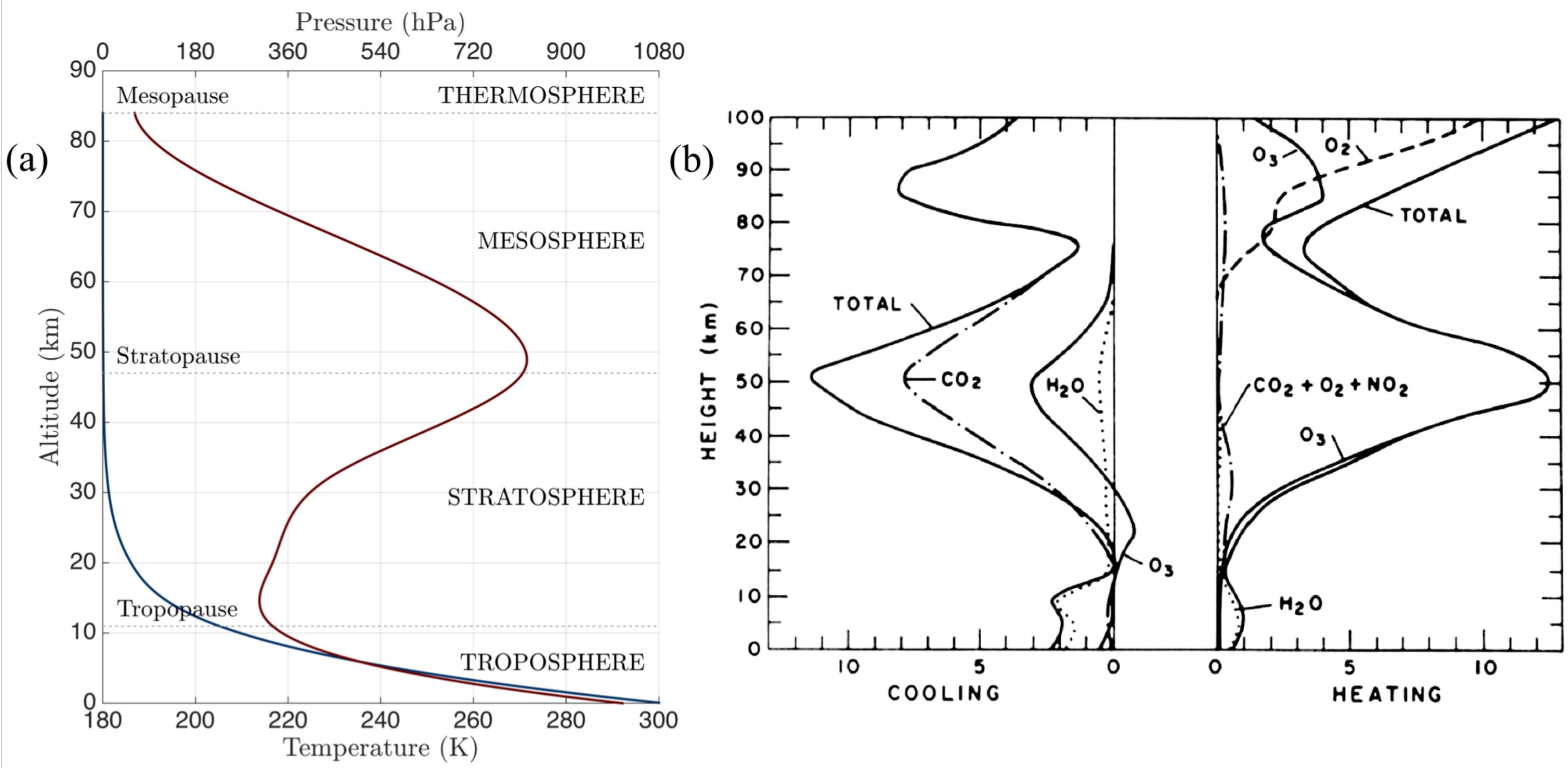

Figure 1.1: Atmospheric temperature (red) and pressure (blue) against height from the International Standard Atmosphere (a) (e.g. Cavcar, 2000) and heating/cooling rate (K/day) as a function of height (b) for trace chemical constituents, from London (1980).

ing. This negative lapse rate is what makes the stratosphere so vertically stable, such that net flow is generally horizontal only.

At an altitude of around 50 km, the temperature gradient reverses such that temperature once again decreases with increasing altitude, a point called the stratopause. Above this lies the mesosphere, where the temperature continues to decrease. This decrease is due to diminishing concentrations of ozone such that radiative emission from CO_2 begins to dominate, resulting in a net cooling. The mesosphere is home to more exotic atmospheric phenomena such as noctilucent clouds, sprites and the aurora. This layer is bounded by another temperature inversion, the mesopause, at around 90 to 100 km, above which lies the thermosphere. Here, ozone and molecular oxygen begin to override the cooling effect of CO_2, resulting in a net heating.

1.2 Dynamics of the middle atmosphere

The dynamics of the stratosphere, mesosphere and lower thermosphere (collectively referred to as the middle atmosphere) are dominated by atmospheric waves and tides. Here, gravity waves, planetary waves and tides force the atmosphere away from thermal equilibrium, resulting in a large number of important processes with global impacts on atmospheric circulation at all levels.

These waves usually have sources in the lower atmosphere (e.g. Fritts and Alexander, 2003), and as they propagate upward their amplitudes grow exponentially in accordance with decreasing density. In the middle atmosphere, these wave motions become so large that they dominate the zonal wind fields. Large wave amplitudes lead to wave breaking and momentum deposition, which in turn drives circulation not expected under thermal equilibrium.

One of the most important wave-driven circulations is the Brewer-Dobson circulation (e.g. Andrews et al., 1987), illustrated in Fig. 1.2. Planetary wave breaking drives a meridional circulation from the tropical stratosphere to the poles, a flow that is particularly strong in the summer hemisphere. In the mesosphere, gravity wave breaking and momentum deposition reverse this flow, driving a meridional flow of air from the summer pole to the winter pole. This forces an upwelling and subsequent adiabatic cooling of the summer mesopause. This cooling effect is so strong the the summer mesopause is the coldest naturally-occurring place on earth, reaching temperatures of 120 K or lower, some 50 K lower than would be expected under thermal equilibrium.

Figure 1.2(a) and (b) show zonal-mean wind speeds for the June and December solstices

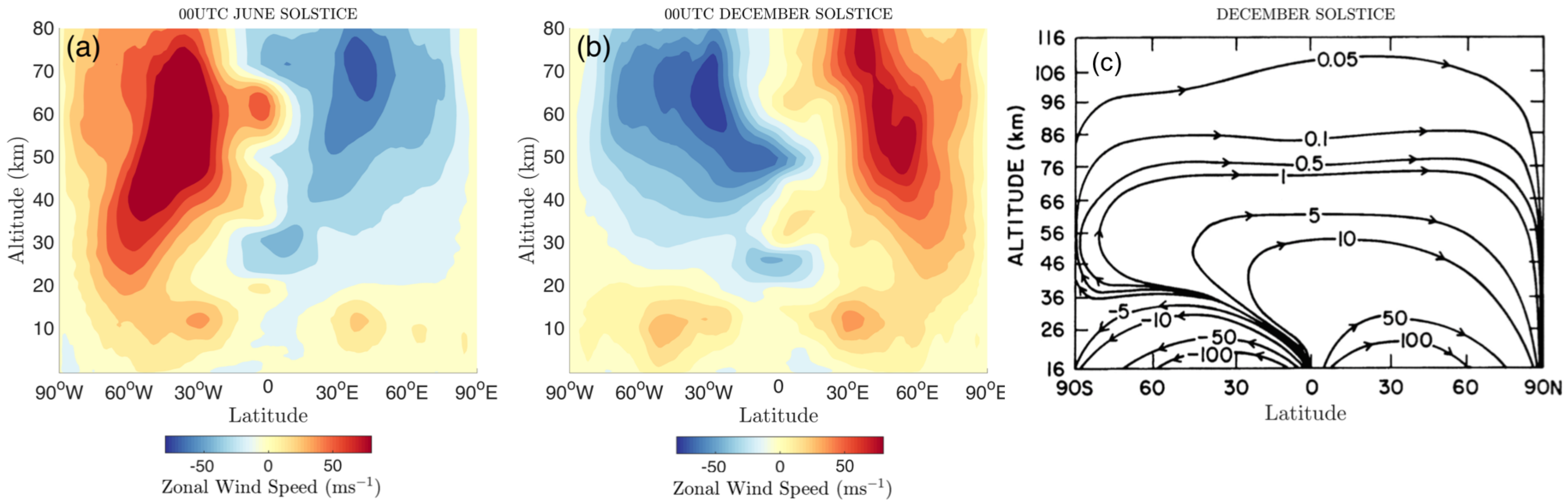

Figure 1.2: Zonal mean wind strength for the June (a) and December (b) solstices from ECMWF operaional analyses and the strength of the Brewer-Dobson circulation at the December solstice (c), illustrated by the meridional mass stream function, with units of kgm^{-1}s^{-1}. The summer to winter mesospheric meridional flow and upwelling over the summer pole results in an adiabatic cooling, causing the cold summer mesopause. From Garcia and Solomon (1983).

in 2012 from European Centre for Medium-Range Weather Forecasts (ECMWF) operational analyses. In the winter hemisphere, an intense vortex of strong westerly zonal winds forms each year in the stratosphere and mesosphere at around 60°S . The formation, strength and subsequent break up of these vortices in spring has far-reaching effects on the circulation and climate of the regions around the globe throughout the year.

1.3 Thesis overview

The work described in this thesis involves the detection and measurement of atmospheric gravity waves using GPS radio occultation and infrared radiometry, together with the development of new spectral analysis techniques for this purpose.

The thesis is ordered into the following chapters:

Chapter 1
We provide a brief description of the structure and dynamics of the terrestrial atmosphere and give an overview of the chapters and key results contained within this thesis.

Chapter 2
We provide an in-depth discussion of atmospheric gravity waves and their properties.

Chapter 3
We describe the instruments and measurements used in our studies.

Chapter 4
We measure gravity wave energies, vertical wavelengths and momentum fluxes over the southern Andes, Drake Passage and Antarctic Peninsula using GPS radio occultation measurements. We also develop a quantitative detection methodology for the automated identification of gravity waves in vertical profiles. Our goal is to better understand the sources and formation of a belt of increased gravity wave energy observed around the edge of the southern stratospheric polar vortex.

Chapter 5
We develop a new analysis method for the measurement of gravity wave amplitudes, horizontal wavelengths and direction of propagation from two-dimensional stratospheric temperature measurements using a two-dimensional Stockwell Transform (S-transform). We introduce alternative windowing functions for use in the

S-transform in two dimensions, and show that the use of and elliptical windowing function can greatly improve the recovery of wave amplitudes. We show that our two-dimensional S-transform method makes significant improvements on current analysis methods in terms of robustness, reliability and accuracy, particularly when our alternative window is used.

Chapter 6
We briefly summarise the findings of the studies described in this thesis.

Chapter 7
Finally, we discuss proposals for future studies.

1.4 Key results

Listed below is brief summary of the principal and, to the best of our knowledge, new results obtained during the course of this PhD that deliver a positive impact to the wider atmospheric science community.

- First application of the two-dimensional Stockwell transform (2DST) in the atmospheric sciences, and first to apply the 2DST to measurements from the satellite-borne AIRS instrument (Chapter 5).
- First use of a new elliptical spectral windowing function in the 2DST in the atmospheric sciences, significantly improving the measurement of key gravity wave parameters in 2-D data (Chapter 5).
- First clear observations of the meridional advection of stratospheric gravity waves from the southern Andes into the southern wintertime polar vortex. This is an observational confirmation of a high-resolution modelling prediction by Sato et al. (2012) (Chapter 4).
- First application of a quantitative gravity-wave detection method (Wave-ID) for one-dimensional (1-D) atmospheric temperature profiles based upon the continuous wavelet transform (CWT). Our approach provides a correlation coefficient of how wave-like the detrended measurements are by correlating them with the analysing wavelet, whereas most previous methods have used only an amplitude threshold (Chapter 4).
- First use of the Wave-ID method to reveal the distribution of gravity wave populations in and around the southern stratospheric polar vortex. We find higher concentrations of infrequent, large-amplitude waves over and around mountain ranges

that tend to dominate wave energy measurements, with more frequently-occurring smaller-amplitude waves generally found over open ocean. Our results provide new and useful insight into gravity wave sources and short-timescale variability in the region (Chapter 4).

- One of only a handful of studies to make measurements of stratospheric gravity-wave momentum flux from GPS radio occultation (GPS-RO) measurements.
- As a co-author of Wright et al. (2016), we present the first combination of 2-D and 1-D measurements from the satellite-borne AIRS and Aura-MLS instruments to obtain full three-dimensional measurements of stratospheric gravity wave parameters from space, where this author provided the 2DST analysis code and key methodological innovative steps.
- As a co-author of Wright et al. (2015a) and a follow-up study currently in preparation, this author provided COSMIC GPS-RO energy and momentum flux results, analysis and interpretation as part of this large two-part multi-instrument study.

Chapter 2

Gravity Waves

2.1 Introduction

Internal Gravity Waves (GWs) are propagating mesoscale disturbances which transport energy and momentum in fluid environments. They have long been understood to be a ubiquitous and vital component of the atmospheric system, providing a key driving mechanism for atmospheric processes in the middle and lower atmosphere through drag and diffusion processes (e.g. Fritts and Alexander, 2003, and citations therein). Of particular importance is their ability to transport momentum both horizontally and vertically far away from their sources and, upon breaking, deposit this momentum into the mean flow, resulting in a driving or drag force.

Gravity wave drag plays an integral role in a variety of key dynamical processes, such as the Quasi-Biennial Oscillation (QBO) (Dunkerton, 1997; Baldwin et al., 2001) and the Brewer-Dobson circulation (Andrews et al., 1987; Garcia and Randel, 2008) in the tropics, and Sudden Stratospheric Warmings (SSWs) at high latitudes (e.g. Liu and Roble, 2002; Wright et al., 2010; France et al., 2012).

Even small amounts of gravity wave drag can significantly impact the thermal structure in the stratosphere through downward control (Egger and Hoinka, 2005). In polar regions, this thermal forcing can have a strong effect on chemical processes that are particularly sensitive to slight changes in ambient temperature, such as the formation of polar stratospheric clouds and the subsequent conversion of chlorine into active forms, a key process in annual ozone depletion (Garcia and Boville, 1994). Höpfner et al. (2006) presented a clear link between intense orographic gravity wave activity over the Antarctic Peninsula and the formation of a belt of nitric acid trihydrate polar stratospheric clouds during

Figure 2.1: A surrealistic representation of atmospheric gravity waves (from Hines (1974)).

austral winter. Gravity wave drag is also responsible for pushing the winter stratosphere away from radiative equilibrium and forcing meridional circulation from the summer pole to the winter pole, making the summer polar mesosphere the coldest place on earth. Polar mesospheric clouds, additional factors in annual ozone depletion rates, form much more frequently in the cold summer mesosphere but not in winter, where the meridional circulation is reversed.

Gravity waves are so ubiquitous in the earth's atmosphere that Hines (1974) suggested that all dynamical perturbations from the background state at any location at a single point in time can, in a subtle analogy of Fourier theory, be considered as the superposition of an infinite number of gravity wave perturbations at that location at that time. Under this consideration, a surrealistic view of the atmospheric gravity waves (Fig. 2.1) of all scales, phases and amplitudes was proposed by Hines (1974). Although a non-scientific representation, this surrealistic viewpoint succinctly conveys the tireless complexity of atmospheric gravity wave motion.

2.1.1 Gravity waves in general circulation models

Despite decades of close observation, considerable uncertainty remains regarding the accurate representation of gravity wave sources, fluxes, propagation and variability in numerical modelling. Nearly all general circulation models (GCMs) have a systematic bias in their representation of the southern stratosphere. In the models, the breakdown of the winter polar vortex occurs too late in the year, the polar vortex winds are too strong and the polar vortex temperatures are too low. This “cold pole” bias is a long-standing problem and has been identified as a serious impediment to model progress, leading to discrepancies in properties including simulated Antarctic ozone trends and southern-hemisphere climate (e.g. McLandress et al., 2012). It has been suggested that these problems may arise due to a deficiency in stratospheric gravity wave drag at latitudes near 60°S. This deficiency may be due to waves in the real atmosphere waves propagating into this latitude belt from sources further north or south, or because the sources of gravity waves in the models under-represent the in situ generation of waves. Characterising the nature of gravity waves at latitudes near 60°S is thus a significant problem.

However, quantitatively identifying and describing the location, magnitude and short-timescale variability of gravity wave sources observationally has proved exceptionally challenging. It is perhaps for this reason that the current generation of GCMs also exhibit strong disagreement in the magnitude and distribution of the gravity wave momentum flux in the southern hemisphere stratosphere during austral winter compared to observations (Geller et al., 2013). Particularly large discrepancies are found over the mountains of the southern Andes and Antarctic Peninsula, suggesting even orographic wave drag is not simulated consistently.

For the majority of operational GCMs used in numerical weather prediction (NWP), many gravity waves are sub-gridscale phenomena and their effects must be parametrized. Parametrizations vary greatly between GCMs, but tuning parameters may for example be chosen in order to produce comparable monthly-mean zonal-mean wind fields to observations (Geller et al., 2013) or obtain a realistic quasi-biennial oscillation (QBO) (e.g. Scaife et al., 2000) while remaining physically plausible. However, a current scarcity of robust observations of key gravity wave parameters means that these parametrizations are poorly constrained (Alexander et al., 2010). With the advent of increased computing power in recent years, high spatial resolution GCMs without the need for gravity wave parametrizations are becoming available (e.g. Watanabe et al., 2008). Such high-resolution modelling studies (e.g. Sato et al., 2012) are promising, but discrepancies between observed and modelled parameters still remain. A useful review of the current state of gravity wave modelling is given by Preusse et al. (2014).

2.2 Wave motion in the atmosphere

To better understand the physical nature of gravity waves, it is useful to derive a mathematical formulation that describes their motion through the atmosphere. A dispersion relation is an expression which relates a gravity wave's spatial characteristics to its intrinsic frequency. Such a relation can be used to learn a great deal about the expected behaviour and observational constraints of gravity waves in the atmosphere.

Here we derive dispersion relations for wave motion in an idealised atmosphere, closely following the derivations of Vallis (2006) and Hines (1960).

We begin by considering the earth's atmosphere as a stratified, isothermal and compressible fluid. We also consider only a two-dimensional Cartesian framework, ignoring spherical and rotation effects.

In this framework, suppose that we have a stationary, well-mixed parcel of air at rest in hydrostatic equilibrium, such that the hydrostatic assumption

$$\frac{\delta p(z)}{\delta z} = -\rho(z)g, \tag{2.1}$$

where $p(z)$ and $\rho(z)$ are pressure and density with respect altitude z and g is acceleration due to gravity, is satisfied. Consider that this "resting" state has the state vector $\Theta(\theta, \rho, p, u, w)$, to which a perturbation is applied, resulting in the perturbed state vector $\Theta'(\frac{\theta'}{\theta}, \frac{\rho'}{\rho}, \frac{p'}{p}, u', w')$, where θ',θ, ρ',ρ, p',p, u',u and w',w are the perturbed and unperturbed quantities of potential temperature, density, pressure and horizontal and vertical wind velocities respectively.

Under linear theory, this perturbation state is related to the resting state via five fundamental equations of motion, namely

$$\rho\frac{\delta u'}{\delta t} = -\frac{\delta p'}{\delta x} \qquad \text{conservation of horizontal momentum} \tag{2.2}$$

$$\rho\frac{\delta w'}{\delta t} = -\frac{\delta p'}{\delta z} - \rho' g \qquad \text{conservation of vertical momentum} \tag{2.3}$$

$$\frac{\delta \rho'}{\delta t} + w'\frac{\delta \rho}{\delta z} = -\rho\left(\frac{\delta u'}{\delta x} + \frac{\delta w'}{\delta z}\right) \qquad \text{conservation of mass} \tag{2.4}$$

$$w'\frac{\delta \theta}{\delta z} = -\frac{\delta \theta'}{\delta t} \qquad \text{thermodynamic equation} \tag{2.5}$$

$$\frac{\theta'}{\theta} + \frac{\rho'}{\rho} = \frac{1}{\gamma}\frac{p'}{p} \qquad \text{equation of state.} \tag{2.6}$$

$$\tag{2.7}$$

Our first goal is to combine these equations such that they can be satisfied by a single wave-like solution. To do this, we first take an alternative formulation of the thermodynamic equation:

$$\frac{\delta p'}{\delta t} - w'\frac{p_0}{H} = -\gamma p\left(\frac{\delta u'}{\delta x} + \frac{\delta w'}{\delta z}\right) \tag{2.8}$$

We assume an isothermal basic state, so we have $p(z) = \rho(z)RT$, where R is the gas constant, for all z. This means pressure p and density ρ can vary smoothly with altitude as

$$p(z) = p_0 e^{\frac{-z}{H}} \qquad \text{and} \qquad \rho(z) = \rho_0 e^{\frac{-z}{H}}, \tag{2.9}$$

where p_0 and ρ_0 are pressure and density at some reference level, and $H = \frac{RT}{g}$ is the scale height of the atmosphere.

By differentiating Eqn. 2.2 with respect to t and using 2.8 above, we can write

$$\left(\frac{\delta^2}{\delta t^2} - c_s^2 \frac{\delta^2}{\delta x^2}\right) u' = c_s^2 \left(\frac{\delta}{\delta z} - \frac{1}{\gamma H}\right) \frac{\delta}{\delta x} w' \tag{2.10}$$

where $c_s^2 = \gamma RT = \frac{\gamma p}{\rho}$ is the local speed of sound squared and $\gamma = \frac{c_p}{c_v} = \frac{1}{1-R/c_p} = \frac{1}{1-\kappa}$ is the ratio of specific heats for constant pressure and volume (c_p and c_v respectively).

In a similar manner, if we differentiate Eqn. 2.3 with respect to t, substitute $\frac{\delta}{\delta t} p'$ with a rearranged form of 2.10 and combine with Eqn. 2.4, we can write

$$\left(\frac{\delta^2}{\delta t^2} - c_s^2 \left[\frac{\delta^2}{\delta z^2} - \frac{1}{H}\frac{\delta}{\delta z}\right]\right) w' = c_s^2 \left(\frac{\delta}{\delta z} - \frac{\kappa}{H}\right) \frac{\delta u'}{\delta x} \tag{2.11}$$

Here we have essentially found expressions for the rate of change of horizontal and vertical momentum with time, using the conditions imposed by the thermodynamic and conservation of mass equations. We combine the result in 2.11 with Eqn. 2.10 to eliminate u', and after some lengthy algebra we can write

$$\frac{\delta^4 w'}{\delta t^4} - c_s^2 \frac{\delta^2}{\delta t^2} \left(\frac{\delta^2}{\delta x^2} + \frac{\delta^2}{\delta z^2} - \frac{1}{H}\frac{\delta}{\delta z}\right) w' - c_s^2 \frac{\kappa g}{H} \frac{\delta^2 w'}{\delta x^2} = 0 \tag{2.12}$$

We can see that we are getting close, but we need to eliminate the single $\frac{\delta}{\delta z}$ derivative in order to avoid a complex result when we apply our wave-like solution. To do this, we assume wave amplitude increases exponentially with height, such that the perturbed vertical velocity $w' = W(x,z,t)e^{\frac{z}{2H}}$, where $W = w\left(\frac{\rho}{\rho_0}\right)^{\frac{1}{2}}$. With this, the single $\frac{\delta}{\delta z}$ is eliminated, which gives

$$\frac{\delta^4 W}{\delta t^4} - c_s^2 \frac{\delta^2}{\delta t^2} \left(\frac{\delta^2}{\delta x^2} + \frac{\delta^2}{\delta z^2} - \frac{1}{4H^2}\right) W - c_s^2 \frac{\kappa g}{H} \frac{\delta^2 W}{\delta x^2} = 0 \tag{2.13}$$

All the derivatives in the above result will produce real coefficients, which means we are now ready to try a wave-like solution. But before we do, we replace $\frac{\kappa g}{H}$ with the square of the Brunt-Väisälä frequency N^2. We do this recalling the definition of potential temperature θ, the theoretical temperature a parcel of air would have if it were brought adiabatically from some pressure p to a reference pressure p_0, and using our idealised

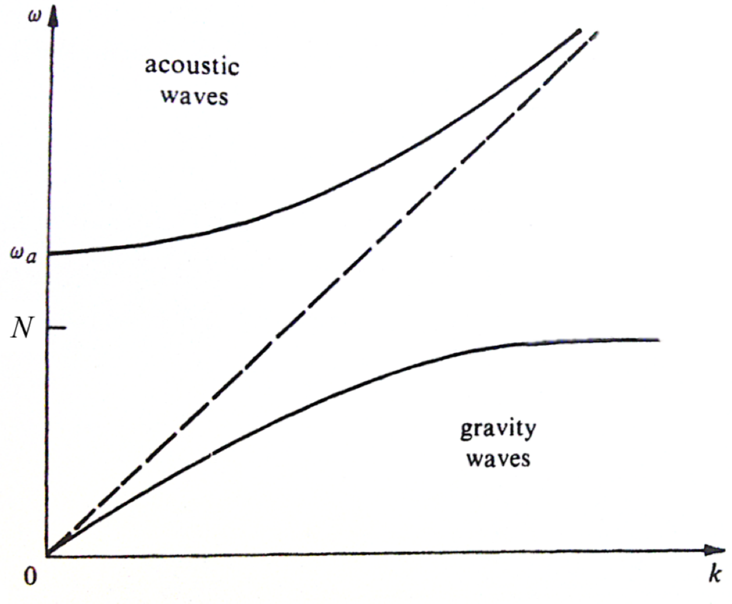

Figure 2.2: Acoustic and gravity wave dispersion curves for wave frequency ω and horizontal wavenumber k_H. Solid lines correspond to vertical wavenumber $m = 0$, ω_a is the acoustic resonance frequency and N is the Brunt-Väisälä frequency . From Houghton (1977).

pressure and density relations in Eqn. 2.9 to show that

$$\theta = T\left(\frac{p_0}{p}\right)^{\kappa} = Te^{\frac{\kappa z}{H}} \tag{2.14}$$

$$N \equiv \sqrt{\frac{g}{\theta}\frac{d\theta}{dz}} = \sqrt{\frac{g}{\theta}\frac{\kappa}{H}\theta} = \sqrt{\frac{\kappa g}{H}} \tag{2.15}$$

where $\kappa = R/c_p$ and c_p is the specific heat capacity for constant pressure. We now assume a wave-like solution for Eqn. 2.13 of the form

$$W = Ae^{i(k_H x + mz + \omega t)} \tag{2.16}$$

where $A = \Re[A]$ is the magnitude of the complex wave amplitude A, k_H and m are the horizontal and vertical wavenumbers respectively and ω is angular frequency. By substituting 2.16 into Eqn. 2.13, performing the differentiation and cancelling the exponentials, we achieve the result

$$\omega^4 - c_s^2\omega^2\left(k_H^2 + m^2 + \frac{1}{4H^2}\right) + c_s^2 N^2 k_H^2 = 0 \tag{2.17}$$

which is the dispersion relation for acoustic-gravity waves. Three kinds of waves satisfy Eqn. 2.17: acoustic (and Lamb) waves and gravity waves (Vallis, 2006). Acoustic waves rely on the compressibility of the atmosphere to provide their restoring force, whereas gravity waves are restored by buoyancy. Here we consider only gravity waves.

In order to go further and separate acoustic wave solutions from gravity wave solutions, we deviate from the derivation of Vallis (2006) and rewrite Eqn. 2.17 as

$$\omega^4 - c_s^2\omega^2\left(k_H^2 + m^2\right) - \frac{c_s^2\omega^2}{4H^2} + c_s^2N^2k_H^2 = 0, \tag{2.18}$$

Dimensional analysis requires that the $c_s^2/4H^2$ term must have units of frequency squared. If we define this as

$$\omega_A = \frac{c_s^2}{4H^2}, \tag{2.19}$$

and rearrange Eqn. 2.18 for vertical wavenumber, we can write

$$m = \sqrt{\frac{\omega^2}{c_s^2}\left(1 - \frac{\omega_A^2}{\omega^2}\right) - k_H^2\left(1 - \frac{N^2}{\omega^2}\right)}. \tag{2.20}$$

For a vertically propagating wave, m must be real. We can see from 2.20 that this only occurs when $\omega > \omega_A$ or when $\omega < N$. This requirement conveniently separates solutions for 2.17 into (i) acoustic waves (where $\omega > \omega_A$) and (ii) gravity waves (where $\omega < N$). Of course, ω_A is not an arbitrary frequency but the acoustic frequency, defined as above.

Figure 2.2, from Houghton (1977), illustrates the different regimes occupied by acoustic and gravity waves in terms of their ratio frequencies and horizontal wavenumbers.

If $\omega_A > \omega > N$, there is an imaginary or complex solution for m in Eqn. 2.20. These waves are evanescent; they may propagate horizontally but not vertically, such that their is no change in phase with height. External gravity waves, which propagate along the horizontal boundary between two fluids between two veritcally stratified fluids, are an example of such waves. These waves can be seen at the interface between adjacent atmospheric layers as a result of shear flow, or on the surface of the oceans.

The work in this thesis concerns internal gravity waves, i.e. waves with real vertical wavenumbers, so in order to acquire a gravity-wave-only dispersion relation for our idealised atmosphere, we neglect acoustic solutions for Eqn. 2.17 by dividing by c_s^2 and letting $c_s \to \infty$, after which we can write

$$\omega^2 = \frac{N^2k_H^2}{k_H^2 + m^2 + 1/4H^2} \tag{2.21}$$

This dispersion relation relates a gravity wave's spatial characteristics k_H and m to its angular frequency ω for our idealised atmosphere. Recall that our idealised atmosphere has no rotational, spherical of background flow effects. Fritts and Alexander (2003) showed that, in order to include effects of the earth's rotation, the terms $\omega^2f^2 + c_s^2f^2\left(m^2 + \frac{1}{4H^2}\right)$ must be added to the left hand side of our acoustic-gravity wave dispersion relation in

Eqn. 2.17, where $f = 2\Omega \sin\phi$ is the Coriolis parameter, Ω is the earth's angular rotation rate and ϕ is latitude. By adding these terms we now recover our gravity-wave-only dispersion relation as

$$\omega^2 = \frac{N^2 k_H^2 + f^2\left(m^2 + \frac{1}{4H^2}\right)}{k_H^2 + m^2 + \frac{1}{4H^2}} \tag{2.22}$$

We now complete our derivation by considering the full three dimensional framework, where $k_H^2 = k^2 + l^2$ where k and l are the orthogonal horizontal wavenumbers in the zonal and meridional directions respectively.

In Eqn. 2.22, we have derived gravity wave angular frequency ω for a perturbed fluid parcel that is stationary with respect to horizontal axes. If we introduce a background flow vector $(\bar{u}, \bar{v})$ in the zonal and meridional directions respectively, we see that our derived frequency in 2.22 is the frequency of the wave in its own inertial reference frame, the intrinsic frequency $\hat{\omega} = \omega_0 - k\bar{u} - l\bar{v}$, where ω_0 is the ground-based frequency for the wave, as measured by an observer at rest on the ground.

The gravity wave dispersion relation now takes its more familiar form as derived by Fritts and Alexander (2003):

$$\hat{\omega}^2 = \frac{N^2(k^2 + l^2) + f^2\left(m^2 + \frac{1}{4H^2}\right)}{k^2 + l^2 + m^2 + \frac{1}{4H^2}}. \tag{2.23}$$

2.3 The gravity wave spectrum

The spatial characteristics of a gravity wave packet are well-governed by the relation in 2.23. Thus, broadly speaking, gravity waves can be divided into three main groups according to the value of their intrinsic frequency $\hat{\omega}$, namely: high-frequency, mid-frequency and low-frequency (inertia-) gravity waves. Our discussions below closely follow the review of Fritts and Alexander (2003).

High-frequency waves

For high-frequency waves, where $\hat{\omega} \gg f$ or $\hat{\omega} \to N$, the Coriolis parameter f term in 2.23 becomes negligible such that the dispersion relation can be simplified to

$$\hat{\omega}^2 \approx \frac{N^2(k^2 + l^2)}{k^2 + l^2 + m^2}. \tag{2.24}$$

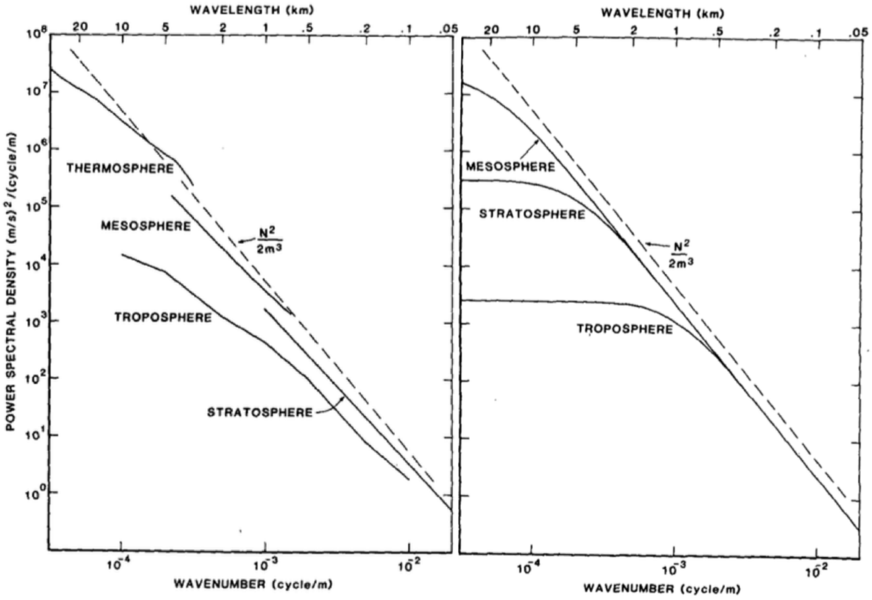

Figure 2.3: Power spectral density of horizontal wind velocity against vertical wavenumber m for (left) observed and (right) modelled gravity waves, suggesting a "universal" saturated spectrum. Values for N^2 in the right panel have been scaled to appropriate values for the troposphere, stratosphere and mesosphere. From Smith et al. (1987).

for vertical wavenumbers $\mathrm{m}^2 \gg 1/4H^2$ (vertical wavelengths shorter than about 60 km). Gravity wave group velocities $\hat{c}_{gx}$, $\hat{c}_{gy}$ and $\hat{c}_{gz}$ in the zonal, meridional and vertical directions are given by

$$(c_{gx}, c_{gy}, c_{gz}) = \left(\frac{\delta \hat{\omega}}{\delta k}, \frac{\delta \hat{\omega}}{\delta l}, \frac{\delta \hat{\omega}}{\delta m} \right) + (\bar{u}, \bar{v}, \bar{w}), \tag{2.25}$$

and vertical group propagation angle for a gravity wave packet, defined with respect to the horizontal plane, is given by

$$\varphi = \tan^{-1} \left(\frac{c_{gz}}{\sqrt{c_{gx}^2 + c_{gy}^2}} \right) \tag{2.26}$$

For the high-frequency approximation we made in 2.24, we see that that such waves have large vertical group velocities, propagating very close to the vertical (recalling in our sign convention that vertical wavenumber m is negative for an upwardly propagating wave). These high-frequency gravity waves are very commonly found in strong winds directly above mountain ranges, where they can propagate to great heights. Such waves are discussed further in section 2.7.1.

Figure 2.4 shows a schematic diagram of a generalised, upwardly-propagating high-frequency gravity wave packet as dispersed from a gravity wave source in a stable, stratified atmosphere with background flow $\bar{u} = 0$. Energy propagation is aligned with group propagation close to the vertical, while phase propagation is orthogonal and downward.

Mid-frequency waves

A large portion of the gravity wave spectrum falls in within the mid-frequency range, where $N \gg \hat{\omega} \gg f$. For these waves, the factor $1/4H^2$ and the inertial frequency are negligible, $m^2 \gg (k^2 + l^2)$ such that $\left(1 - \frac{m^2}{k^2+l^2}\right) \sim \frac{-m^2}{k^2+l^2}$, such that Eqn. 2.23 can be simplified to

$$\hat{\omega}^2 \approx \left| \frac{N^2(k^2 + l^2)}{m^2} \right|. \tag{2.27}$$

Taking this simplified form allows valuable insight into the characteristics of a large subset of gravity waves. In particular, we can infer generalised behavioural changes that result from changes in background wind and stability, with implications relating to wave breaking and momentum deposition mechanisms. Such mechanisms are discussed in section 2.6.

Low-frequency (inertia-gravity) waves

For gravity waves with very low intrinsic frequencies, where $\hat{\omega} \to f$, the effect of the earth's rotation (i.e. inertia) becomes important. As above, the approximations $\left(1 - \frac{m^2}{k^2+l^2}\right) \sim \frac{-m^2}{k^2+l^2}$ and $m^2 \gg 1/4H^2$ can once again be made, but terms in 2.23 involving inertial frequency f are no longer negligible, such that the dispersion relation simplifies to

$$\hat{\omega}^2 \approx \frac{N^2(k^2 + l^2)}{m^2} + f^2. \tag{2.28}$$

A generalised low-frequency, inertia-gravity wave packet is shown in Fig. 2.4. Vertical propagation is slow for these waves; they have long dwell-times in the lower and middle atmosphere. Thus for pseudo-instantaneous measurements made by instruments sensitive to these waves, their likelihood of detection is increased. Furthermore, horizontal group propagation is most significant for inertia-gravity waves, meaning they can transport energy and momentum to large horizontal distances from their sources before breaking.

2.3.1 The observational filter

The exact region of the gravity wave spectrum within which an instrument can detect gravity waves is determined by the horizontal and vertical resolution capabilities of the given instrument (e.g. Alexander, 1998; Alexander and Barnet, 2007; Alexander et al.,

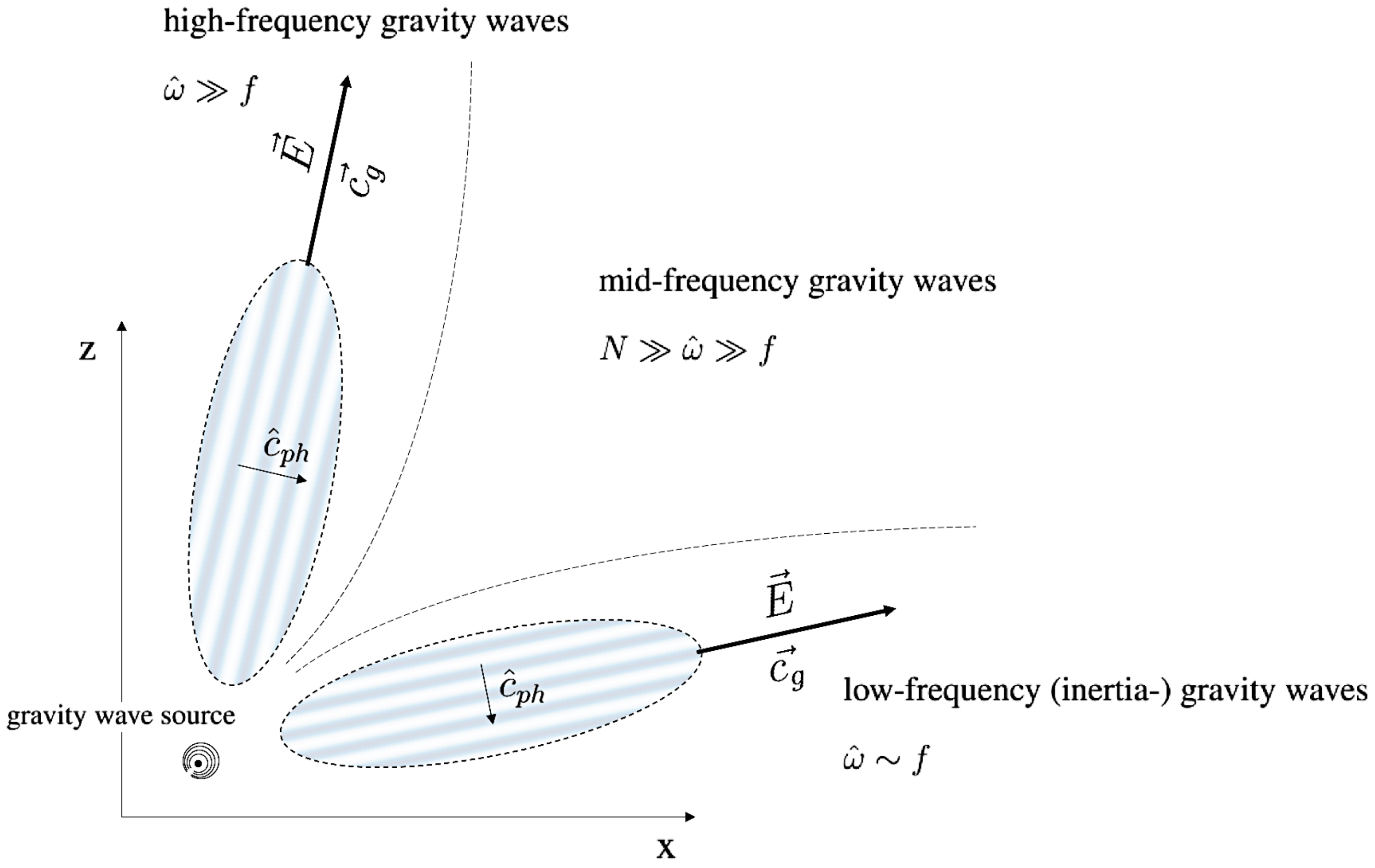

Figure 2.4: Sketch of different frequency regimes of wave packets in a generic non-orographic gravity wave field. Black arrows show directions of energy propagation $\vec{E}$, group velocity $\vec{c_g}^2 = (\delta\omega/\delta k_H)^2 + (\delta\omega/\delta m)^2$ and phase velocity $\hat{c}_{ph}^2 = (\hat{\omega}^2/k_H^2 + \hat{\omega}^2/m^2)$ for a high- and a low-intrinsic frequency $\hat{\omega}$ gravity wave packet. Figure 3.2 shows a similar sketch, but for wave packets in a generic orographic wave field.

2010; Wright et al., 2015a). This “observational filter” effect is one of the reasons why atmospheric gravity wave study has proved so challenging. Currently, no single instrument can measure the whole gravity wave spectrum. Figure 2.5 shows a cross-section of modelled temperature perturbations resulting from a convective gravity wave source. This is a more developed version of the generalised schematic diagram in Fig. 2.4. In order for an instrument to detect a gravity wave, its cigar-shaped weighting function must fit between adjacent wavefronts. Limb sounders (with the exception of UARS-MLS) are generally sensitive to low-frequency, longer horizontal wavelength waves while nadir sounders are more sensitive to higher frequency, short horizontal wavelengths.

2.4 Potential Energy

Gravity waves are an important mechanism for energy transportation between different regions and layers in the atmosphere. As such, measurement of gravity wave energies is not only of importance to circulation and mixing but also a good proxy for gravity wave activity.

Total gravity wave energy E_T (e.g. Tsuda et al., 2000) per unit mass is defined as

$$\begin{aligned} E_T &= E_K + E_p \\ &= \frac{1}{2}\left[u'^2 + v'^2 + w'^2\right] \quad + \quad \frac{g^2}{2N^2}\left(\frac{T'}{\bar{T}}\right)^2 \end{aligned} \tag{2.29}$$

where E_K and E_p are kinetic and potential energies, u', v' and w' are perturbation wind velocities in the zonal, meridional and vertical respectively and T' is temperature perturbation from a background state $\bar{T}$. Strictly, the definition for the total energy of a monochromatic gravity wave requires these perturbation measurements to be averaged over an integer number of wavecycles, such that they may correspond to mean wave amplitude. This is often not possible, which is where the instantaneous amplitude measurements, as a function of frequency, derived from by spectral analysis methods such as the Stockwell transform (Chapter 5), can prove useful.

Measurement of total gravity wave energy require simultaneous, and thus technically challenging, measurement of wind speeds and temperature. Fortunately, u', v', w' and T' are coupled via polarisation relations (e.g. Fritts and Alexander, 2003), and linear theory predicts the ratio of kinetic to potential energy E_K/E_p to be constant (e.g. VanZandt, 1985). Thus E_T can, if necessary, be estimated from measurements of gravity wave temperature perturbations only (Tsuda et al., 2000). This is useful, as it implies that any measurement technique or instrument that can measure temperature as a function of time

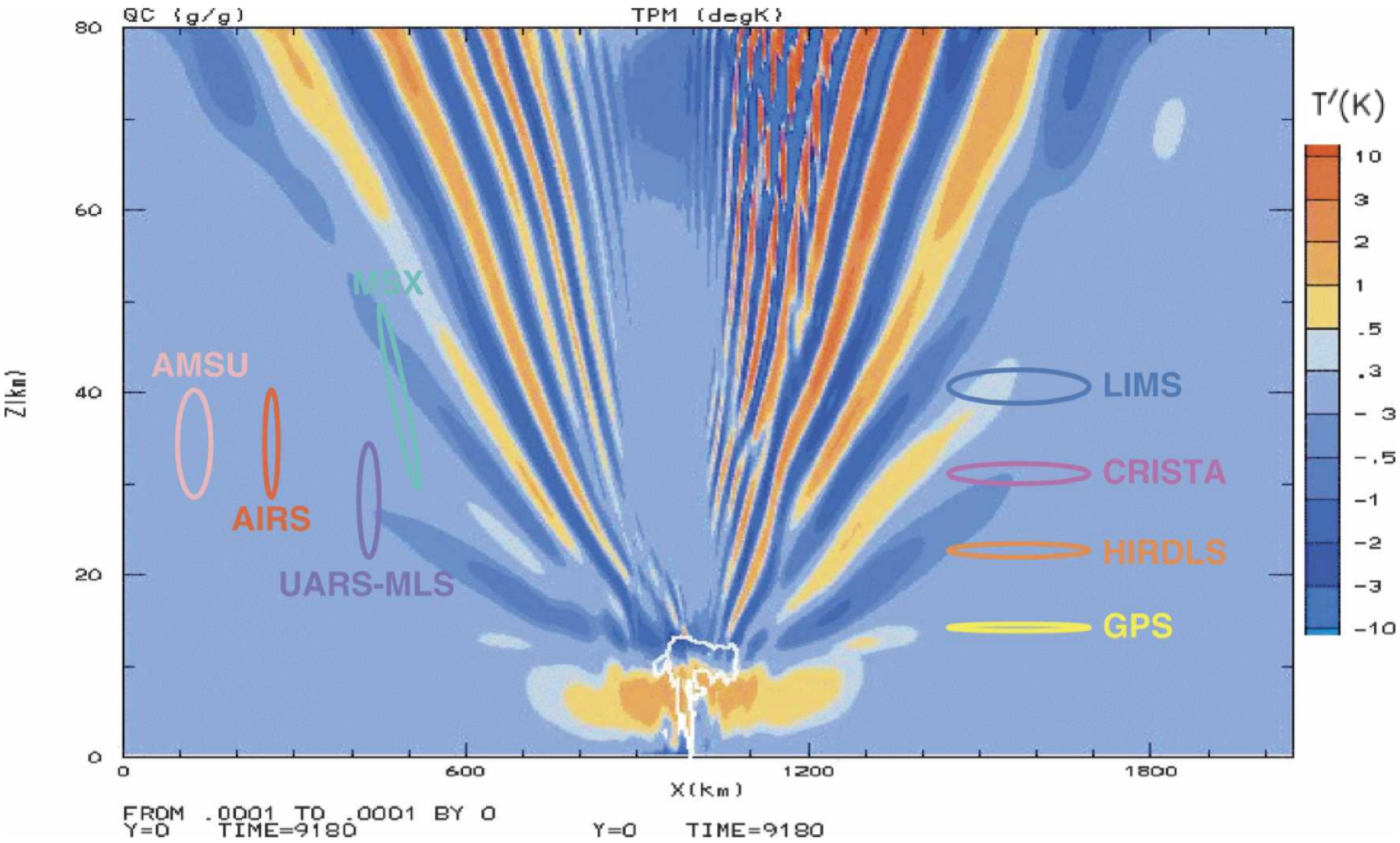

Figure 2.5: Cross-section of modelled temperature perturbations (colours, units of Kelvin (K)) resulting from a convective gravity wave source (Holton and Alexander, 1999) at bottom-centre of the image (from Alexander and Barnet, 2007). Coloured cigars show the approximate size of the weighting functions of selected spaceborne instruments (from left to right, pink – AMSU; red – AIRS; purple – UARS-MLS; cyan – MSX; blue – LIMS; magenta – CRISTA; orange – HIRDLS; yellow – GPS-RO). For an instrument to be able to resolve a gravity wave in the generalised field above, its cigar must fit between adjacent wavefronts in both the horizontal and vertical directions. For example, the limb-sounding instruments on the right are able to resolve waves with long horizontal wavelengths, but not the short horizontal wavelengths found directly above the source.

or distance can, in theory, be used to detect and estimate gravity wave energies.

2.5 Momentum flux

One of the key properties of gravity waves is their ability to transport horizontal momentum vertically, depositing it into the mean flow upon breaking.

Momentum flux or, more specifically, the vertical flux of horizontal pseudo-momentum, is an important coupling mechanism between the troposphere, stratosphere and mesosphere. In the case of ascending waves, decreasing density with altitude means that even low levels of tropospheric momentum flux can have far-reaching effects on the circulation of the middle atmosphere if transported and deposited by gravity waves. For gravity waves propagating against the mean flow in the horizontal direction, this exerts an acceleration (usually a “drag” force) upon circulation.

Unlike potential energy described above, gravity wave momentum flux is a vector quantity, and can be defined (e.g. Fritts and Alexander, 2003) in terms of zonal, meridonal and vertical wind perturbations u', v' and w' as

$$(F_x, F_y) = \bar{\rho}\left(1 - \frac{f^2}{\hat{\omega}^2}\right)\left(\overline{u'w'}, \overline{v'w'}\right). \tag{2.30}$$

where $\overline{\rho}$ is the mean background density, F_x and F_y are momentum flux in the zonal and meridional directions respectively, and overbars denotes the time average of wind perturbations and density over one wave period.

The definition in 2.30 implies that, in order to precisely measure a gravity wave’s momentum flux, three-dimensional wind perturbation measurements over a full wave cycle are required. As with the measurement of gravity wave energies above, this is very technically challenging and in many cases not possible using the current generation of instruments. By using the polarisation relations of Fritts and Alexander (2003) and the mid-frequency approximation defined above, Ern et al. (2004) showed that a good approximation for gravity wave momentum flux can be derived in terms of more easily measurable quantities: temperature perturbations T' and horizontal and vertical wavenumbers k_H and m. This derivation is outlined below.

We begin by recalling our definition for a wavelike perturbation state $\Theta'(z)$ with increasing amplitude with altitude as $e^{\frac{z}{2H}}$, introduced above:

$$\Theta'(z) = \left(u', v', w', \frac{\theta'}{\overline{\theta}}, \frac{p'}{\bar{p}}, \frac{\rho'}{\bar{\rho}} \right) e^{\left(i(kx+ly+mz-\hat{\omega}t) + \frac{z}{2H} \right)} \tag{2.31}$$

Following the form in Eqn. 2.30, we now take the product of the real parts of the time averages of the zonal, meridional and vertical wind perturbations u', v' and w' to obtain

$$F_x = \bar{\rho} \left(1 - \frac{f^2}{\hat{\omega}^2} \right) \frac{1}{\tau} \int_0^{\tau} \Re\{u'\}\Re\{w'\} dt \tag{2.32}$$

$$F_y = \bar{\rho} \left(1 - \frac{f^2}{\hat{\omega}^2} \right) \frac{1}{\tau} \int_0^{\tau} \Re\{v'\}\Re\{w'\} dt \tag{2.33}$$

where $\tau = 2\pi/\omega$ is the period of the wave. Evaluating integration over one wave period then gives us

$$F_x = \frac{\bar{\rho}}{2} \left(1 - \frac{f^2}{\hat{\omega}^2} \right) \Re\{uw^* e^{z/H}\} \tag{2.34}$$

$$F_y = \frac{\bar{\rho}}{2} \left(1 - \frac{f^2}{\hat{\omega}^2} \right) \Re\{vw^* e^{z/H}\} \tag{2.35}$$

$$\tag{2.36}$$

where w^* is the complex conjugate of background vertical wind speed w. Next we convert the wind perturbation terms above into functions of intrinsic frequency $\hat{\omega}$ and potential temperature θ perturbations via the polarisation relations given by Fritts and Alexander (2003)

$$-i\hat{\omega}\frac{\theta'}{\overline{\theta}} + \frac{N^2}{g} w' = 0 \tag{2.37a}$$

$$p' = \left(\frac{\hat{\omega}^2 - f^2}{\hat{\omega} + ifl} \right) u' = \left(\frac{\hat{\omega}^2 - f^2}{\hat{\omega}l - ifk} \right) v' \tag{2.37b}$$

and the expression

$$w' = \frac{-\hat{\omega}}{N^2 - \hat{\omega}^2} \left(m + i \left(\frac{1}{2H} - \frac{g}{c_s^2} \right) \right) p'. \tag{2.38}$$

After some algebra (Ern et al., 2004) we can write:

$$(F_x, F_y) = \left(-1 + \frac{\hat{\omega}^2}{N^2} \right) \left(\frac{1 + \frac{fl}{mk\hat{\omega}} \left(\frac{1}{2H} - \frac{g}{c_s^2} \right)}{1 + \frac{1}{m^2} \left(\frac{1}{2H} - \frac{g}{c_s^2} \right)^2} \right) \times ... \tag{2.39}$$

$$\times \frac{\bar{\rho}}{2} \left(\frac{g}{N} \right)^2 \left(\frac{\theta'}{\theta} \right)^2 \left(\frac{k}{m}, \frac{l}{m} \right) e^{\left(\frac{z}{H} \right)} \tag{2.40}$$

$$\tag{2.41}$$

Fortunately, the two large terms in brackets on the top row of this precise representation describe the deviation from the mid-frequency approximation of the value found by the terms in the bottom row (Ern et al., 2004). Hence, by taking the mid-frequency approximation we can take the bottom row of terms only. Ern et al. (2004) showed that momentum flux values calculated using this approximation deviated by around 10% from values found using the full form. The exponential term $e^{\frac{z}{H}}$ describes the increase of amplitude with altitude, hence it can be neglected if we only consider momentum flux measurements at a single height. If large vertical averages of flux are required, this factor should be included in order to avoid biases.

Since the magnitude of momentum flux $|F_{x,y}|$ is given by

$$|F_{x,y}| = \sqrt{F_x^2 + F_y^2}, \tag{2.42}$$

and fractional perturbations in potential temperature θ correspond exactly to fractional perturbations in temperature T under normal conditions as

$$\frac{\theta'}{\bar{\theta}} = \frac{T' e^{\frac{\kappa z}{H}}}{\bar{T} e^{\frac{\kappa z}{H}}} = \frac{T'}{\bar{T}} \tag{2.43}$$

we can rewrite 2.39 in the somewhat more practical form:

$$(F_x, F_y) = \frac{1}{2}\bar{\rho}\frac{g^2}{N^2}\left(\frac{T'}{\bar{T}}\right)^2\left(\frac{k}{m}, \frac{l}{m}\right) \tag{2.44}$$

and express the magnitude of momentum flux as

$$|F_{x,y}| = \frac{1}{2}\bar{\rho}\frac{g^2}{N^2}\left|\frac{k_H}{m}\right|\left(\frac{T'}{\bar{T}}\right)^2. \tag{2.45}$$

The units momentum flux are force per unit area; i.e. Pascals. We can see from 2.45 above that absolute values of momentum flux scale proportionally with vertical wavelength and inversely with horizontal wavelength. High-frequency gravity waves, such as those found directly above sources as shown in Figures 2.4 and 2.5, generally carry much larger momentum fluxes higher and faster into the middle atmosphere than their low-frequency counterparts. In the real atmosphere, typical values for gravity wave momentum fluxes can range from a few ones to a few hundreds of millipascals (mPa).

2.6 Wave breaking

Analogous to the breaking of ocean surface waves upon a tropical shore, gravity waves break when they encounter atmospheric conditions that force them to become unstable.

For any vertically propagating gravity wave, there exist horizontal atmospheric levels through which the wave cannot propagate. As waves approach these "critical levels", changing atmospheric conditions mean that energy and/or momentum can no longer be conserved within the wave-like motion and the wave becomes unstable. At this point, the intrinsic frequency of the wave asymptotically approaches zero or infinity and the wave "breaks".

The most common kind of critical level encountered by vertically propagating waves in the real atmosphere involves a rapid decrease or increase of horizontal wind speed. When the ground-based horizontal phase speed c_h equals the wind speed $\bar{U}$ in both magnitude and direction, the wave is absorbed. Waves asymptotically approach these critical levels until the spatial characteristics exceed what can be stably supported by the local atmospheric conditions, at which point they break. Generally speaking, internal gravity waves usually break via convective instability, dynamic (shear) instability, wave-wave interaction or some combination of these processes, though latter two may not necessarily involve a critical level. A useful review of gravity wave breaking mechanisms is provided by Fritts and Rastogi (1985).

2.6.1 Convective instability

Figure 2.6 shows a upwardly propagating wave approaching a critical level, marked by the dashed red line. Here, horizontal wind velocity (light blue arrows) is decreasing with height, meaning that the wave is propagating "downshear" towards the critical level. As the intrinsic phase speed decreases, the vertical wavelength and intrinsic frequency decrease such that $\lambda_z \to 0$ and $m \to \infty$ at the critical level. For waves propagating "upshear", vertical wavelengths increase with increasing height.

Wave amplitude T' also increases with altitude as $T'_2 = T'_1 e^{\frac{z_2 - z_1}{2H}}$. This, in addition to the decrease in vertical wavelength, means that the phase fronts of the wave become so tightly packed with such large amplitudes that they start to exceed the dry adiabatic lapse rate Γ_d, given as

$$\Gamma_d = -\frac{dT}{dz} = \frac{g}{c_p} \approx 9.8\,\text{K/km} \tag{2.46}$$

At this point, the wave becomes super-adiabatic, and thus convectively unstable. In Fig. 2.6, diagonal dotted lines show the dry adiabatic lapse rate Γ_d. The wave ascends (solid black line), with increasing amplitude and decreasing λ_z until it becomes super-adiabatic (dashed black line) as it approaches the critical level. Large amplitude, long vertical wavelength waves are most likely to break via this mechanism.

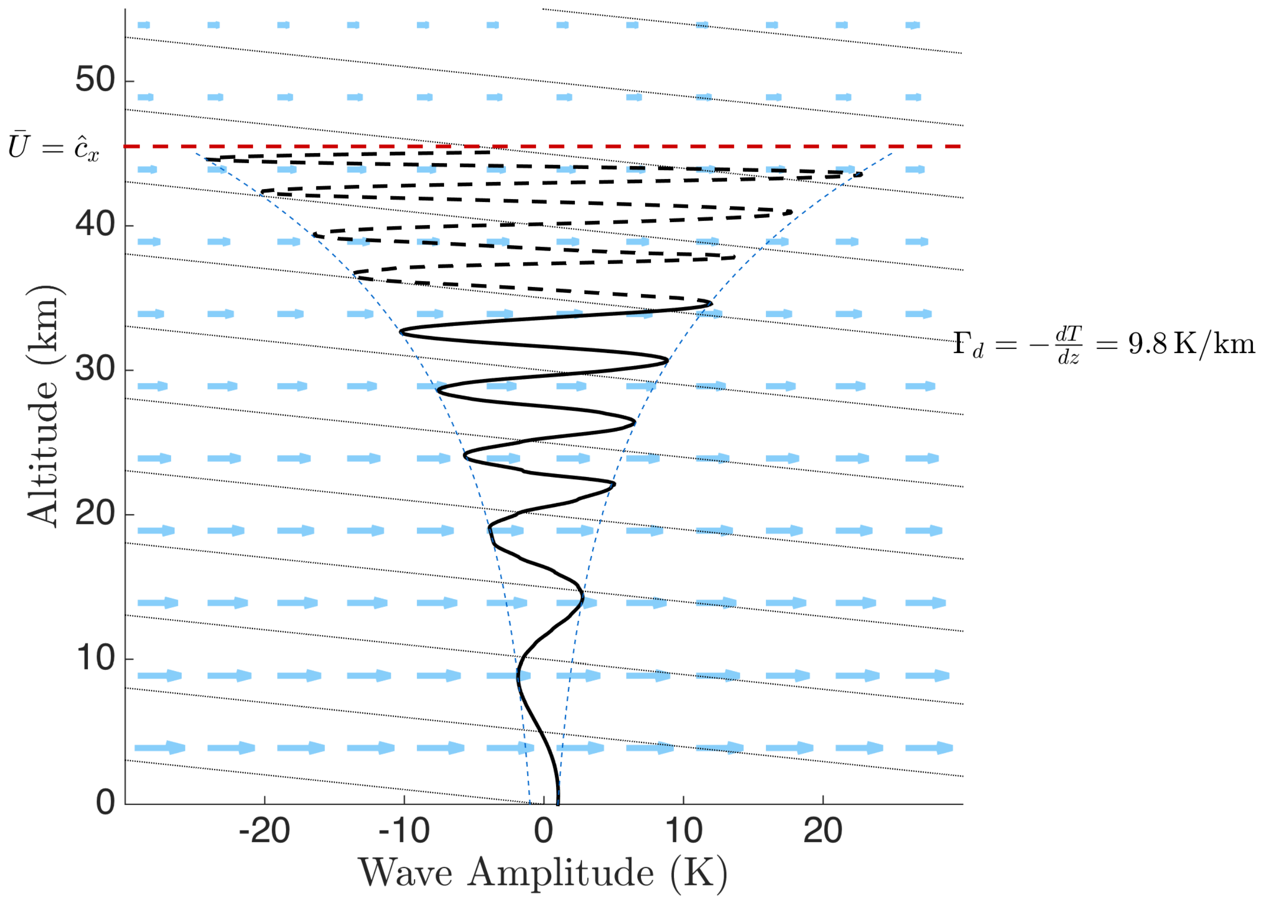

Figure 2.6: Sketch of an idealised, vertically-ascending gravity-wave packet with $k_H \sim 0$ (near-horizontal wavefronts) experiencing convective instability as it approaches a critical level (dashed red line) where the mean wind speed $\bar{U}$ (indicated by blue arrows) equals the wave's intrinsic phase speed $\hat{c}_x$. As the wave approaches the critical level, the vertical wavelength decreases such that the temperature gradient between adjacent phase fronts exceeds the dry adiabatic lapse rate Γ_d, shown by dashed grey lines, and the wave becomes unstable (dashed black line).

2.6.2 Dynamic instability

Dynamic, or shear instability, occurs when the vertical shear of horizontal wind $\delta u/\delta z$ becomes so great that the Richardson number, defined as

$$R_i = -\frac{N^2}{(\delta u/\delta z)^2} \tag{2.47}$$

falls to below $1/4$ (Fritts and Rastogi, 1985). Flow is dynamically stable for a Richardson number $R_i > 1/4$. Under this conditions, a perturbed parcel of air is able to oscillate in wave-like motion stably. If the vertical shear of horizontal wind speed is too great, the atmosphere cannot conserve energy and/or momentum under such motion and the wave becomes unstable. Figure 2.7 shows dynamic instability resulting from increasing wind shear about a fixed horizontal level (from Rosenhead, 1931). Transverse (vertical) perturbations from the equilibrium state follow a curved path as they oscillate, with the extra longitudinal component proportional to this shear in wind. If this shear is too great, wave crests "over-turn" and oscillations become unstable. This is analogous to ocean surface waves breaking as they approach the shore, where friction on the flow at the sea floor creates a shear between it and the flow at the surface, preventing further oscillation. Up to the point of breaking, a gravity wave's intrinsic frequency $\hat{\omega} \to 0$. Hence this kind of instability is more commonly encountered by large amplitude, long horizontal wavelengths waves such as inertia-gravity waves. Figure 2.8 shows a conceptual diagram of power spectral densities of oceanic internal gravity wave breaking mechanisms (from Liu et al., 2010). Although it concerns oceanic gravity waves, the analogy of such waves to atmospheric gravity waves is good. The long, flattened wave fronts in the overlaid wave pattern can be interpreted as analogous to low-frequency inertia gravity waves in the atmosphere, with relatively long horizontal wavelengths and short vertical wavelengths, for which dynamic (shear) instabilities are the primary breaking mechanism. The broader wave fronts in the centre can be interpreted as gravity waves with high-frequencies and relatively long vertical wavelengths, for which convective instability is the primary breaking mechanism. The power spectral densities on the right show the deviations in modelled distributions of breaking mechanisms for linear and non-linear theory considerations in a high-resolution modelling study. An interesting result is the significant overlap in breaking mechanisms when non-linear theory is considered.

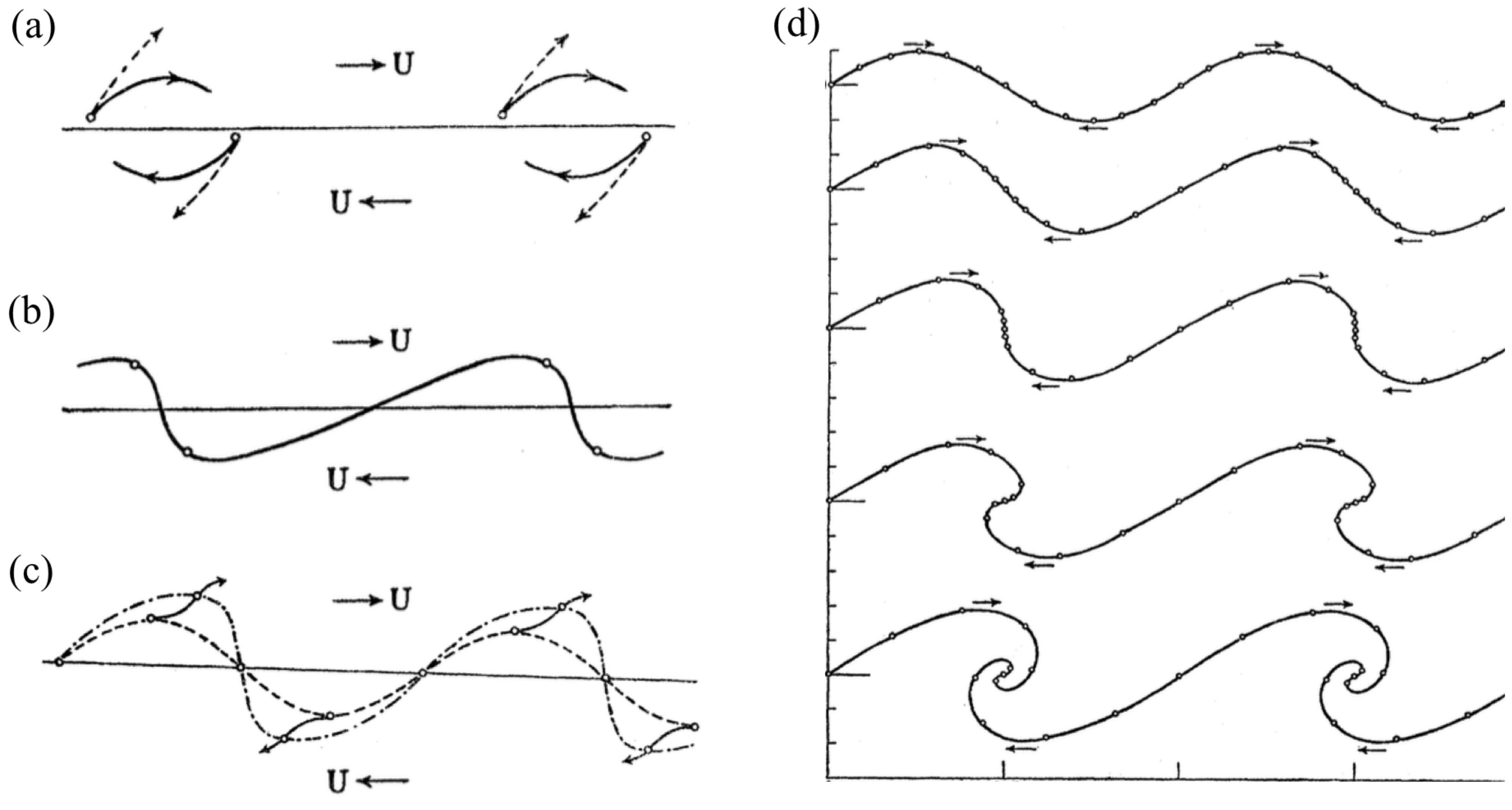

Figure 2.7: Schematics of vortex formation at the interface between two shear flows, describing dynamic instability for gravity waves propagating in a shear flow. Regular perturbations from the stratified interface (a) follow a curved path, resulting in an undulating interface (b), further perturbations to which follow even more curved paths (c) eventually resulting in instabilities and vortices at the wave crests (d). From Rosenhead (1931).

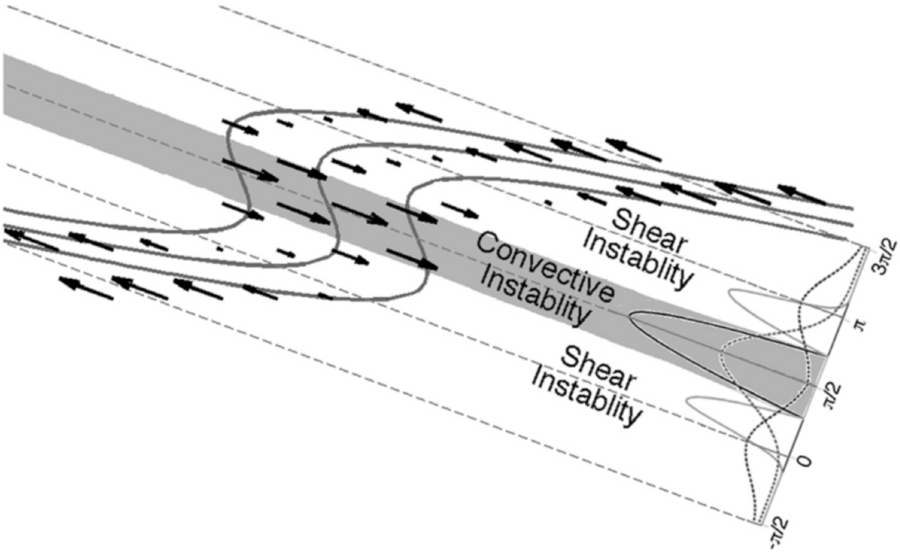

Figure 2.8: Conceptual diagram showing probability density functions (PDFs) of the general regimes of convective and shear instabilities for modelled oceanic internal gravity waves, with estimated phase fronts alignments for each case overlaid as the black solid lines. Black arrows indicate the background flow. Solid and dashed lines for PDFs indicate those calculated under linear and non-linear theory respectively. From Liu et al. (2010).

2.7 Gravity wave sources

As described in the previous section, a gravity wave is fundamentally a sinusoidal oscillation of a fluid parcel that has been perturbed from its equilibrium position by some action, the sources of which can take a very wide variety of forms.

Broadly speaking, gravity wave sources can be divided into two groups: orographic, resulting from flow over topography, and non-orographic, resulting from atmospheric convective processes, adjustment processes and/or wind shear.

2.7.1 Orographic gravity waves

Orographic gravity waves, also known as lee waves or mountain waves, occur when flow is forced over topography. Analogous to the “wake” pattern formed downstream of a pebble placed in a shallow, fast-moving stream, mountain waves occur frequently in a stable atmosphere. Viewed from above, these wave patterns generally consist of a group of large amplitude, short horizontal wavelength perturbations directly over the source and a long “wing-shaped” pattern of longer horizontal wavelengths extending downstream in a chevron pointing back upstream. Orographic gravity waves occupy the largest portion of the global gravity wave momentum budget, with strikingly intense sources tied to fixed locations (Hoffmann et al., 2013). In the two-dimensional framework, mountain waves

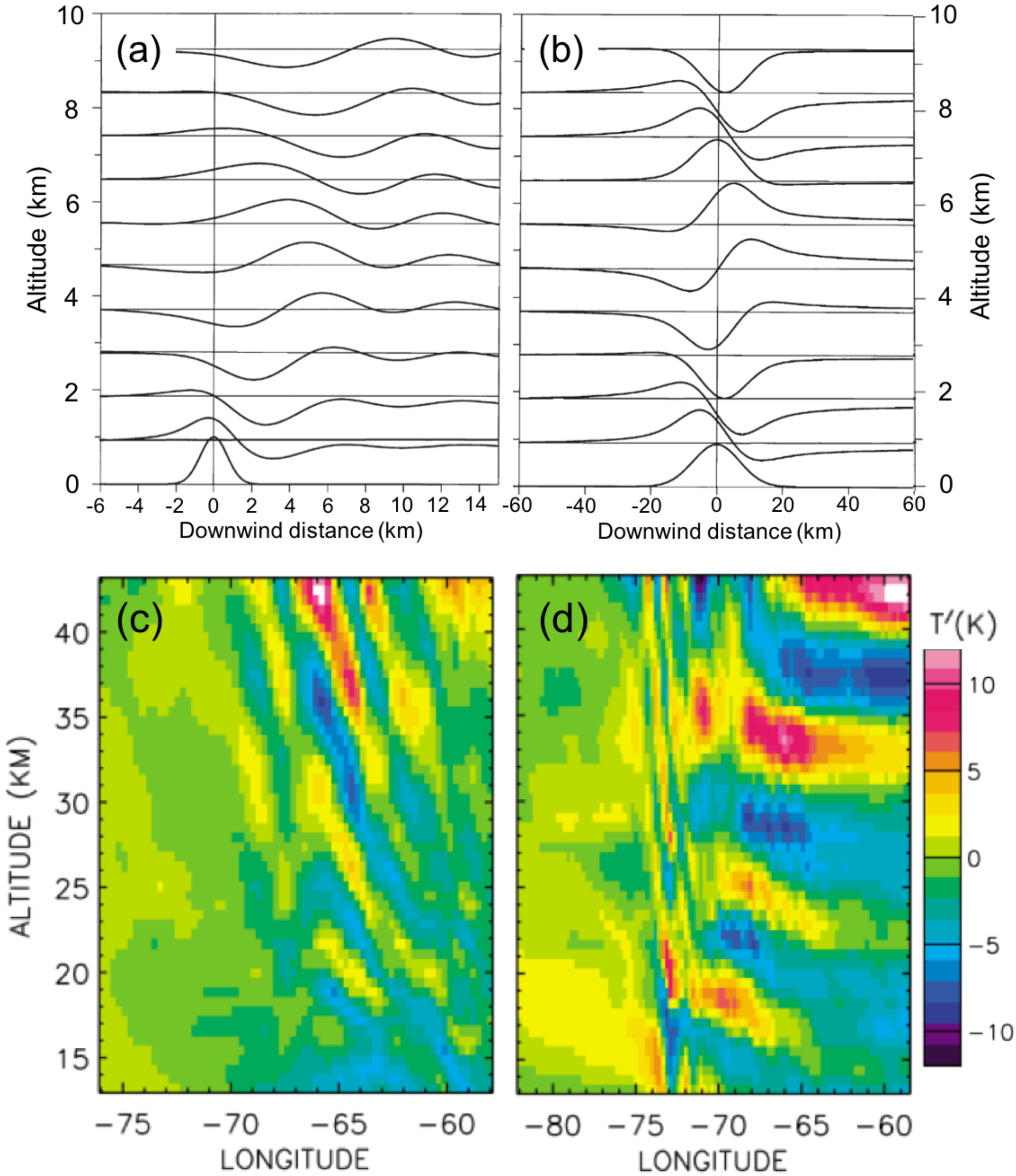

Figure 2.9: Horizontal cross-sections of mountain wave fields. Panels (a) and (b) show sketches of horizontal wind divergence resulting from flow over isolated Gaussian-shaped ridges with FWHM values of 1 and 10 km respectively, from Nappo (2013). Panels (c) and (d) show modelled mountain wave temperature perturbation fields over the Andes mountain range at around 50°S , from Alexander and Teitelbaum (2011).

are well understood. Figure 2.9 shows modelled flow patterns over a simulated mountain range (Durran, 2003; Nappo, 2013) and a model representation of the Andes mountain range (Alexander and Teitelbaum, 2011). Note that the alignment of the modelled wavefronts is quite different to the generic arrangement we have discussed in Figures 2.4 and 2.5, which more closely resemble convective gravity wave sources.

A crucial aspect of mountain waves is that they are "stationary" with respect to the ground, i.e. their intrinsic phase speed $\hat{c}_h$ is equal in magnitude but opposite in direction and sign to the horizontal wind vector $\bar{u}$ such that their ground-based horizontal phase speed c_h is zero. Likewise, their ground-based frequency $\omega = \hat{\omega} + k_H\bar{u} = 0$. Hence, these waves always propagate in the "upstream" direction against the mean flow.

This leads to a some convenient simplifications of the dispersion relation. If we consider the long vertical wavelengths found directly over the mountain wave source, such as those in Fig. 2.9d, and apply the dispersion relation for high-frequency gravity waves in Eqn. 2.24, we can write

$$\begin{aligned}
\hat{\omega}^2 = (\omega - k_H\bar{u})^2 &= \frac{N^2k_H^2}{k_H^2 + m^2} \\
(k_H^2 + m^2)(0 - k_H\bar{u})^2 &= N^2k_H^2 \\
k_H^2 + m^2 &= \frac{N^2}{\bar{u}^2} \\
m &= \pm\sqrt{\frac{N^2}{\bar{u}^2} - k_H^2} \qquad (2.48)
\end{aligned}$$

Likewise, using the mid-frequency and low-frequency approximations respectively we can write

$$m = \pm\frac{N}{\bar{u}} = \frac{N}{|\hat{c}_h|} \qquad (2.49)$$

and

$$m = \pm\sqrt{\frac{N^2}{\left(\bar{u}^2 - \frac{f^2}{k_H^2}\right)}}, \qquad (2.50)$$

which we can use to describe other portions of the mountain wave field. These relations prove useful in describing the vertical characteristics of the complete wave field by using *a priori* knowledge of the background wind vector.

2.7.2 Non-orographic gravity waves

Non-orographic wave sources may be described simply as all processes other than flow over topography that can excite gravity waves. These include a very broad range of processes including deep convection and adjustment mechanisms around jets and fronts, a useful overview of which is provided by Fritts and Alexander (2003). Although spatially diffuse, the integrated momentum flux budget of these waves can be comparable to fluxes from intense orographic sources, for example over the Southern Ocean during austral winter (e.g. Hertzog et al., 2008). In the tropics, deep convection is dominant, whereas adjustment mechanisms from jet flows and fronts dominate non-orographic gravity wave generation at mid-latitudes.

A common tropospheric source of gravity wave-inducing deep convection is the latent heat release from intense rainfall during storms. Figure 2.10, from Gong et al. (2015), shows concentric ring-shaped non-orographic gravity waves measured by the Atmospheric Infrared Sounder (AIRS) aboard the Aqua satellite. These are likely attributable to deep convection above an intense thunderstorm over the continental United States (Fig. 2.10(b)). When viewed from above, gravity wave phase fronts arranged in concentric rings like in Fig. 2.10, which radiate outward from their source over time, are very characteristic of non-orographic gravity waves. Contrarily, orographically-generated gravity waves generally have plane or chevron-shaped phase fronts (Figures 2.9 and 5.6). In the close vicinity of their sources, convectively-generated gravity waves tend to have relatively high intrinsic frequencies and relatively short horizontal scales, where flow is forced upward sharply either as a part of or flowing over an intense updraft. Such up-wellings may occur as a result of latent heat result from intense rainfall or uneven solar heating of different surface terrains, resulting in wide regions of temperature imbalance. At larger horizontal distances, horizontal scales can grow and intrinsic frequencies can decrease to form broad inertia-gravity wave fields.

In the middle atmosphere, low-frequency inertia-gravity waves can be generated as a result of spontaneous (and sometimes, but not necessarily, geostrophic) adjustment resulting from unbalanced flow around jet streams and fronts (Guest et al., 2000; Fritts and Alexander, 2003; Plougonven and Zhang, 2014). Unbalanced flows effectively "radiate" gravity waves upstream as they adjust from instabilities (O'Sullivan and Dunkerton, 1995). Historically such generation of gravity waves has proved difficult to study in modelling simulations as the common assumption of a balanced flow suppresses their formation (Plougonven and Zhang, 2014).

Crucially, horizontal phase speeds and propagation directions of non-orographic gravity waves can be very different from those permitted by orographic generation, being related only to the transient and unique process or combination of processes that excite each wave. Unlike orographic waves, which always propagate upstream against the mean flow, non-orographic waves can propagate in virtually any direction where local wind conditions permit. Although statistical information can be obtained to suggest preferential excitation

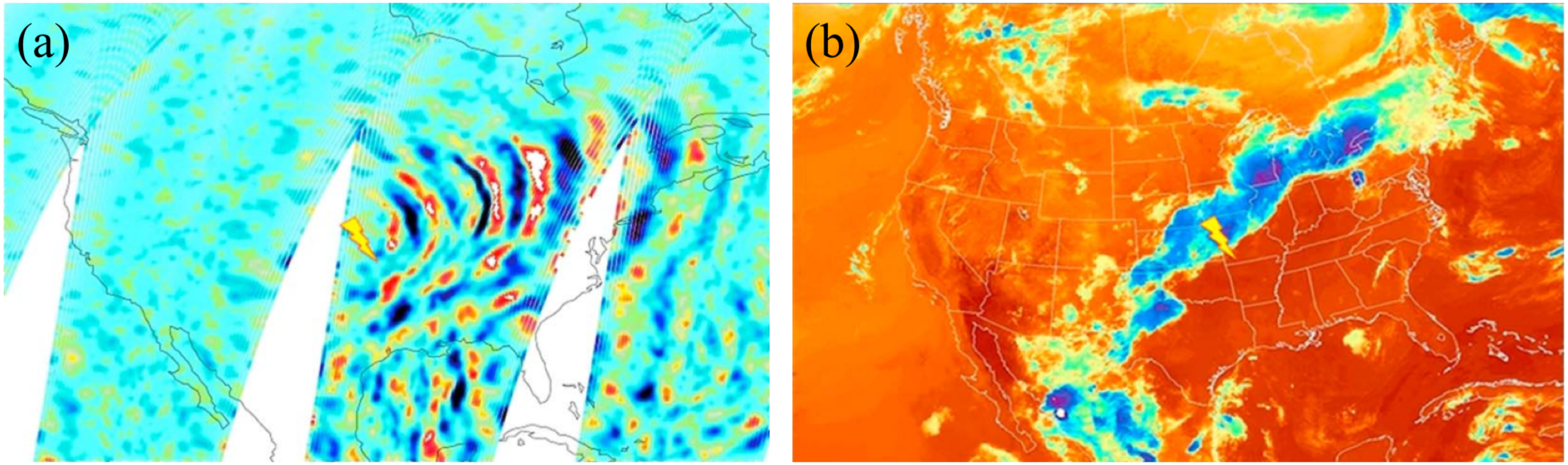

Figure 2.10: Non-orographic gravity waves revealed in AIRS (channel 75, z∼40 km) measurements (a) over the continental United States, and temporally coincident cloud-induced radiance perturbations observed by GEOS-15 (b). The estimated location of the gravity-wave source is shown by the gold thunderbolt. From Gong et al. (2015).

methods, this aspect of non-orographic gravity waves makes prediction of the impact of their momentum transport and subsequent drag/driving effects on the middle atmosphere more conceptually challenging than it is for their orographic counterparts.

Chapter 3

Instruments and Data

Overview

The principal remote-sensing data analysed in this thesis are measurements from the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) mission (Anthes et al., 2008), which uses GPS radio occultation to make vertical profiles of atmospheric temperature. We also analyse measurements from the High-Resolution Dynamics Limb-Sounder (HIRDLS) (Gille et al., 2003), a limb-sounding infrared radiometer aboard the Aura satellite, and the Atmospheric Infrared Sounder (AIRS) (Aumann et al., 2003), a multi-spectral infrared sounder aboard the Aqua satellite.

In this section we provide a brief overview of satellite remote-sensing techniques and introduce the COSMIC, HIRDLS and AIRS instruments.

3.1 Remote sensing of the Earth's atmosphere

A remote-sensing experiment is one that obtains measurements of some atmospheric quantity remotely without coming into contact with the medium in which the quantity resides. This is in contrast to *in-situ* measurements, where the quantity is measured directly in its environment by an observer in the same environment (e.g. a thermometer or anemometer at the earth's surface or aboard a radio- or rocketsonde). In the context of the atmospheric sciences, remote-sensing techniques can be roughly divided into two types: active and passive.

An active remote-sensing technique is one that involves the illumination of the atmosphere by some artificial radiation source. By measuring changes to the properties of the illuminating radiation, atmospheric properties along the ray-path can be inferred. Active remote-sensing techniques include lidar (e.g. Mitchell et al., 1991) and meteor radar (e.g. Beldon and Mitchell, 2009), which are usually ground-based or housed aboard aircraft, rather than aboard satellites, due to size and power limitations.

Passive remote-sensing techniques involve inferring atmospheric properties by measuring the natural radiance of the atmosphere, or the changes to illuminating radiation from some natural source as it travels through the atmosphere. Passive instruments are more commonly used for spaceborne remote sensing due to relatively low power-requirements and their ability to be relatively easily miniaturised.

3.1.1 Viewing geometry of spaceborne remote sensing

Satellite-based remote-sensing instruments can be broadly divided into two classes based upon their viewing geometry, each with distinct advantages and disadvantages.

Limb viewing

A limb-viewing instrument (Fig. 3.1) aboard an orbiting satellite makes observations of the atmospheric limb at an angle near-perpendicular to the radially downward direction. Observations made using this geometry have a symmetric horizontal weighting function along the line-of-sight centred at a tangent point half way through the atmospheric limb. Measurements of atmospheric properties made using this geometry are generally confined to a cigar-shaped region around this central tangent point, the size of which determines the instrument's vertical and horizontal resolutions. By adjusting the angle of the horizontal viewing direction in a scanning motion, vertical profiles of the atmosphere, centred at the tangent point, can be derived.

Figure 3.2 shows a schematic of a gravity wave field similar to that shown in Fig. 2.5 but for mountain waves. In order for an instrument to resolve a wave, the schematic representation of its cigar-shaped weighting function must be able to fit between adjacent wavefronts. Measurements made by limb-viewing instruments generally have good vertical resolutions but relatively coarse horizontal resolutions. As such, they are primarily sensitive to mid- or low-frequency (inertia-) gravity waves as shown by the red dashed line and cigar in Fig. 3.2.

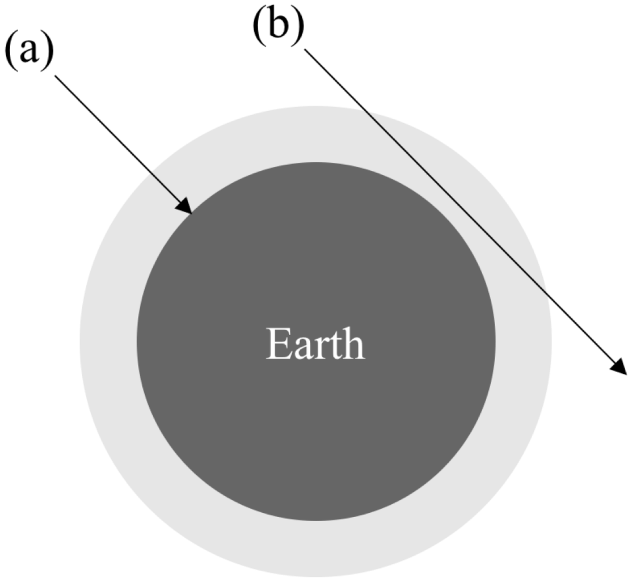

Figure 3.1: Schematic of two main viewing geometries for satellite remote-sensing: (a) nadir viewing and (b) limb-viewing. After Andrews (2000).

Nadir viewing

A nadir-viewing instrument makes observations of the atmosphere at an angle near to the radially downward direction Fig. 3.1, looking down towards the earth's surface.

Nadir-viewing instruments are usually optical or near-infrared spectral imagers or interferometers, measuring the integrated radiance of the atmosphere along a line-of-sight down to the earth's surface. By selecting different spectral channels, and thus different chemical species or temperature regimes, different weighting functions (approximately log-normal Gaussian-shaped (e.g. Wright et al., 2015a, their Fig. 1)) can be obtained which localise measurements to a specific (though quite deep) altitude region of the atmosphere. In a similar manner to the limb-viewing method, by adjusting the viewing angle from the vertical in a scanning motion, horizontal gradients in vertically measured properties can be found. As the satellite moves along its orbital path, this results in a two-dimensional “swath” of measurements.

Consequently, the horizontal resolution of nadir-viewing instruments is relatively high. The vertical resolution is relatively coarse however, due to the aforementioned depth of the vertical weighting function. As a result, nadir-viewing instruments are generally sensitive to high- or higher mid-frequency gravity waves as shown by the blue cigar-shaped weighting function in Fig. 3.2.

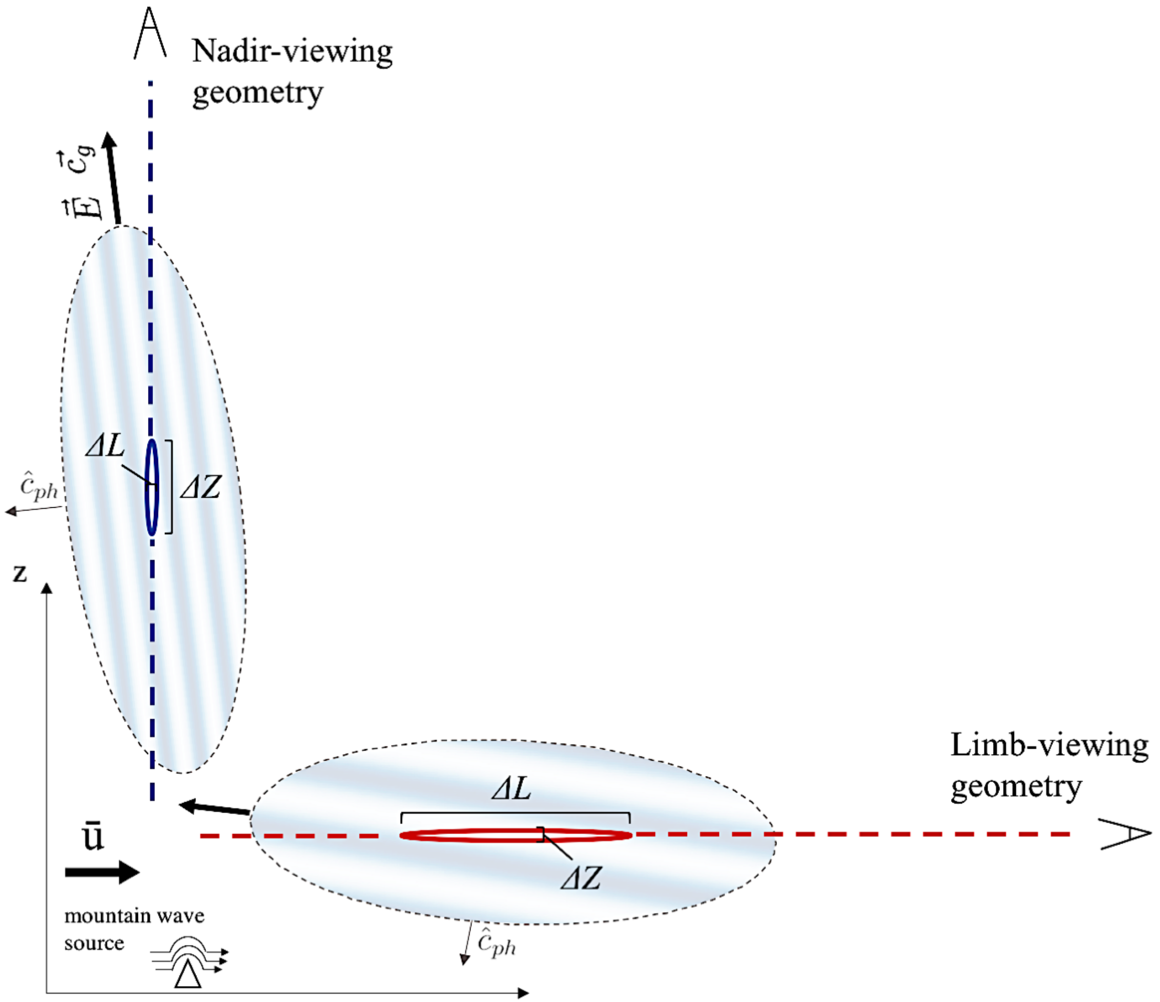

Figure 3.2: Sketch of different mountain wave packets in a generic orographic wave field. Schematic nadir- and limb-viewing geometries and weighting functions are shown by blue and red dashed lines and cigar-shapes respectively. The lengths and widths of the cigars represent the vertical resoltion ΔZ and horizontal resolution ΔL of their respective methods. As in Fig. 2.5, the cigar-shaped weighting function of an instrument must fit between adjacent wavefronts in order for a wave to be resolved. Figure 2.4 shows a similar sketch, but for wave packets in a generic non-orographic wave field.

3.2 GPS radio occultation

Radio occultation (RO) is an active limb-sounding remote sensing technique. The technique uses a radio receiver on board a low-earth-orbit (LEO) satellite to track radio signals of an emission source as it occults above or below below a planet's atmospheric limb. If the emission source is artificial, e.g. a another satellite, the occultation technique is active sounding. By contrast, solar occultation, where sunlight (a natural radiation source) is used to illuminate the atmosphere, is a passive remote-sensing technique.

As the occulting satellite rises (sets) above (beneath) the horizon, the radio signals pass through the planet's atmosphere, where they are refracted by changing refractivity gradients along the line-of-sight axis between the two satellites. Using the precisely-determined locations of the occulting and receiving satellite, together with the measured phase delay (or excess) of the radio signals, a vertical profile of atmospheric refractivity at a tangent point between the two satellites can be measured.

Radio occultation was first used in the 1960s to probe the neutral atmosphere of Mars as part of the Mariner III and IV missions (Kliore et al., 1965; Fjeldbo and Eshleman, 1965) and later as part of the Mariner V mission to Venus (Fjeldbo et al., 1971; Fjeldbo and Eshleman, 1969). Since then, radio occultation been used to probe the atmospheres of nearly every planet in the solar system, and many of their moons. With the advent of the global positioning system (GPS) in the 1970s and 1980s onwards, a radio occultation method using the GPS satellite constellation (GPS-RO) became a viable atmospheric measurement technique for the earth.

In 1995, the proof-of-concept GPS-RO mission GPS/MET (Kursinski et al., 1996; Rocken et al., 1997; Hajj and Romans, 1998; Schreiner et al., 1999; Steiner et al., 1999) flew aboard the NASA MicroLab 1 spacecraft with a single forward-facing antenna, measuring $\sim$100,000 occultations during its lifetime. In the following years, the Argentine Satelite de Aplicaciones Cientificas-C (SAC-C) (Hajj et al., 2004), the Challenging Microsatellite Payload (CHAMP) (Wickert et al., 2001) and the Gravity Recovery and Climate Experiment (GRACE) (Beyerle et al., 2005) were launched, featured both forward and rearward facing GPS-receiving antennae, each measuring many hundreds of refractivity profiles per day by tracking both rising and setting occultations.

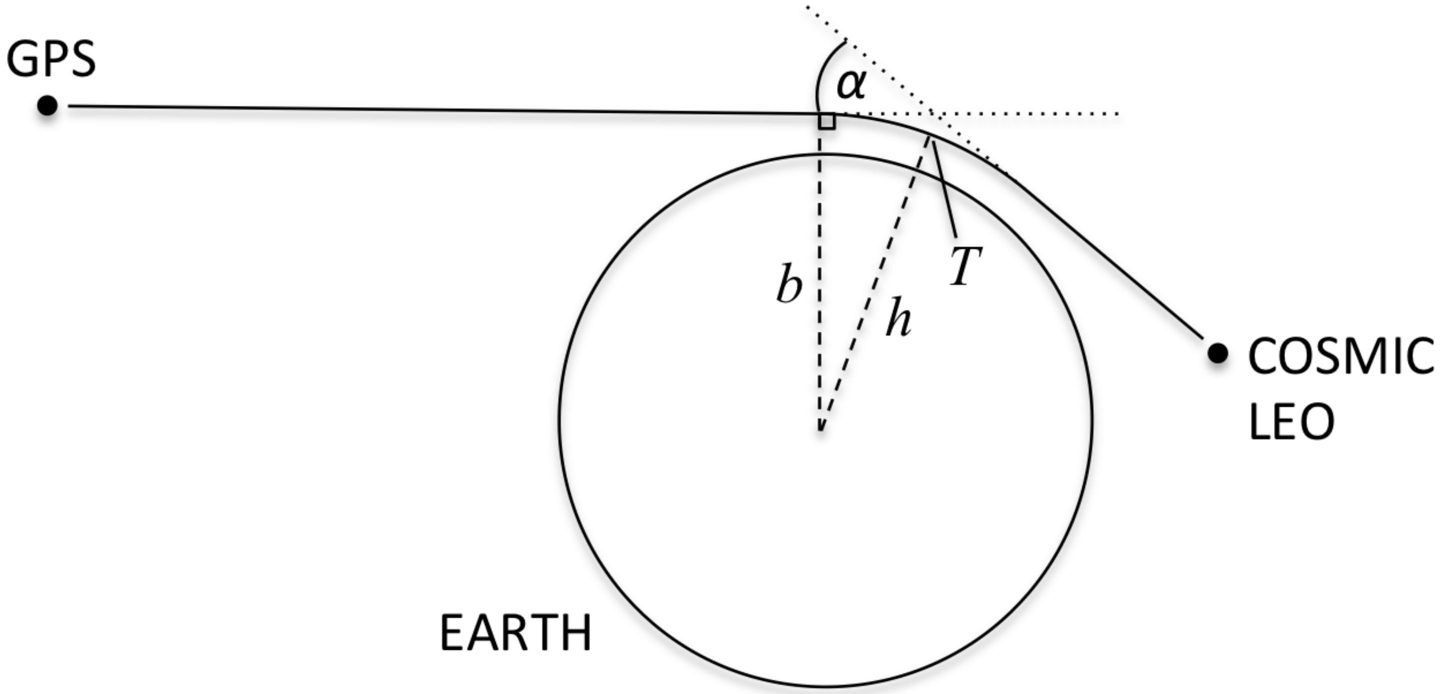

Figure 3.3: Schematic of the geometry of a COSMIC satellite tracking an occulting GPS satellite, with impact parameter b, bending angle α and radius h to the ray periapsis tangent point T. After Kursinski et al. (1996).

3.3 COSMIC

The Constellation Observing System for Meteorology, Ionosphere and Climate/Formosa Satellite 3 (COSMIC/FORMOSAT-3) (Rocken et al., 2000; Schreiner et al., 2007; Liou et al., 2007) is the world's first dedicated GPS-RO constellation. Launched in 2006, the space segment of the joint US/Taiwan mission consists of six identical satellites with a final orbital altitude $\sim$800 km, orbital period $\sim$90 minutes, 72° inclination and 30° separation. Tracking both rising and setting occultations, the constellation measures approximately 1000-2000 pseudo-randomly distributed refractivity profiles per day.

The design of each COSMIC satellite (Fig. 3.4) is morphologically similar to NASAs MicroLab I that housed GPS/MET, which in turn was adapted from JPLs TurboRogue geodetic receiver. Each satellite is a flattened cylinder, with both flat faces opening outwards 90 to face radially away from the Earths surface, with solar panels mounted on their outer face. Each satellite houses three science payloads: (1) the Integrated GPS Occultation Receiver (IGOR), (2) the Tiny Ionospheric Photometer (TIP), and (3) the TriBand Beacon transmitters (TBB) (e.g. Rocken et al., 2000). The TIP instrument points radially downward, providing measurements horizontal gradients of electron density at the peak of the F/F2 layer along the satellites path. The TBB instrument transmits phase data at 150, 400 and 1067 MHz to ground receiving stations worldwide, which is converted into line-of-sight Total Electron Content (TEC) counts for ionospheric observations.

The IGOR features two high-gain limb-viewing antennas for GPS tracking and subsequent atmospheric profiling, and two single patch antennas at a 15 degree elevation for 1Hz ionospheric profiling and precise orbit determination (Schreiner et al., 2007). Each satellite tracks the L1 and L2 GPS signals at 1575.42 MHz and 1227.6 MHz respectively, conducts some basic buffering before transmitting the data back to earth at one of two high-latitude receiving stations. Occultation profiles are processed and archived at the COSMIC Data And Archive Center (CDAAC) in Boulder, Colorado, with a secondary archive centre based in Taiwan.

A basic overview of the operational temperature retrieval employed at CDAAC is outlined in Fig. 3.5 (Derber, 2014). During an occultation event, each COSMIC satellite measures the amplitude and phase of the L1 and L2 signals from the occulting GPS satellite (Fig. 3.3), which are geometrically converted to a vertical profile of L1 and L2 bending angles α_1 and α_2 (e.g. Hajj et al., 2002) centred at the tangent point T. For this calculation, precise knowledge of the COSMIC satellite's orbit is required, which is obtained by simultaneously tracking a non-occulting GPS satellite during each occultation, a method known as Precise Orbit Determination (POD). The bending angles α_1 and α_2 are converted to a

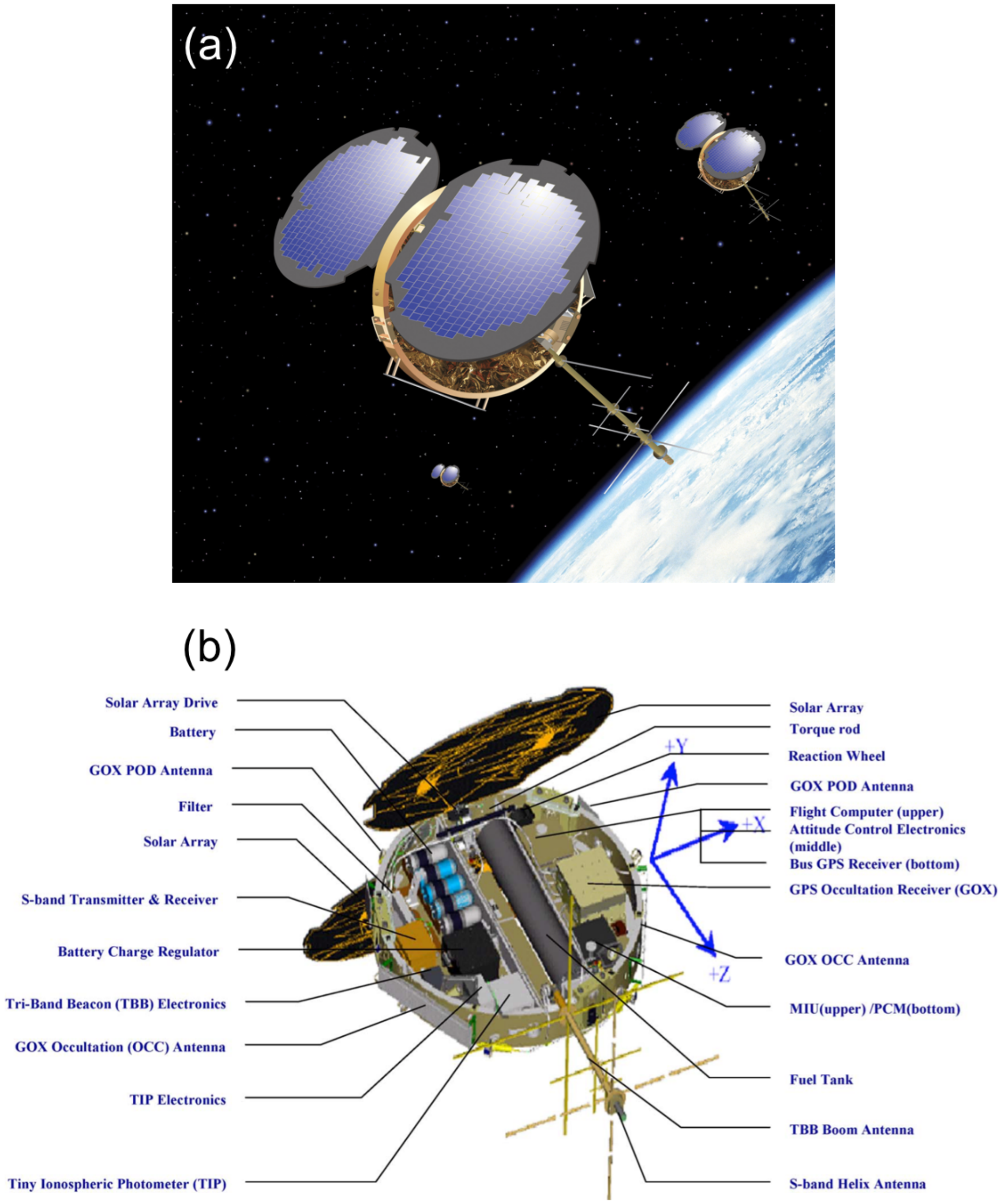

Figure 3.4: Artistic impression of COSMIC satellite (a, image credit: NSPO) and technical diagram of the payloads aboard each COSMIC satellite (b, from Fong et al. (2008)).

neutral bending angle α via ionospheric correction using data from the TIP instrument. An Abel transform (e.g. Kursinski et al., 2000; Hajj et al., 2002) then converts the the bending angle profile to one of refractivity.

Optimal estimation from a numerical weather prediction (NWP) model and an assumption of hydrostatic balance are then used to convert refractivity to profiles of dry temperature, pressure and other parameters. The dry temperature conversion breaks down in the presence of water vapour, but works well in the stratosphere, where water vapour is negligible. Kursinski et al. (1997) estimated a temperature retrieval accuracy of ~0.3 K between 5-30 km, while Tsuda et al. (2011) verified multiple profiles with nearby radiosonde flights, returning discrepancies typically less than 0.5 K between 5-30 km.

Here we use COSMIC level 2 (version 2010.2640) post-processed dry temperature data (e.g. Kursinski et al., 1996) from launch in April 2006 to the end of 2012.

Vertical and horizontal resolutions of GPS-RO

The expected vertical and horizontal resolutions of GPS-RO are discussed at length by Kursinski et al. (1997). They showed that in the stratosphere, where reasonable spherical symmetry of the local atmosphere can be assumed, the vertical resolution ΔZ is primarily limited by Fresnel diffraction as

$$\Delta Z \approx 2\left(\lambda L_T\right)^{\frac{1}{2}} \approx 1.4\,\mathrm{km} \tag{3.1}$$

where $\lambda = 19\,\mathrm{cm}$ is the GPS L1 wavelength and $L_T \approx 28500\,\mathrm{km}$ is the distance from the GPS satellite to the tangent point. The vertical resolution of GPS-RO improves significantly around the tropopause and below due to the exponential increase of the refractivity gradient with decreasing altitude, which increases the size of the bending angle α. As a result, operational retrieval is then switched from the geometric optics method, which is Fresnel-limited as described above, to full-spectrum inversion (FSI) (Tsuda et al., 2011), which can achieve vertical resolutions of a few hundreds of metres. However, increasing water vapour content, smaller gravity wave amplitudes and sharp vertical temperature gradients make gravity wave study using GPS-RO below the tropopause prohibitively difficult, despite the vertical resolution improvements offered by FSI.

Kursinski et al. (1997) showed that the stratospheric horizontal line-of-sight resolution ΔL of GPS-RO could be defined as the horizontal distance travelled by the GPS ray as it enters and exits an atmospheric layer with vertical resolution ΔZ. By a first order

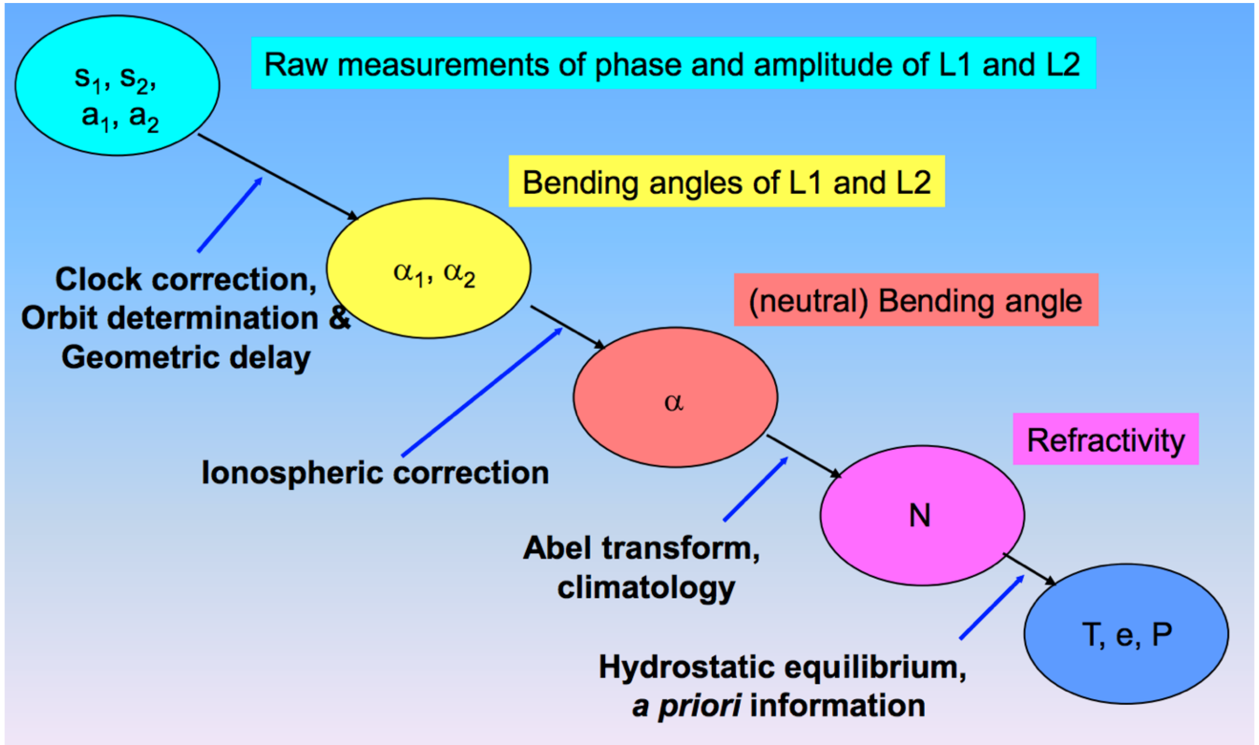

Figure 3.5: Flow chart showing an overview of the processing chain used at CDAAC for the retrieval of temperature profiles from COSMIC GPS-RO. From Derber (2014) and others.

geometric argument, ΔL and ΔZ are approximately related as

$$\Delta L = 2(2R\Delta Z)^{\frac{1}{2}} \tag{3.2}$$

where R is the radius of the atmosphere at the tangent point. For a vertical resolution $\Delta Z \approx 1.4$ km in the stratosphere, the horizontal line-of-sight resolution ΔL is around 270 km. Gravity waves with $\lambda_H \lesssim 270$ km in the line-of-sight are hence less likely to be detected by GPS-RO. However, if the line-of-sight is not aligned with the wave's horizontal wavenumber vector, the projection of λ_H in the line-of-sight may be longer. This means that some waves with $\lambda_H < 270$ km may be resolved (Alexander et al., 2009). The cross-beam horizontal resolution of GPS-RO in the stratosphere is only diffraction limited since horizontal refractivity gradients are generally small, and is also around 1.4 km.

3.4 AIRS

The Advanced Infrared Sounder (AIRS) is an infrared nadir-sounding instrument on NASAs Aqua satellite, launched on the 4 May 2002 and with continuous data availability since shortly after that date (Aumann et al., 2003). Part of NASAs A-Train satellite constellation, Aqua flies in a 98-minute sun-synchronous polar orbit, with an ascending-node equator-crossing local solar time of 1:30 p.m. By scanning ±49° from the nadir, AIRS makes measurements in 2378 spectral channels in a continuous 90-pixel (1800 km) wide "swath" of along the orbital track. The width of the cross-track footprint varies from ~13.5 km at the centre of the instrument track to ~40 km at the edges, while the along-track footprint is consistently ~20 km. Data are processed and archived in two-dimensional 135-pixel along-track sections referred to as "granules", which correspond to 6 minutes of data collection.

Gravity waves can be detected in these granules as perturbations from a background radiance state. AIRS measurements have been used to study atmospheric gravity waves in a range of recent studies (e.g. Alexander and Barnet, 2007; Alexander and Grimsdell, 2013; Hoffmann et al., 2013, 2014; Gong et al., 2015). In chapter 5 we make measurements of gravity wave properties detected in AIRS measurements using a two-dimensional Stockwell transform.

3.5 HIRDLS

The High Resolution Dynamics Limb Sounder (HIRDLS) is a 21-channel limb-viewing radiometer on NASAs Aura satellite (Gille et al., 2003). Aura is also part of NASAs A-Train, with an ascending-node equator-crossing time a few minutes after Aqua. HIRDLS measured around 5500 vertical profiles per day between January 2005 and early 2008, when the failure of the optical chopper prohibited further measurements. Gravity wave study with HIRDLS has proved a particularly productive area of research in the last decade (e.g. Alexander et al., 2008; Hoffmann and Alexander, 2009; Wang and Alexander, 2010; Wright et al., 2010; Yan et al., 2010; Ern and Preusse, 2012; Wright et al., 2013). HIRDLS profiles are typically spaced approximately 70 to 120 km apart in the direction of the Aura orbital track, due to an optical blockage discovered shortly after launch (Gille et al., 2008).

In chapter 4 we use HIRDLS measurements to make estimates of gravity wave momentum flux (GWMF) over the southern Andes, Drake Passage and Antarctic Peninsula region during June-August 2006 using a method originally based upon that of Alexander et al. (2008). We then compare these results to GWMF estimates made by COSMIC GPS-RO using a method presented here.

Chapter 4

The southern stratospheric gravity wave hot spot: individual waves and their momentum fluxes measured by COSMIC GPS-RO

In this chapter we use COSMIC GPS radio occultation data to investigate the nature and sources of stratospheric gravity waves in the southern gravity wave hot spot, namely the mountainous southern Andes/Drake Passage/Antarctic Peninsula. We explore source mechanisms for the long leeward region of enhanced gravity wave energy observed to stretch eastward from the mountains more that half way around the world. We find evidence for meridional wave propagation, develop a quantitative automated identification method for detecting gravity waves in vertical profiles, investigate gravity wave intermittency and finally make estimates of momentum flux using closely spaced pairs of COSMIC profiles.

The key findings of this chapter are published as:

Hindley, N. P., C. J. Wright, N. D. Smith, and N. J. Mitchell, 2015: The southern stratospheric gravity wave hot spot: individual waves and their momentum fluxes measured by cosmic gps-ro. *Atmos. Chem. Phys.*, **15**, 7797–7818, doi:10.5194/acp-15-7797-2015

Abstract

Nearly all general circulation models significantly fail to reproduce the observed behaviour of the Southern wintertime polar vortex. It has been suggested that these biases result from an underestimation of gravity wave drag on the atmosphere at latitudes near 60°S, especially around the "hot spot" of intense gravity wave fluxes above the mountainous Southern Andes and Antarctic peninsula. Here, we use Global Positioning System (GPS) Radio Occultation (RO) data from the COSMIC satellite constellation to determine the properties of gravity waves in the hot spot and beyond. We show considerable southward propagation to latitudes near 60°S of waves apparently generated over the southern Andes. We propose that this propagation may account for much of the wave drag missing from the models. Furthermore, there is a long leeward region of increased gravity wave energy that sweeps eastwards from the mountains over the Southern Ocean. Despite its striking nature, the source of this region has historically proved difficult to determine. Our observations suggest that this region includes both waves generated locally and orographic waves advected downwind from the hot spot. We describe and use a new wavelet-based analysis technique for the quantitative identification of individual waves from COSMIC temperature profiles. This analysis reveals different geographical regimes of wave amplitude and short-timescale variability in the wave field over the Southern Ocean. Finally, we use the increased numbers of closely spaced pairs of profiles from the deployment phase of the COSMIC constellation in 2006 to make estimates of gravity wave horizontal wavelengths. We show that, given sufficient observations, GPS-RO can produce physically reasonable estimates of stratospheric gravity wave momentum flux in the hot spot that are consistent with measurements made by other techniques. We discuss our results in the context of previous satellite and modelling studies and explain how they advance our understanding of the nature and origins of waves in the southern stratosphere.

4.1 Introduction

Gravity waves are propagating mesoscale disturbances that transport energy and momentum in fluid environments. They are a vital component of the atmospheric system and a key driving mechanism in the middle and lower atmosphere through drag and diffusion processes (e.g. Fritts and Alexander, 2003, and citations therein). Despite their importance, considerable uncertainty remains regarding gravity wave sources, fluxes, propagation and variability.

A striking example of the importance of accurately assessing gravity wave fluxes is that nearly all global climate models (GCMs) have a systematic bias in their representation of the southern stratosphere. In particular, in the models the breakdown of the winter polar vortex occurs too late in the year, the polar vortex winds are too strong and the polar vortex temperatures are too low. This "cold pole" bias is a long-standing problem and has been identified as a serious impediment to model progress, leading to discrepancies in properties including simulated Antarctic ozone trends and southern-hemisphere climate (e.g. McLandress et al., 2012). These problems are believed to arise because the models are deficient in gravity wave drag in the stratosphere at latitudes near 60°S. This deficiency may arise because in the real atmosphere waves from other latitudes propagate into this latitude belt, or because the sources of gravity waves in the models under-represent the in situ generation of waves. Determining the nature of gravity waves at latitudes near 60°S is thus a significant problem.

During austral winter, observations have revealed the southern hemisphere stratosphere to be home to some of the most intense gravity wave activity on Earth. At high southern latitudes, the mountains of the southern Andes and Antarctic Peninsula are a hot spot of stratospheric gravity wave momentum flux (e.g. Eckermann and Preusse, 1999; Ern et al., 2004; Alexander and Teitelbaum, 2007; Alexander et al., 2008; Alexander and Teitelbaum, 2011). Several second-order hot spots include South Georgia (Alexander et al., 2009) and other small islands in and around the Southern Ocean (Alexander and Grimsdell, 2013; Hoffmann et al., 2013). Accompanying the momentum flux hot spot is a long leeward distribution of increased gravity wave energy stretching eastwards from the southern Andes, Drake Passage and Antarctic Peninsula far over the Southern Ocean. This feature has puzzled researchers since it was first seen in spaceborne observations. Despite more than a decade of close observation (e.g. Wu and Waters, 1996; Wu and Jiang, 2002; Ern et al., 2004; Hei et al., 2008; Alexander et al., 2008, 2009; Yan et al., 2010; Gong et al., 2012; Hendricks et al., 2014; Preusse et al., 2014) its origins are still not incontestably understood.

It has been suggested that gravity waves in this region may have a number of orographic and non-orographic sources, such as the leeward propagation of mountain waves from the southern tip of South America and/or the northern tip of the Antarctic Peninsula (Preusse et al., 2002; Sato et al., 2009, 2012), baroclinic instabilities from tropospheric storm systems (Hendricks et al., 2014; Preusse et al., 2014) or spontaneous adjustment arising independently from, or as a result of, either or both of these primary processes. It is likely that the gravity waves observed in this region are a result of some or all of these processes overlapping in spatial and temporal regions.

However, quantitatively identifying and describing the location, magnitude and short-timescale variability of each gravity wave source through close observation has proved exceptionally challenging. It is perhaps for this reason that the current generation of General Circulation Models (GCMs) exhibit strong disagreement in the magnitude and distribution of the flux of horizontal psuedomomentum (hereafter referred to as momentum flux) due to gravity waves in the southern hemisphere stratosphere during austral winter compared to observations (Geller et al., 2013). Particularly large discrepancies are found over the mountains of the southern Andes and Antarctic Peninsula suggesting even orographic wave drag is not simulated consistently.

For the majority of operational GCMs used in numerical weather prediction (NWP), many gravity waves are sub-gridscale phenomena and their effects must be parametrized. Parametrizations vary greatly between GCMs, but tuning parameters may for example be chosen in order to produce comparable monthly-mean zonal-mean wind fields to observations (Geller et al., 2013) or obtain a realistic quasi-biennial oscillation (QBO) (e.g. Scaife et al., 2000) while remaining physically plausible. However, a current scarcity of robust observations of key gravity wave parameters means that these parametrizations are poorly constrained (Alexander et al., 2010). With the advent of increased computing power in recent years, high spatial resolution GCMs without the need for gravity wave parametrizations are becoming available (e.g. Watanabe et al., 2008). Such high-resolution modelling studies are promising (e.g. Sato et al., 2012), but discrepancies between observed and modelled parameters still remain. An in-depth review of the current state of gravity wave modelling is presented by Preusse et al. (2014).

All of the above factors highlight the need for accurate measurements of gravity wave sources, energies, fluxes and variability. Here, we use Global Positioning System radio occultation (GPS-RO) data to investigate the nature and origins of waves in the southern stratospheric gravity wave hot spot and associated leeward distribution of enhanced gravity wave energy. In Section 4.2, we present maps and cross-sections of gravity wave energy in the southern hemisphere, with implications for oblique focussing and leeward propagation of gravity waves into the southern stratospheric jet. In Section 4.3, we propose a new method for the quantitative identification of individual waves from GPS-RO profiles. We use this method to investigate the geographical distribution of wave amplitudes and short-timescale variability of individual gravity waves in the wave field over the Southern Ocean. In Section 4.4, we present a method for the estimation of gravity wave momentum flux from GPS-RO measurements over the southern Andes and Antarctic Peninsula using pairs of closely spaced and closely timed profiles. Our results are discussed in the context of other studies in Section 4.5, and in Section 4.7 the key results of the present study are summarised.

4.1.1 COSMIC GPS Radio Occultation

Launched in April 2006, The Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) mission consists of six low Earth orbit (∼800 km) satellites at ∼72° inclination and 30° separation. A detailed description of the COSMIC constellation and the radio occultation process is provided by Liou et al. (2007). Each satellite tracks occulting GPS satellites as they rise above or set below the Earth's horizon. As the GPS signal traverses the atmospheric limb, phase delay measurements attributable to changing vertical gradients of refractivity in the atmosphere are measured. Taking an integral along the line-of-sight, vertical profiles of dry temperature and pressure can be computed at the tangent point of the occultation via an Abel inversion (Fjeldbo et al., 1971). The dry temperature conversion breaks down in the presence of water vapour, but works well in the stratosphere, where water vapour is negligible. Kursinski et al. (1997) estimated a temperature retrieval accuracy of ∼0.3 K between 5-30 km, while Tsuda et al. (2011) verified multiple profiles with nearby radiosonde flights, returning discrepancies typically less than 0.5 K between 5-30 km.

In the present study we use COSMIC level 2 (version 2010.2640) post-processed dry temperature data from launch in April 2006 to the end of 2012. The sampling density of the COSMIC constellation in its final deployment configuration for a typical month in the southern hemisphere is shown in Fig. 4.1. Good coverage at high latitudes and a band of preferential sampling at around 50°S as a result of orbital geometry means that COSMIC GPS-RO is well suited to a study of the southern gravity wave hot spot and the surrounding area.

Vertical and horizontal resolution limits

Currently, no single observational technique can study the entire gravity wave spectrum. Each technique is sensitive to a specific portion of the gravity wave spectrum, referred to as its observational filter (Alexander and Barnet, 2007; Preusse et al., 2008; Alexander et al., 2010).

The expected vertical and horizontal resolutions of GPS-RO are discussed at length by Kursinski et al. (1997). They showed that in the stratosphere, where reasonable spherical symmetry of the local atmosphere can be assumed, the vertical resolution ΔZ is primarily limited by Fresnel diffraction as

$$\Delta Z \approx 2 \left(\lambda L_T\right)^{\frac{1}{2}} \approx 1.4\,\mathrm{km} \tag{4.1}$$

where $\lambda = 19\,\mathrm{cm}$ is the GPS L1 wavelength and $L_T \approx 28500\,\mathrm{km}$ is the distance from the GPS satellite to the tangent point. The vertical resolution of GPS-RO improves significantly below the tropopause due to the exponential increase of refractivity gradient with decreasing altitude, but the combination of sharp vertical temperature gradient changes, increased humidity and smaller wave amplitudes make gravity wave study in this region difficult with GPS-RO via traditional methods.

Kursinski et al. (1997) showed that the horizontal line-of-sight resolution ΔL of GPS-RO could be defined as the horizontal distance travelled by the GPS ray as it enters and exits an atmospheric layer with vertical resolution ΔZ. By a first order geometric argument, ΔL and ΔZ are approximately related as

$$\Delta L = 2(2R\Delta Z)^{\frac{1}{2}} \tag{4.2}$$

where R is the radius of the atmosphere at the tangent point. For a vertical resolution $\Delta Z \approx 1.4\,\mathrm{km}$ in the stratosphere, the horizontal line-of-sight resolution ΔL is around 270 km. Gravity waves with $\lambda_H \lesssim 270\,\mathrm{km}$ in the line-of-sight are hence less likely to be detected by GPS-RO. However, if the line-of-sight is not aligned with the wave's horizontal wavenumber vector, the projection of λ_H in the line-of-sight may be longer. This means that some waves with $\lambda_H < 270\,\mathrm{km}$ may be resolved. As discussed by Alexander et al. (2009), orographic waves generated by the mountains of the southern Andes and Antarctic Peninsula may tend to have roughly westward orientated horizontal wavenumber vectors, and the majority of COSMIC occultations in this region tend to be preferentially aligned towards the north-south axis. As a result, the projection of λ_H in the COSMIC line-of-sight is longer and the likelihood of orographic wave detection over this region is increased.

The cross-beam horizontal resolution in the stratosphere is around 1.4 km, being only diffraction limited since horizontal refractivity gradients are generally small. This is of importance to our momentum flux study in section 4.4.

4.2 The gravity wave hot spot and leeward region of increased E_p

In this section, we investigate the seasonal variability and distribution of potential energy per unit mass E_p in the southern hemisphere using COSMIC GPS-RO. E_p is a fundamental property of the gravity wave field and can provide a useful proxy for gravity wave activity.

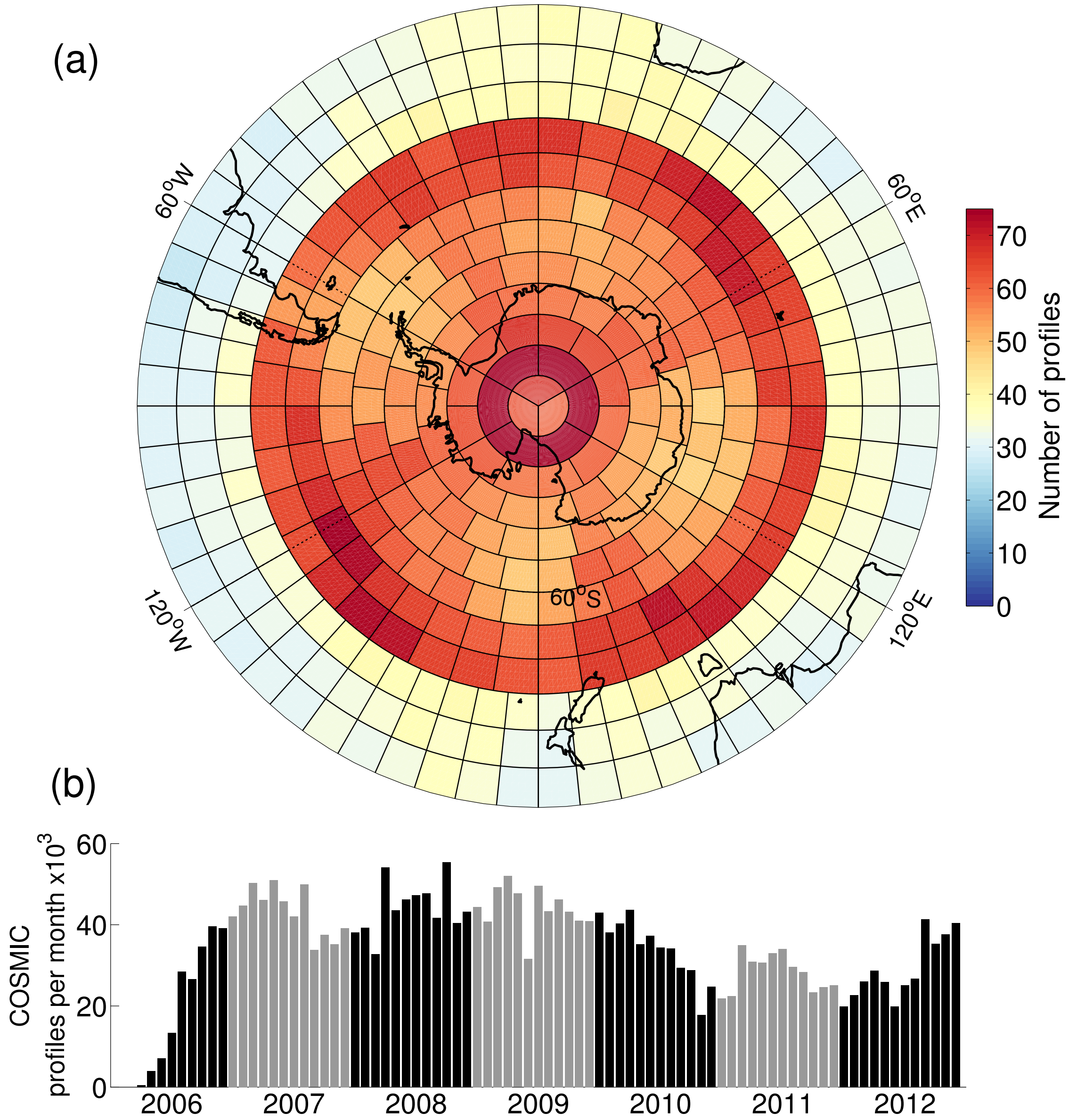

Figure 4.1: Polar stereographic projection of monthly-mean COSMIC sampling density for the period 2007-2012 (a), and total number of occultations per month for the period 2006-2012 (b). Each box in (a) represents an equal area of approximately 550×550 km^2. Alternating years in (b) are shown by black and gray bars.

In satellite observations, E_p is often derived from temperature perturbations around a background mean and can hence be calculated independently in each temperature profile. To calculate E_p, we first interpolate each dry temperature profile $T(z)$ to 100m resolution over the altitude range 0-50 km. We obtain a background temperature profile $\bar{T}(z)$ by low-pass filtering $T(z)$ with a 2nd order Savitzky-Golay filter (Savitzky and Golay, 1964) with an 18 km frame-size and compute $T(z) - \bar{T}(z)$ to yield a temperature perturbation profile $T'(z)$.

Features with vertical scales less than ~2 km cannot be reliably disassociated with noise in GPS-RO temperature profiles (Marquardt and Healy, 2005), therefore we apply a 2$^{\text{nd}}$ order Savitzky-Golay low-pass filter with a 3 km frame size to suppress these small-scale features in $T'(z)$. Note that this step has virtually no effect on vertical wavelengths greater than ~4 km. The transmission functions for each step in our analysis is shown in Fig. 4.6, and discussed further in Section 4.3.1. For the calculation of E_p in this section, only the blue and green dashed lines in Fig. 4.6 apply.

This analysis provides a dynamic cut-off for vertical features in $T'(z)$. Features with vertical scales ~3-14 km are generally transmitted with a factor of at least 0.5, however transmission of vertical wavelengths longer than $\sim$ 13 km (shorter than 4 km) decreases with increasing (decreasing) wavelength. It is important to note that no digital filter can provide a perfect cut-off in the frequency domain without introducing ringing artifacts into the spatial domain via the Gibbs phenomenon. We select the Savitzky-Golay filter as a reasonable trade-off between Gibbs ringing in the spatial domain and a sharp transition into the frequency stop band.

We use $T'(z)$ and $\bar{T}(z)$ to compute $E_p(z)$ as

$$E_p(z) = \frac{1}{2}\left(\frac{g}{N}\right)^2 \left(\frac{T'(z)}{\bar{T}(z)}\right)^2 \tag{4.3}$$

where g is acceleration due to gravity and N is the local Brunt-Väisälä frequency. It is not meaningful to take E_p at a single height z from a single profile since a full wave cycle does not exist (Alexander et al., 2008). Hence, E_p is often taken as an integral over a specified height interval when used as a proxy for gravity wave activity (e.g. Hei et al., 2008).

Unlike previous studies such as Alexander et al. (2009), no planetary wave removal techniques are applied to these data. At high latitudes, planetary waves typically have vertical scales much longer than 10 km, hence they are generally removed by our filtering method. We recognise however that some low-level planetary wave features may remain in the post-processed data.

4.2.1 Geographic distribution of E_p in the southern hemisphere

Figure 4.2 shows E_p in the southern hemisphere for each month in 2010 over the height interval 26-36 km. This 10 km region of the lower-middle stratosphere is selected as being one in which we are most confident of not only the accuracy of COSMIC temperature measurements (Tsuda et al., 2011), but also confident that perturbations detected in these temperature measurements are most likely to correspond to atmospheric gravity waves. Extending the region further upwards into the stratosphere would reduce confidence of this due to increased noise in measurements above z $\approx$ 38 km. We also implicitly assume that λ_Z does not change much with altitude, which forces us to select a relatively small vertical region, though we acknowledge that this may decrease the contribution of waves with $\lambda_Z > 10$ km.

We observe increased levels of E_p in austral winter and lower values in austral summer, consistent with other GPS-RO studies (e.g. Hei et al., 2008; Alexander et al., 2009). Between June and November, we see in Fig. 4.2 a long leeward region of increased E_p stretching clockwise from the Southern Andes, Drake Passage and Antarctic Peninsula at around 70°W to around 180°E. This long leeward region of increased E_p is consistent with studies using other limb sounders such as the Upper Atmosphere Research Satellite Microwave Limb Sounder (UARS-MLS) (e.g. Wu and Waters, 1996), Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA) (e.g. Ern et al., 2004) and the HIgh Resolution Dynamics Limb Sounder (HIRDLS) (e.g. Yan et al., 2010).

The magnitude and distribution of E_p in Fig. 4.2 is also consistent with results from a high-resolution modelling study by Sato et al. (2012) using the T213L256 ”Kanto” GCM developed by Watanabe et al. (2008). This is significant since Sato et al. used no gravity wave parametrizations, such that all resolved waves effects were spontaneously generated. They showed a long leeward distribution of E_p at 10 hPa ($\sim$31 km) stretching clockwise around the southern ocean from the southern Andes and Antarctic Peninsula to around 180°W during June-October. They proposed a downwind propagation mechanism for orographic waves from the mountains of the southern Andes and Antarctic Peninsula, whereby a wave could be freely advected by the component of the mean wind perpendicular to the wave's horizontal wavenumber vector, and primarily attributed the long leeward distribution to this mechanism.

Some differences in our observed distribution of gravity wave E_p are apparent, however.

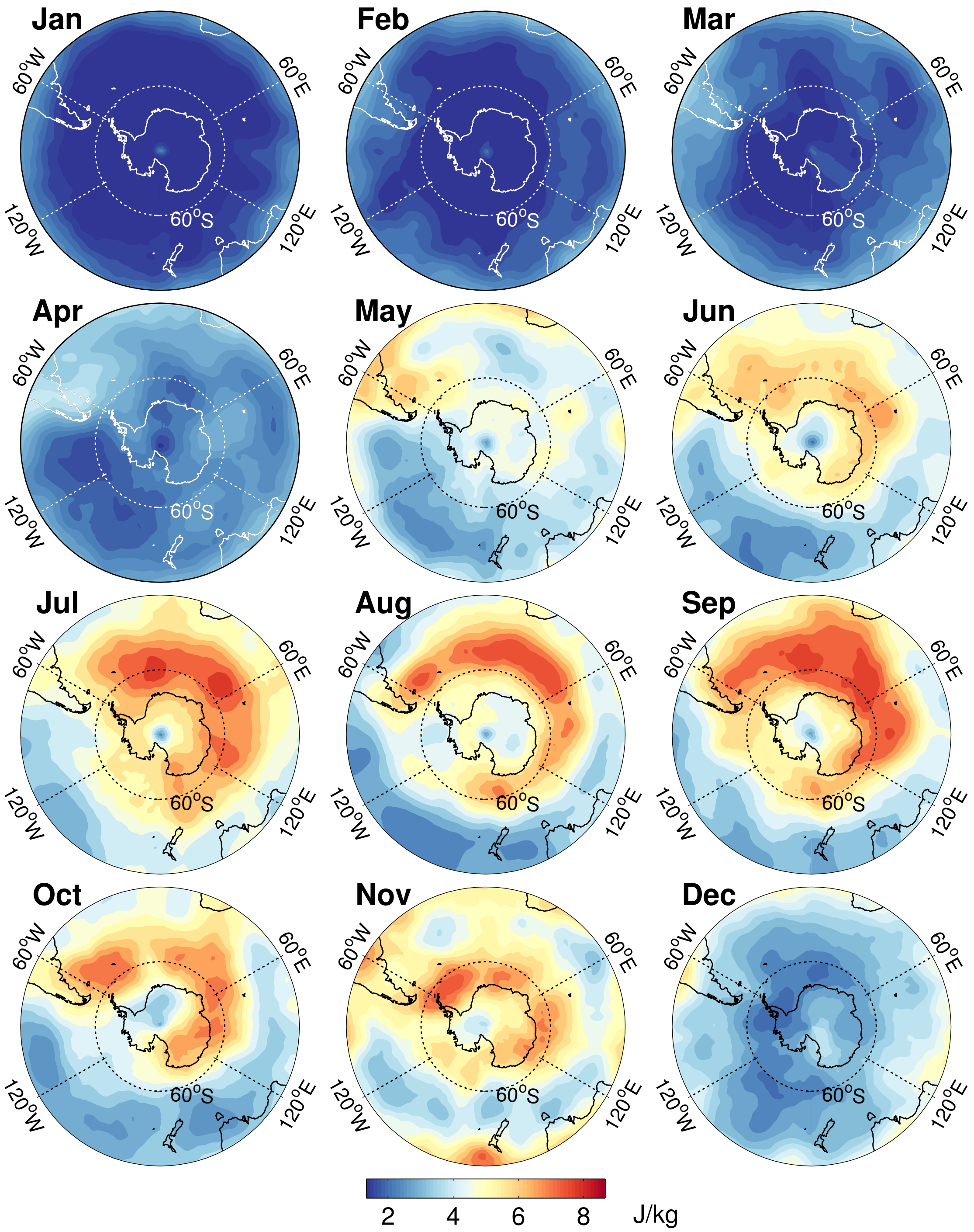

Figure 4.2: Polar stereo projections of monthly-mean potential energy per unit mass E_p in the southern hemisphere averaged over the height range 26-36 km (~20-5 hPa) for each month in 2010.

Sato et al. (2012, their Figure 2) showed maximum E_p directly over the mountains of the southern Andes at 10 hPa. Likewise, Yan et al. (2010, their Figure 5) revealed similar distributions of mean gravity wave amplitude T' from HIRDLS data that also maximised over the mountains and slowly decreased eastward. In our results, we see some enhancement over the mountains in the height range 26-36 km ($\sim$22-5 hPa) in Fig. 4.2, but maximum values are usually observed well to the east over the oceans during 2010. Other years show similar distributions (omitted for brevity).

One possible explanation may relate to the range of vertical wavelengths to which our analysis method is sensitive. For mountain waves, vertical wavelengths directly over the mountains can be quite long (e.g. Alexander and Teitelbaum, 2011). As previously discussed, our analysis method is primarily sensitive to waves with $4 \gtrsim \lambda_Z \gtrsim 13$ km, and significant amplitude underestimation occurs for waves with $\lambda_Z \gtrsim 13$ km. It could be that the contribution of these long λ_Z waves directly over the mountains to the E_p distributions in Fig. 4.2 is underestimated by our analysis method.

Sato et al. (2012) also observed regions of downward energy flux. In particular they found that, in the region immediately eastward of the southern tip of South America, up to 10% of the E_p distribution consisted of downward propagating waves. This suggests that some of the E_p in our observed distribution may correspond to waves that are propagating downward.

The sources of waves in the long leeward region of increased E_p are currently a topic for debate. As mentioned above, Sato et al. (2012) suggested that increased E_p over 70°W-180°E could be primarily due to mountain waves from the southern Andes and Antarctic Peninsula that have been advected downwind, but the rest of the enhancement was likely the result of other mechanisms. Other studies suggest that much of the enhancement is primarily the result of non-orographic wave sources in and around the Southern Ocean (e.g. Hendricks et al., 2014; Preusse et al., 2014). Furthermore, the distribution of increased E_p in Fig. 4.2 is very reminiscent of southern hemisphere storm tracks (Hoskins and Hodges, 2005). It is thus likely that the observed distribution of E_p is the result of a number of orographic and non-orographic processes, each playing different roles in different geographical regions. In the next section we use an extended altitude range to build vertical cross-sections of stratospheric E_p in the long leeward distribution to investigate this further.

4.2.2 Vertical distribution of E_p over the southern Andes and Antarctic Peninsula

An interesting result discussed by Sato et al. (2009) and presented in Sato et al. (2012) was the apparent focusing of gravity waves into the southern stratospheric jet in the Kanto GCM. In a meridional cross-section from 30°S-70°S centred on 55°W (their Fig. 13), Sato et al. showed increased E_p values in a distinct slanted vertical column over the southern Andes during 5 days in August. Energy flux vectors showed a large flow of energy $\sim$1500-2000 km southward over the height region 100 hPa ($\sim$16 km) to 1 hPa ($\sim$48 km). The flow appeared to focus towards the centre of the jet, where mean zonal winds were strongest.

In our Fig. 4.3, we select a thin meridional cross-section of normalised monthly-mean COSMIC E_p for August 2010 centred on 65°W. This is close to the cross-section used by Sato et al. (2012). Since raw E_p is expected to increase with increasing altitude and decreasing pressure, each height level in Fig. 4.3 has been normalised such that the lowest value is equal to 0 and the highest value is equal to 1 (Wright and Gille, 2011). This approach highlights the vertical structure. Although temperature profiles from COSMIC typically exhibit increased noise above around 40 km, the normalisation and the increased number of measurements in the month-long time window potentially allow us to resolve large persistent features at higher altitudes, albeit with caution.

A slanted vertical column of increased E_p in the height region 22-35 km and a near vertical column from 35-50 km is evident in Fig. 4.3. The lower section of the column traverses nearly 1500 km southward over the height region 22-35 km. This suggests a clear focusing effect similar to the one suggested by Sato et al., although we cannot recover energy flux information from COSMIC. A monthly-mean zonal wind field $\bar{U}$ (thick contours) from ECMWF operational analysis is used in Fig. 4.3 to show the approximate position of the southern stratospheric jet over the southern Andes and Antarctic Peninsula during August 2010. The gradient of the southward slant in E_p is greatest when the horizontal gradient in zonal wind speed is greatest, such that waves appear to be focused into the centre of the stratospheric jet. This observation is consistent with meridional ray-tracing analyses in the Kanto model (Watanabe et al., 2008; Sato et al., 2009, 2012). Above $\sim$35 km the horizontal gradient in zonal mean wind speed is low and waves appear to generally propagate upward without further latitudinal drift.

This result suggests that waves observed at around 30-40 km over the southern tip of South America and the Drake Passage may have sources further north. In a ray-tracing analysis for an idealized background zonal wind field, Sato et al. (2012, their Figure 5) showed that waves with horizontal ground-based phase velocity $c_H = 0$ and $\lambda_H = 300$ km launched from the southern Andes could propagate eastward and southward by up to

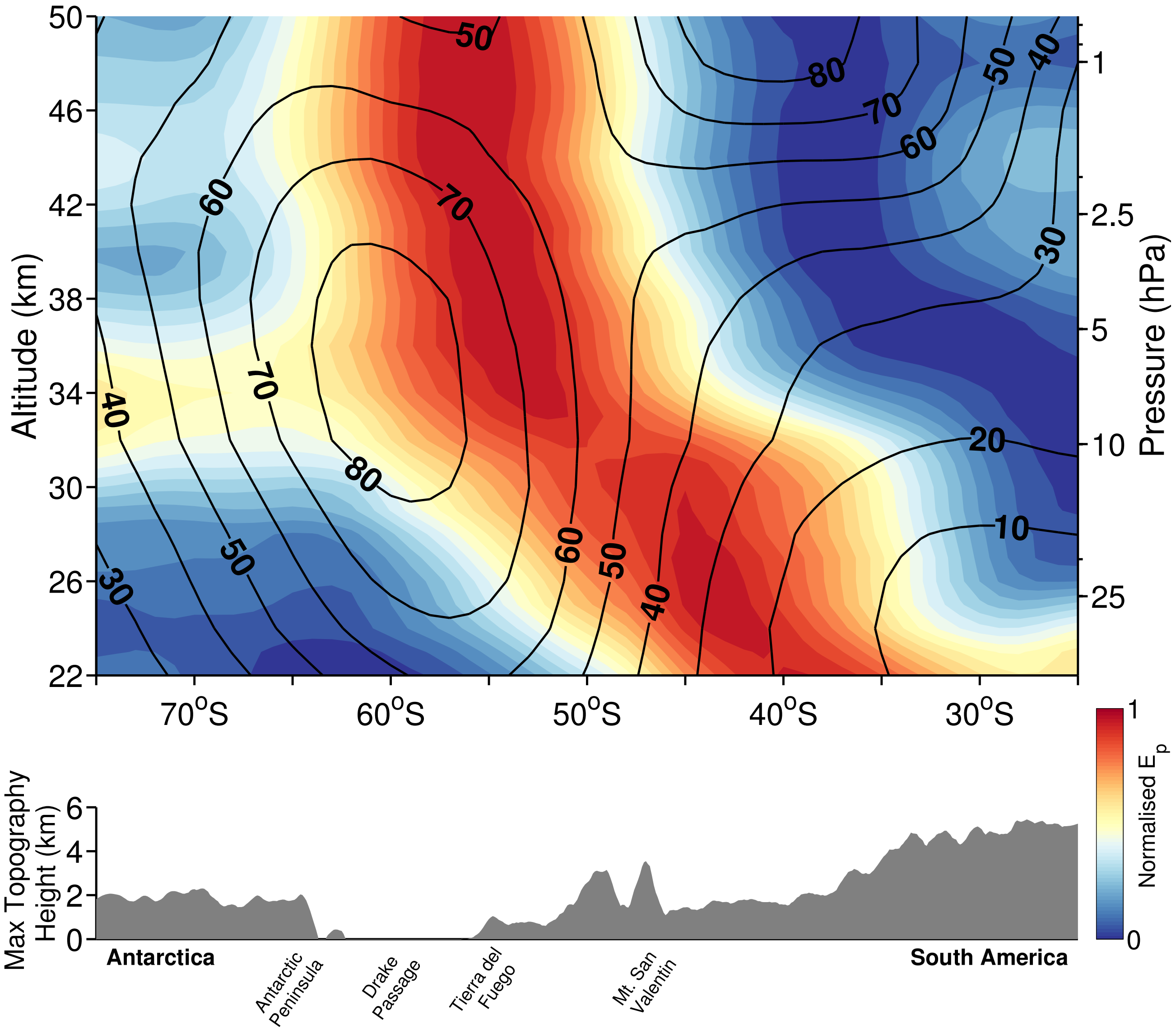

Figure 4.3: Normalised monthly-mean meridional cross-section of E_p in August 2010 over the southern Andes and Antarctic Peninsula (top panel) and maximum topography height (bottom panel) in a ±5° slice centred on 65°W. Monthly-mean zonal-mean winds from ECMWF operational analyses are shown by thick contours in the top panel, at intervals of 10 m/s. Note the E_p has been normalised at each height level to highlight the vertical structure.

around 2500 km and 1000 km respectively before reaching an altitude of 40 km. They found that waves launched from north of 45°S did not propagate upward due the mean wind being too weak. Our results suggest that such waves may indeed propagate from sources north of 45°S, since the slanted column in Fig. 4.3 is observed all the way down to 22 km over 30-45°S . This could imply that there are significant time periods where the tropospheric zonal winds are strong enough to allow vertical propagation of mountain waves from these sources.

An important consideration of this work is the effect of the range of gravity wave vertical wavelengths to which our observations and analysis are limited. For mountain waves

$$\lambda_Z \approx \frac{2\pi U_{||}}{N} \tag{4.4}$$

where $U_{||}$ is the component of mean wind speed $\bar{U}$ parallel to the wave's horizontal wavenumber vector and N is the local Brunt-Väisälä frequency (Eckermann and Preusse, 1999, Eq.1). Our analysis is primarily sensitive to waves with $4 \gtrsim \lambda_Z \gtrsim 13\,\mathrm{km}$ (Fig. 4.6). From Equation 4.4, mountain waves could have vertical wavelengths too long to be detected for $U_{||} \gtrsim 40 - 50\,\mathrm{ms}^{-1}$. However, it has been shown (e.g. Alexander and Teitelbaum, 2011) that horizontal wavenumber vectors of mountain waves over the southern tip of South America can be aligned more southwards poleward of 45°S over the Drake Passage. Here, the mean wind vector $\bar{U}$ and the horizontal wavenumber vector are no longer parallel and shorter vertical wavelengths are not precluded since $U_{||}$ is reduced. Indeed, similar observational studies (Yan et al., 2010; Ern et al., 2011, e.g.) estimate stratospheric vertical wavelengths of 10-13 km at these latitudes on zonal average. These values just fall within the detectable range of our analysis. Therefore the slanted vertical column of E_p in Fig. 4.3 could be due to mountain waves and could suggest meridional propagation. Using similar approaches to those presented here, other observational studies have also attributed slanted vertical columns of gravity wave variances (Jiang et al., 2004) and momentum fluxes (Ern et al., 2013) to the meridional propagation of gravity waves in the tropics and mid-latitudes, but for waves generated by convection.

Sato et al. also suggested that a symmetric northward focusing effect may occur for orographic waves from the Antarctic Peninsula. We investigated such an effect using COSMIC data. Though a slight suggestion of northward meridional focussing may be evident (not shown), we could not find an effect so clear as is observed over the southern Andes.

4.2.3 Vertical distribution of E_p over the Southern Ocean

We also investigate the vertical distribution of wave energies over the Southern Ocean. Fig. 4.4 shows altitude-normalised E_p in a zonal cross-section from $40 - 60°$S during August 2010. As in Fig. 4.3, E_p is normalised at each height level in order to highlight the vertical structure.

The vertical column of increased E_p located around 70°W in Fig. 4.4 is the projection in the zonal domain of the vertical column evident in Fig. 4.3. This column is highly suggestive of intense localised mountain wave activity from the southern Andes. The relative intensity of this column at lower altitudes suggests that, within the observational filter of our COSMIC analysis, the southern Andes is the dominant source of orographic wave activity in this latitude band. If small mountainous islands in the Southern Ocean are also significant orographic sources, as has been suggested in recent studies (Alexander et al., 2009; McLandress et al., 2012; Alexander and Grimsdell, 2013), then it is likely that waves from these islands either (1) fall outside the observation filter of our analysis; (2) have small amplitudes; or (3) are too intermittent over monthly time-scales to be revealed in our analysis.

The column at 70°W appears to persist over the full height range in Fig. 4.4. However, between 25-35 km the largest values are observed well eastward, between 60°E-60°W. These peaks are located in a deep region of increased E_p between 20-40 km and 30°W-90°E, which is the projection in the vertical domain of the long leeward region of increased E_p seen in Fig. 4.2.

At first glance, Fig. 4.4 suggests that this long leeward region of increased E_p is strongly associated with mountain waves from the southern Andes and Antarctic Peninsula region. The lack of significant gravity wave energies upwind (westward) of the mountains and the intensity of energies downwind (eastward) is clear. Sato et al. (2012) suggested that the leeward distribution of increased E_p might be the result of primary mountain waves from the southern Andes and Antarctic Peninsula that have been advected downwind. However, Sato et al. also showed that waves with $\lambda_H < 350$ km rarely travelled further east than the prime meridian via this mechanism, even under ideal conditions. This suggests that if downwind-advected orographic waves do contribute to the region of increased E_p eastwards of the prime meridian, they likely have $\lambda_H \gtrsim$350 km. Primary orographic waves from the southern Andes and Antarctic Peninsula may also contribute through secondary mechanisms, such as the local generation of waves around the stratospheric jet through breaking or other wave-mean flow interactions (Bacmeister and Schoeberl, 1989). Waves generated by in-situ instabilities and spontaneous adjustment around the stratospheric may also play a part.

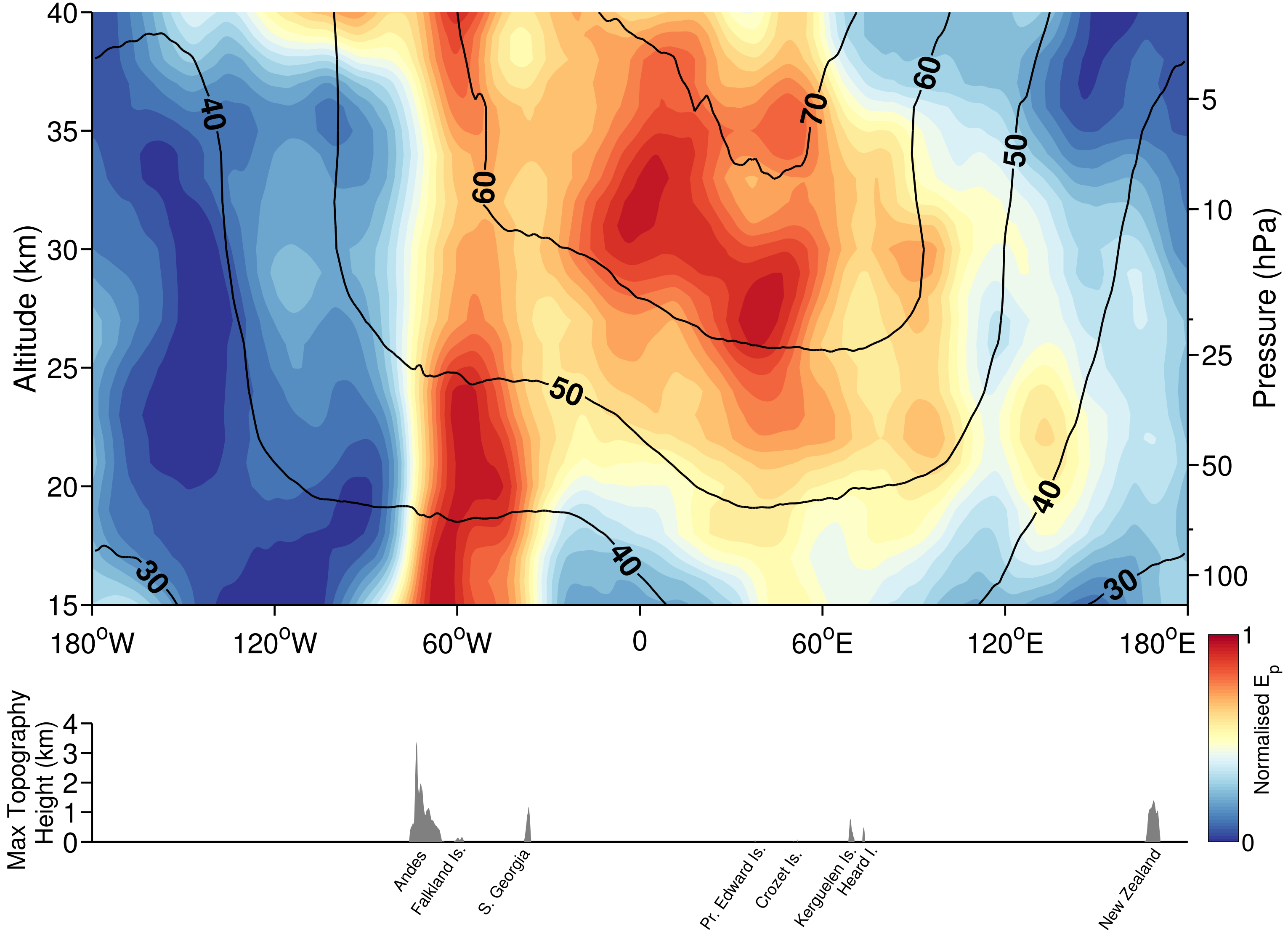

Figure 4.4: Normalised monthly mean zonal cross-section of E_p for August 2010 over Southern Ocean (top panel) and maximum topography height (bottom panel) in a ±10° slice centred on 50°S. Monthly mean zonal mean winds from ECMWF operational analyses are shown by thick contours in the top panel, at intervals of 10 m/s. Note the E_p has been normalised at each height level to highlight the vertical structure, and that the height range differs from Fig. 4.3 due to increased measurement noise near to the tropopause and stratopause over this larger geographic region.

Hoskins and Hodges (2005) presented a detailed view of southern hemisphere storm tracks in ECMWF Re-Analysis (ERA-40) data. During austral winter, storms generally tended to maximise over the southern Atlantic and Indian sectors, spiralling poleward and eastward over the Pacific sector. Such storm-tracks may indicate intense sources of non-orographic wave activity. O'Sullivan and Dunkerton (1995) showed that non-orographic waves generated around the tropospheric jet can have vertical wavelengths of a few kilometres. These wavelengths may be too short to be resolved by COSMIC. Mean zonal wind generally increases with height as seen in Fig. 4.4. This may refract these waves to longer vertical wavelengths such that they become visible to COSMIC. This could explain the relative reduction of wave activity at low altitudes over 30°W to 120°E in Fig. 4.4. The spiralling effect of the storm-tracks might also mean that these intense sources of non-orographic waves may begin to move poleward out of the latitude band used in Fig. 4.4. This may explain the relative decrease in intensity further eastward.

We suspect therefore that the leeward region of increased E_p over 70°W-90°E in Fig. 4.4 is likely dominated by (1) primary orographic waves with $\lambda_H > 350\,\text{km}$ from the southern Andes or Antarctic Peninsula that have been advected downwind; (2) secondary waves with non-zero phase speeds generated in the breaking zones of primary orographic waves; and (3) non-orographic wave activity associated with storm tracks over the oceanic sectors. Note that due to vertical resolution limitations, these results may underestimate the contribution of (3). They also do not preclude the existence of other non-orographic sources in the region such as jet instabilities or spontaneous adjustment mechanisms.

4.3 Individual waves

The long leeward region of increased E_p observed over the southern Atlantic and Indian Oceans is a persistent feature each year during austral winter, though some interannual variability exists. Multiple year averages are one way to learn about dominant processes in a region, but in order to investigate properties of a specific wave field, such as vertical wavelength or wave amplitude, a key question must first be answered: is a wave present? Once this has been answered, it becomes possible to investigate the distribution and species of individual gravity waves in a geographical region.

4.3.1 Wave identification (Wave-ID) methodology

This section describes our methodology for identifying individual gravity waves from COSMIC GPS-RO temperature profiles. The method is illustrated for an example profile in Fig. 4.5.

We begin by extracting temperature perturbations $T'(z)$ from each profile (Fig. 4.5(a)) as described in Section 4.2. We then window the profile with a Gaussian of Full Width at Half Maximum (FWHM) 22 km centred at a height of 30 km (Fig. 4.5(b)). The purpose of this step is to focus on the height range of the profile most appropriate for gravity wave study using COSMIC GPS-RO data. This height range is chosen to generally correspond to the largest vertical region where (1) the error in bending angle is low; (2) we are unlikely to encounter spurious temperature perturbation anomalies due to incomplete background removal around the tropopause; and (3) retrieval errors associated with ionospheric effects are low (see Tsuda et al., 2011). This corresponds to a region typically between 20-40 km. The choice of a Gaussian window minimises edge effects that may arise in subsequent spectral analysis.

We normalise the windowed profile such that the root-sum-square (RSS) "energy" of the profile is equal to 1 (Fig. 4.5(c)). Note that the term "energy" is defined as the sum-square of the values of the windowed profile and does not take any other physical meaning here. We then set the average of the profile to zero, and compute the Continuous Wavelet Transform (CWT) of the windowed, normalised and zero-averaged profile. For scale parameter a and position along the profile b, the spectral coefficients $C(a, b)$ of the CWT are given as

$$C(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} T'(z) \psi^* \left(\frac{z-b}{a} \right) dz \tag{4.5}$$

where $T'(z)$ is our normalised, windowed and zero-averaged perturbation profile and ψ^* is the complex conjugate of the analysing wavelet. We choose an 8^{th} order complex Gaussian wavelet for ψ, such that phase information is retained. For such a Gaussian wavelet, scale parameter a, which corresponds to a "stretching" of the wavelet, is approximately related to wavelength λ by

$$\lambda \approx \frac{a \Delta z}{f_{cent}} \tag{4.6}$$

where Δz is sampling interval and f_{cent} is the dominant central frequency of the wavelet

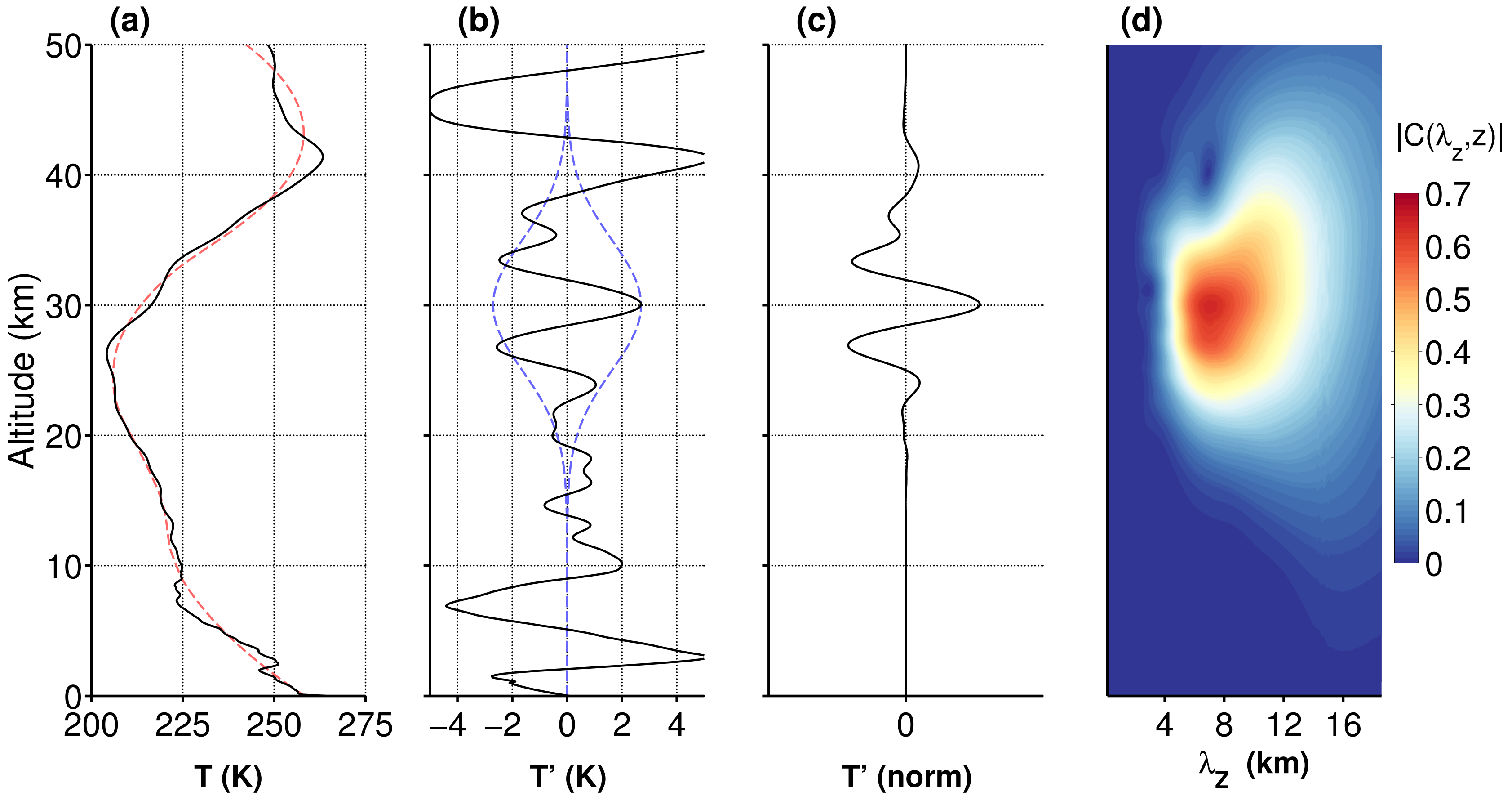

Figure 4.5: Wave identification (Wave-ID) methodology for an example COSMIC profile at 2319 UTC on 1st August 2010 at 53°S, 50°W. Panels show (a) raw temperature profile T (black solid) and filtered background temperature profile $\bar{T}$ (red dashed), (b) Temperature perturbation profile T' (black solid) and a Gaussian window centred on 30 km (blue dashed), (c) windowed, RSS normalised and zero-averaged perturbation profile T'_{norm}, (d) magnitudes of the spectral coefficients of the Continuous Wavelet Transform $|C(\lambda_Z, z)|$ of T'_{norm}. For details, see text.

for scale parameter $a = 1$ and unit integer interval spacing ($\Delta z = 1$). Position along the profile b corresponds to altitude z at intervals of Δz.

We are thus able to describe the spectral coefficients $C(a, b)$ in terms of vertical wavelength λ_Z and altitude z. The absolute magnitudes of the spectral coefficients $|C(\lambda_Z, z)|$ are plotted in Fig. 4.5(d).

As a result of the normalisation and zero-averaging, the absolute magnitudes of these coefficients can be interpreted as coefficients of pseudo-correlation, describing the correlation between the profile and a wavelet of wavelength λ_Z at altitude z. Ranging between 0 and 1, high (low) values of $|C(\lambda_Z, z)|$ imply the presence (absence) of a clear wave-like feature in the profile.

If the profile and the analysing wavelet are both real, both have RSS "energy" equal to 1 and are both zero-averaged, then the coefficients of the CWT can be exactly interpreted as coefficients of correlation in the usual sense. The coefficients in our analysis are strictly pseudo-coefficients of correlation, due to our choice of a complex wavelet. This choice has the advantage of producing one single peak per wave-like feature in Fig. 4.5(d), which is easier to interpret than a series of peaks corresponding to correlation/anti-correlation which would result from a non-complex wavelet.

To positively identify a wave in the profile, we require that the absolute magnitude of the spectral peak coefficient $C_{\text{max}} = \max(|C(\lambda_Z, z)|)$ is greater than or equal to 0.6. This choice is somewhat arbitrary, but it can be interpreted as a requirement that the profile is pseudo-correlated with a wavelet of wavelength λ_Z at position z with coefficient greater than 0.6. If this condition is satisfied, the identification is positive and we record the vertical wavelength λ_{peak} and altitude z_{peak} at C_{max}. As a result of the Gaussian windowing, z_{peak} is almost always located within one wavelength λ_{peak} of 30 km altitude, hence it is reasonable to consider this analysis method as sensitive to gravity waves at a height of around 30 km.

C_{max} can thus be regarded as a confidence metric for the existence of wave-like features in COSMIC perturbation profiles. In the example in Fig. 4.5(d), $C_{\text{max}} \approx 0.64$ such that a wave with $\lambda_Z \approx 7.1$ km is positively identified at an altitude near 30 km. Information regarding the wave's amplitude T' cannot strictly be obtained from the CWT, so in order to obtain an estimate of T' we find the maximum amplitude of the temperature perturbation profile $T'(z)$ over the height region $z_{\text{peak}} \pm \lambda_{\text{peak}}/2$. In the example in Fig. 4.5(b), $T' \approx 2.3$ K.

To summarise our requirements for a positive wave identification, we require that the wave (1) has an amplitude $1 < T' < 10$ K; (2) has a vertical wavelength $2 < \lambda_{\text{peak}} < 20$ km; (3)

is located such that $20 < z_{\text{peak}} < 40\,\text{km}$; and (4) has a confidence metric $C_{\text{max}} \geqslant 0.6$ as described above. Note that the horizontal wavelengths of waves we detect here are also limited by the observational filter of the GPS-RO technique to $\lambda_H \lesssim 270\,\text{km}$ as discussed in Chapter 3.

Using these criteria, we find that on global year-long average around 20-40% of profiles contain an identifiable gravity wave signal. In some regions and seasons, as will be seen later, this fraction can be as high as $\sim 80\%$. This wave identification method will be henceforth described as the Wave-ID method for convenience. Note that we currently limit this Wave-ID methodology to one (the dominant) wave per profile.

We note that this method preferentially selects profiles that contain a single large amplitude monochromatic wave with low levels of disassociated noise. A superposition of two waves of equal amplitude may result in neither being identified due to the confidence metric described above. This may also affect our amplitude estimation. However on average it is equally likely that the amplitude will increase or decrease as a result of any superposition. Hence if a sufficient number of profiles are measured, this effect should average out. Wright and Gille (2013) showed that in the southern hemisphere during austral winter, and particularly in the vicinity of the southern Andes and Antarctic Peninsula, there were typically fewer overlapping waves than any other geographical region. Hence, wave identification problems associated with wave superposition are likely minimised in our geographical region of interest.

The choices we have made in our Wave-ID processing will also affect the range of vertical wavelengths we detect. Fig. 4.6 shows transmission curves as a function of wavelength for each processing step in the Wave-ID method. As shown by the net transmission curve (black solid) in Fig. 4.6, the combined analysis method is generally sensitive to gravity waves with $4 < \lambda_z < 13\,\text{km}$, with a sharp cut-off below 4 km and a more gradual cut-off above 13 km.

The histogram in Fig. 4.6 shows vertical wavelengths of gravity waves identified by this method in the region 35-75°S and 0-90°W during June-August 2006-2012. The distribution of observed vertical wavelengths generally follows the net transmission curve of synthetic waves, with peak observations at $7 < \lambda < 9\,\text{km}$.

A primary limitation of the Wave-ID method is the limited vertical window, which limits maximum resolvable vertical wavelength. This is due to the limited vertical extent of the high-accuracy temperature retrieval of COSMIC GPS-RO. Extending the region upwards would reduce confidence in any resolved waves due to increased noise in measurements above z $\approx$ 38 km (Tsuda et al., 2011). If we extend the region down much further, sharp gradients in temperature around the tropopause risk introducing spurious artifacts via tra-

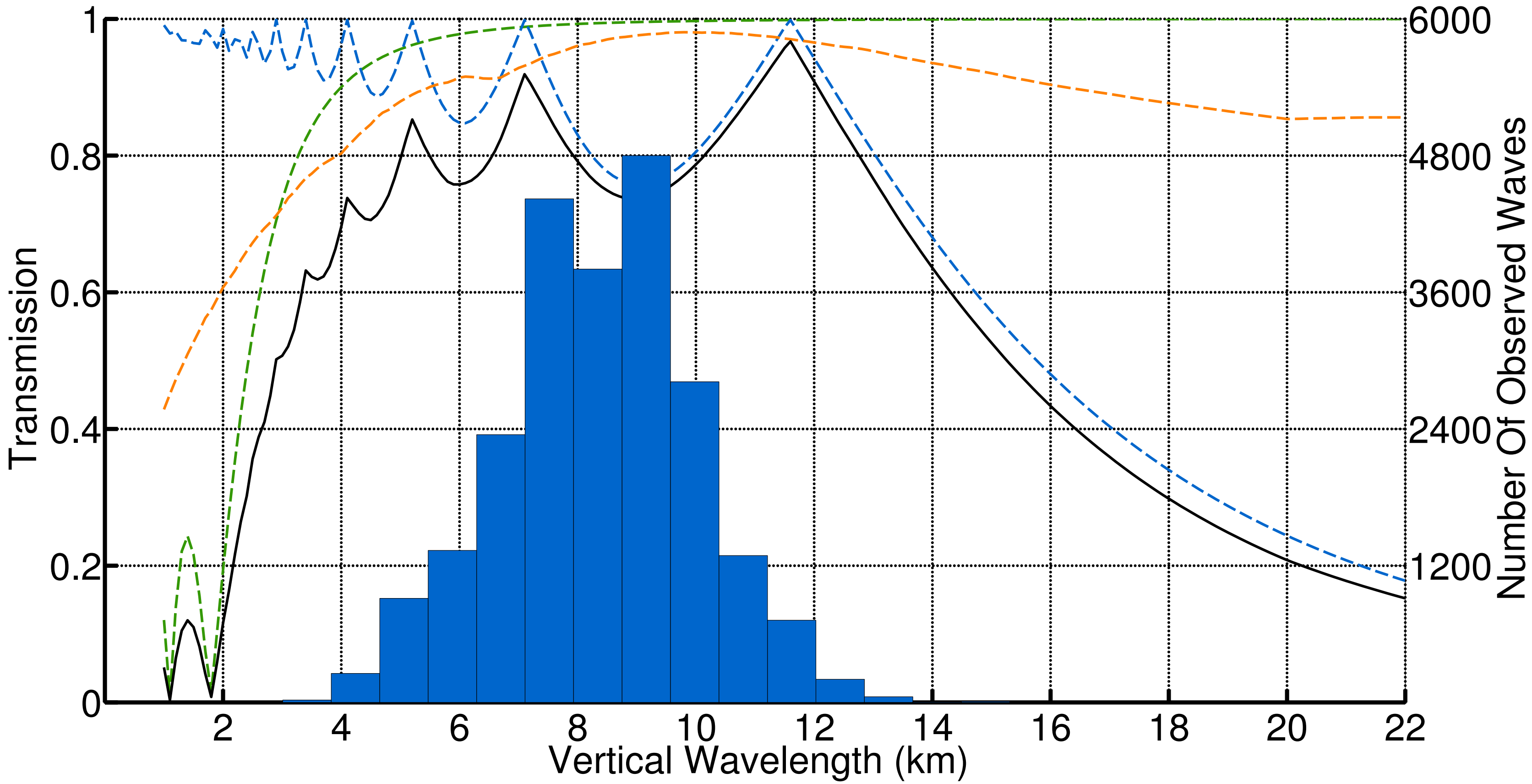

Figure 4.6: Transmission against vertical wavelength for each step in our Wave-ID processing for synthetic waves with $\lambda_Z < 22$ km centred at 30 km altitude: background subtraction (blue dashed); noise reduction (green dashed); Gaussian windowing and CWT (orange dashed); and the combined transmission (black solid). Blue bars show a histogram (right axis) of number of waves identified in COSMIC data in the region 35-75°S 0-90°W during June-August 2006-2012 using this method.

ditional filtering methods (Alexander and de la Torre, 2011). Furthermore, decreasing wave amplitudes with increasing pressure in addition to the presence of water vapour makes gravity wave study below the tropopause difficult via GPS-RO. We also implicitly assume that λ_Z does not change much with altitude, which might not hold true for the real atmosphere. This could decrease the probability that we will identify waves with longer λ_Z, which may help to explain the slight mis-match between the histogram in Fig. 4.6 and the range of permitted wavelengths (solid black curve) for longer λ_Z waves.

Future work may involve (1) optimising this vertical window so as to resolve the maximum possible range of vertical wavelengths; (2) investigating the optimum threshold value above which to consider a wave identification as positive; and (3) employing methods to identify overlapping waves as described by Wright and Gille (2013).

4.3.2 Wave identification results

In Fig. 4.7(a) we present a multi-year composite plot of E_p for June-August 2006-2012 at 30 km over the southern hemisphere. In this analysis, we take the mean E_p from all available profiles, including those where no significant waves are present. In Fig. 4.7(b) we produce another composite plot of E_p but calculated using only waves identified via the Wave-ID method described above. In other words, Fig. 4.7(a) is a time-averaged climatology of E_p in the region whereas Fig. 4.7(b) is the mean E_p of individual waves detected using the Wave-ID method during this period.

An initial observation is that much higher E_p values are apparent in Fig. 4.7(b) than in Fig. 4.7(a). This is expected, since mean E_p values in Fig. 4.7(b) are skewed by the exclusion of profiles for which no wave-like feature was detected.

The same long leeward region of increased E_p sweeping around Antarctica is present in both panels of Fig. 4.7. The largest values in both panels are generally observed just east of the southern tip of South America and the Antarctic Peninsula, decreasing eastward and reaching a minimum just west of the Drake Passage. By comparison of the two maps from the two different methods in Fig. 4.7, information about wave intermittency can be inferred. The peak of the distribution of E_p in Fig. 4.7(b) resides much closer to the mountains of the southern Andes and Antarctic Peninsula, but the rest of the distribution remains broadly co-located with the results in Fig. 4.7(a). The westward shift of the peak implies that waves close to the southern Andes and Antarctic Peninsula have on average larger amplitudes, but are more intermittent since this peak is diminished in the average of all available profiles. The rest of the distribution may therefore be less intermittent,

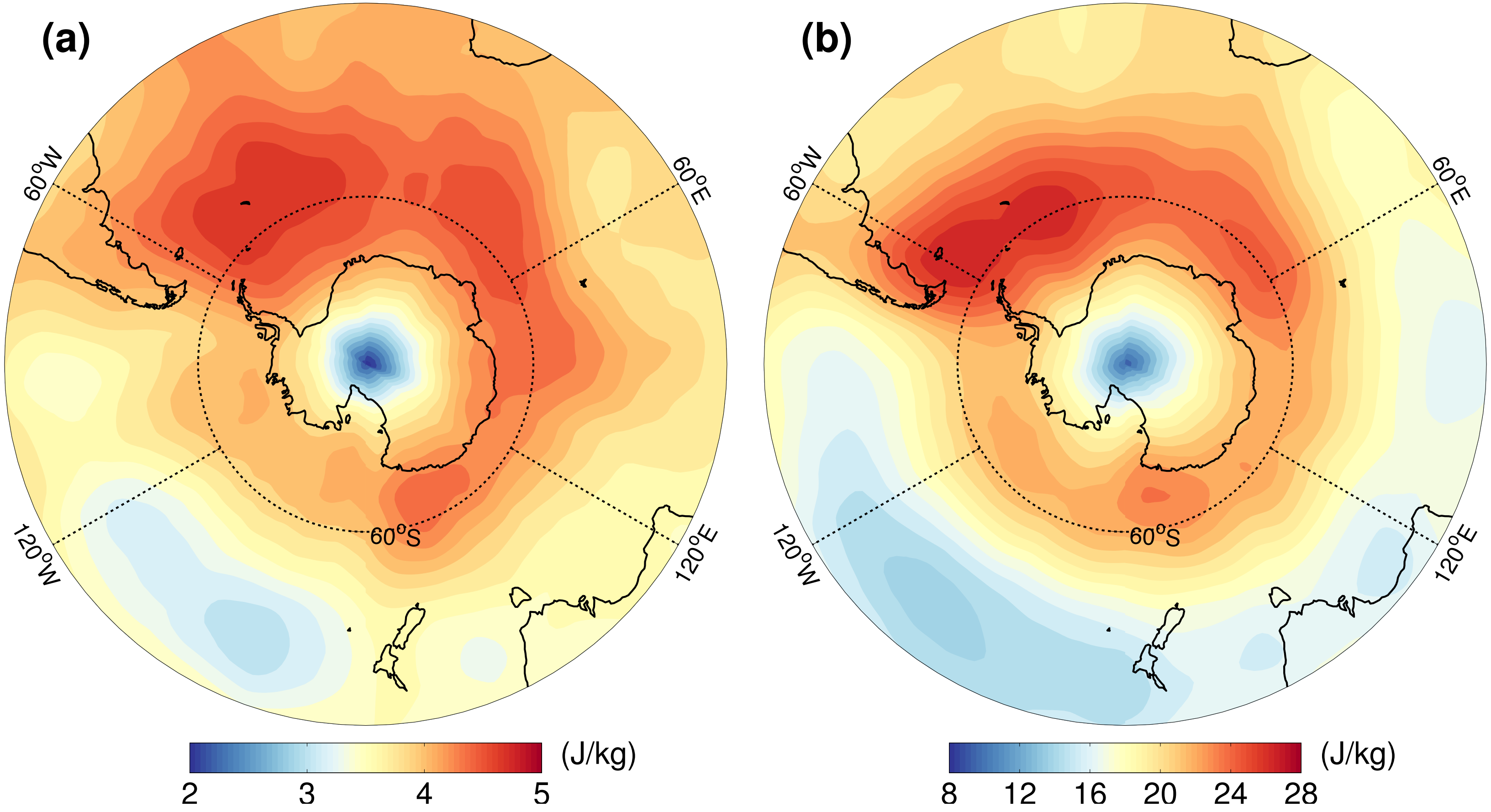

Figure 4.7: Polar stereo projections of E_p at 30 km (∼10 hPa) for June-August 2006-2012 using (a) all available COSMIC profiles and (b) only individually identified waves using the Wave-ID method. Note that (a) and (b) have different colour scales, and that we expect E_p values in (b) to be larger than in (a) since we preferentially select profiles that contain a wave (see text).

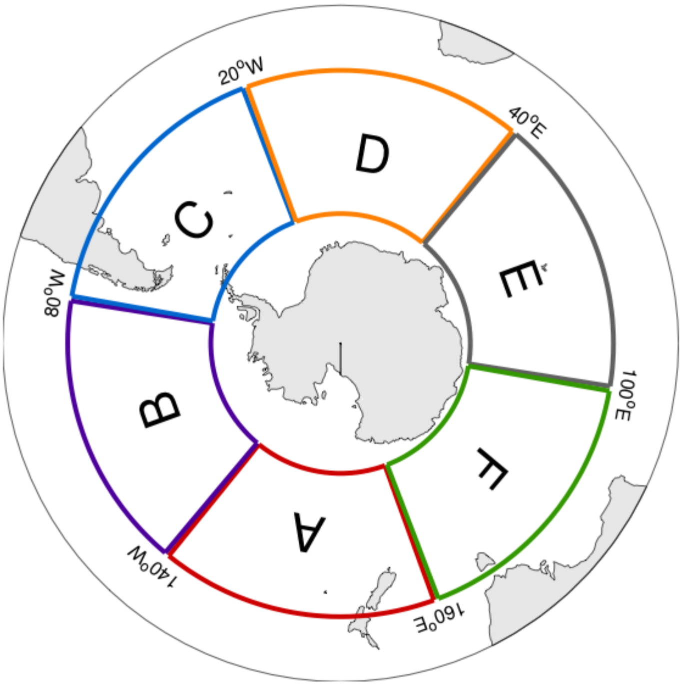

Figure 4.8: Polar stereo projection showing longitudinal Sectors A-F in the latitude band 40-65°S used in Fig. 4.9.

since it remains broadly co-located in both panels. This is consistent with the hypothesis that the region immediately east of the mountains is dominated by waves from orographic sources, which have been shown to be generally more intermittent than non-orographic sources in this region (Hertzog et al., 2008, 2012; Plougonven et al., 2013; Wright et al., 2013). A small enhancement is also evident at around 160°E 65°S that may be suggestive of a contribution from orographic waves from the Transantarctic Mountains.

To further investigate the nature of the wave field in this long leeward region of increased E_p, we divide the latitude band 40-65°S into six longitudinal sectors A-F, and examine the population of waves in each sector. Sector C contains the mountains of southern Andes, Antarctic Peninsula and South Georgia. Sector B is oceanic and upwind (westward) of these mountains. Sector D is also oceanic but immediately downwind (eastward) of the mountains. Sectors A, E and F are predominantly oceanic. Figure 4.9 presents histograms of individual wave amplitudes identified using the Wave-ID method in each of these six sectors during June-August 2006-2012. Note that these waves are from the same profiles used to produce the E_p distribution in Fig. 4.7(b).

At first glance, the histograms of wave amplitudes in each sector in Fig. 4.9 appear broadly similar. Approximately 20000 waves are identified in each sector and the modal amplitude is between 2-3 K. Upon closer inspection however, some important differences become apparent.

Despite containing around 4.5% and 12% fewer profiles than Sector B respectively, Sectors C and D contain around 13% and 6% more identified waves respectively. This indicates that the sectors containing and immediately downwind of the southern Andes and Antarctic Peninsula (C,D) contain significantly more identifiable waves than sectors immediately upwind. Furthermore, Sector B has the highest number of available profiles, yet the lowest number of identified waves of any sector.

We next investigate the relative distribution of wave amplitudes in each sector compared to the zonal mean to highlight any longitudinal variation in wave amplitude populations. The rightmost panel in Fig. 4.9 shows the difference between the histogram in each sector and the zonal mean histogram of wave amplitudes. The curves in this panel indicate that the sectors containing and downwind of the southern Andes and Antarctic Peninsula (C,D) contain significantly more large amplitude ($3 < T' < 8\,\mathrm{K}$) waves and fewer small amplitude waves ($T' < 2.5\,\mathrm{K}$) than the zonal mean, whereas upwind Sectors A, B and F contain fewer large amplitude waves and more small amplitude waves.

Three interesting conclusions are indicated by this analysis. Firstly, the geographical region downwind (eastward) of the mountains of the southern Andes and Antarctic Peninsula up to around 40°E contains significantly more identifiable gravity waves than a region of equal size upwind (westward) of the mountains.

Secondly, this downwind region contains significantly more large amplitude waves with $3 < T' < 8\,\mathrm{K}$ than the corresponding upwind region, though these large amplitude waves are still relatively rare. Since $E_p \propto (T')^2$, it is likely that the structured distribution of E_p in Fig. 4.7(b) is hence the result of an increased number of large amplitude mountain waves immediately downwind of the southern Andes and Antarctic Peninsula. In a recent study involving balloon, satellite and mesoscale numerical simulations above Antarctica and the Southern Ocean, Hertzog et al. (2012) showed that rare, large amplitude waves are not only more commonly observed above mountains in this region but that these events represent the main contribution to the total stratospheric momentum flux during the winter regime of the stratospheric circulation. Hertzog et al. also showed that gravity waves populations over open ocean tend to follow a more log-normal distribution with fewer rare, large amplitude events. Our results reinforce the findings of Hertzog et al..

Thirdly, and perhaps most interestingly, differences in the number of identified waves and the relative distribution of wave amplitudes between sectors are significant, but relatively small in absolute terms. In general, each sector has strikingly similar distributions of wave amplitudes and total numbers of identified waves. This zonal uniformity in the distributions of wave amplitudes may be suggestive of strong, zonally uniform source mechanisms for gravity waves in all sectors, such as spontaneous adjustment or jet instability around the edge of the southern stratospheric jet. This is discussed further in

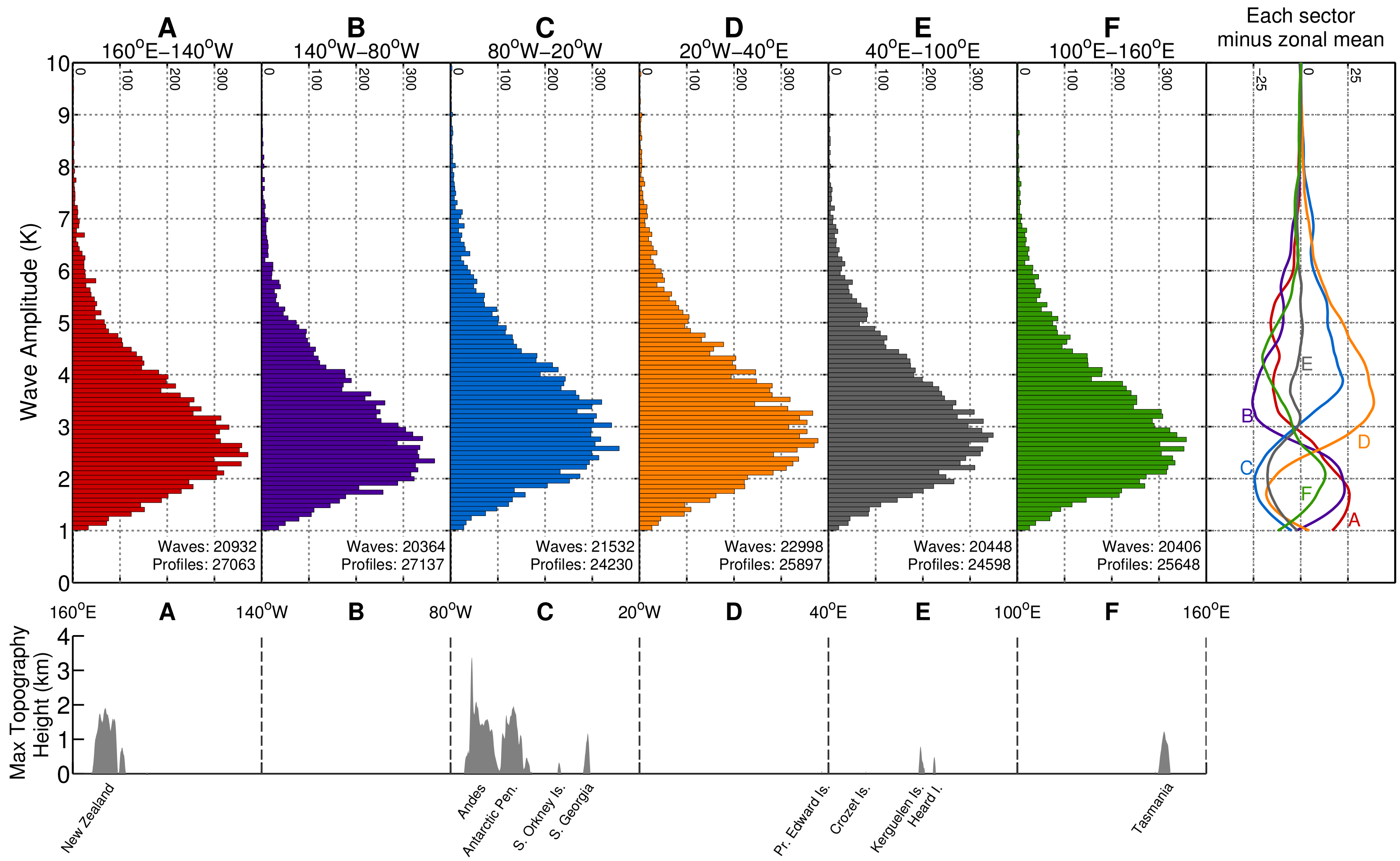

Figure 4.9: Histograms of individual wave amplitudes detected during June-August 2006-2012 in longitudinal sectors A-F in the latitude band 40-65°S using the Wave-ID method (see text). The rightmost panel shows the difference between the wave amplitude distribution in each sector and the zonal-mean distribution. The bottom panel shows maximum topography height in the latitude band 40-65°S

Section 4.5.

4.4 Gravity wave momentum fluxes during JJA 2006 using COSMIC profile-pairs

Gravity wave momentum flux is one of the key parameters characterising the effects of gravity waves in the atmosphere. This is of vital importance to the gravity wave modelling community, but typically difficult to obtain from observations (Fritts and Alexander, 2003; Alexander et al., 2010). Ern et al. (2004) showed that an approximation to the absolute value of momentum flux can be inferred from satellite observations of a gravity wave's amplitude T' and horizontal and vertical wavenumbers k_H and m. In the case of limb-sounding instruments such as HIRDLS and CRISTA, T' and m can be obtained directly from a single vertical temperature profile, while k_H can be estimated using the phase shift between adjacent profiles (Ern et al., 2004; Alexander et al., 2008). However, such k_H estimation methods have not routinely been applied to COSMIC, due to typically large inter-profile spacing. The scarcity of multiple profiles that are both closely spaced and closely timed with near-parallel lines of sight limits the accurate estimation of k_H in this way. Wang and Alexander (2010) investigated the use of 3 or more COSMIC profiles to make estimates of zonal and meridional horizontal wavenumbers k and l. However, as discussed by Faber et al. (2013), limitations in sampling density, aliasing and differing lines of sight restrict their approach being used in the general case.

Here we investigate an alternative approach for estimation of k_H from COSMIC GPS-RO data using a modified form of the method described by Alexander et al. (2008). We take advantage of the deployment phase of the COSMIC constellation, when pairs of satellites were often physically close (Liou et al., 2007). During this phase, a single occulting GPS satellite was often tracked by a close pair of COSMIC satellites, resulting in a significant number pairs of profiles that were closely spaced and closely timed, with near-parallel lines of sight. These particular profile-pairs permit the use of a k_H estimation method and subsequently an estimation of gravity wave momentum flux. k_H has also been determined in a similar manner in recent studies by McDonald (2012) and Faber et al. (2013). In this section, we use this method to make estimates of gravity wave momentum flux from COSMIC GPS-RO during June-August 2006 over the southern Andes, Drake Passage and Antarctic Peninsula.

4.4.1 Profile pair selection and processing

First, we identify profile-pairs during June-August 2006 that are closely spaced, closely timed and have near-parallel lines of sight. We require that the two profiles must (1) be horizontally separated by less than 300 km at a height of 30 km; (2) be separated in time by less than 15 minutes; and (3) have lines of sight aligned within 30° of each other. The line-of-sight requirement is important since we require that waves have $\lambda_H \geqslant 270\,\mathrm{km}$ in the line-of-sight as discussed in Section 4.1.1. If the two viewing angles differ by a large amount, the same wave might not be observable in both profiles. Finally, we require that a clear wave-like feature of approximately the same vertical wavelength ($\pm 1.5\,\mathrm{km}$) is identified in both profiles using the Wave-ID method described in section 4.3.1. A discussion of this vertical wavelength criterion as an identification method for the same wave in both profiles is provided by McDonald (2012).

In practise, we find that the majority of profile-pairs during June-August 2006 have horizontal separations $\sim$10 km (see Fig. 4.12(a)), time separations of less than a minute and lines-of-sight separated by less than 1°. Hence requirements (1), (2) and (3) are usually satisfied. The requirement that both profiles contain the same wave-like feature reduces the number of available pairs from $\sim$75000 to $\sim$14000 globally during June-August 2006. Of these, around 1300 lie in our geographical region of interest.

To estimate k_H in each profile pair, we follow a modified form of the method described by Alexander et al. (2008). We first apply a Gaussian window of FWHM = 22 km centred at 30 km altitude as described in section 4.3.1. We next compute the CWT of each profile. The resulting transform $\tilde{T}(z, \lambda_Z)$ is a complex valued function of altitude z and vertical wavelength λ_Z. For the two profiles a and b, the cospectrum $C_{a,b}$ is computed as

$$C_{a,b} = \tilde{T}_a \tilde{T}_b^* = \hat{T}_a \hat{T}_b e^{i\Delta\phi_{a,b}} \tag{4.7}$$

where $\hat{T}$ is the magnitude and $\Delta\phi_{a,b}$ is the phase difference between the two profiles for each λ_Z at each position z. The covariance spectrum is the absolute value $|C_{a,b}|$. We locate the maximum in the covariance spectrum C_{max} in the height region 20-40 km, for vertical wavelengths less than 18 km. The location of C_{max} in the covariance spectrum corresponds to the dominant vertical wavelength λ_{DOM} common to both profiles at altitude z_{DOM}. We then compute the phase difference between the two profiles $\Delta\phi_{a,b}$ as

$$\Delta\phi_{a,b} = \arctan\left(\frac{\mathrm{Im}(C_{a,b})}{\mathrm{Re}(C_{a,b})}\right) \tag{4.8}$$

where $\mathrm{Re}(C_{a,b})$ and $\mathrm{Im}(C_{a,b})$ are the real and imaginary coefficients of the covariance spectrum $C_{a,b}$. We record the value of $\Delta\phi_{a,b}$ at C_{max}.

We then compute the projection of the horizontal wavenumber k_H along the horizontal axis joining the two profiles a and b as

$$k_H = \frac{\Delta\phi_{a,b}}{\Delta r_{a,b}} \tag{4.9}$$

where $\Delta r_{a,b}$ is the horizontal separation of profiles a and b at around 30 km altitude. We then compute $\lambda_H = 2\pi/k_H$. This projected value of λ_H is typically longer than the true horizontal wavelength, and hence represents an upper-bound estimate (Ern et al., 2004). A useful illustration of this geometry can be found in (Preusse et al., 2009). We require that $100 \leqslant \lambda_H \leqslant 5000$ km to exclude unphysically short or extremely long horizontal wavelengths. This exclusion is discussed in more detail in Section 4.6.

Generally, the horizontal separation Δr of our profile-pairs is much shorter than the line-of-sight horizontal resolution $\Delta L \sim 270$ km. Therefore, any phase difference measured between the profiles is not likely to be the result of phase difference in the direction of the line-of-sight, but the result of phase difference perpendicular to the line-of-sight. For this reason we take Δr to be the perpendicular horizontal separation of the pair with respect to the line-of-sight or the first profile in each pair. This is generally close to the absolute horizontal separation due to the geometry of the constellation during the deployment phase.

To obtain an estimate of wave amplitude T', we find the maximum amplitude in each perturbation profile $T'_a(z)$ and $T'_b(z)$ over the height region $z_{\mathrm{DOM}} \pm \lambda_{\mathrm{DOM}}/2$, and take the mean.

Finally we compute an estimate of the absolute value of momentum flux M_{flux} as

$$M_{\mathrm{flux}} = \frac{\bar{\rho}}{2}\frac{\lambda_Z}{\lambda_H}\left(\frac{g}{N}\right)^2\left(\frac{T'}{\bar{T}}\right)^2 \tag{4.10}$$

where $\bar{\rho}$ is local atmospheric density, g is acceleration due to gravity and N is the Brunt-Väisälä (buoyancy) frequency.

4.4.2 COSMIC momentum flux results

Figure 4.10 shows gravity wave vertical wavelengths, horizontal wavelengths and momentum flux from our COSMIC pair analysis over the southern Andes, Drake Passage and Antarctic Peninsula during June-August 2006. Also shown are coincident results from HIRDLS, using the Stockwell Transform (S-Transform Stockwell et al., 1996) method described by Alexander et al. (2008) modified by Wright and Gille (2013). COSMIC and HIRDLS are sensitive to broadly overlapping parts of the gravity wave spectrum, so we provide results from HIRDLS as a comparison.

In Fig. 4.10(a), our COSMIC analysis shows longer mean vertical wavelengths over the southern tip of South America extending south over the Drake Passage. This southward extension out over the Drake Passage is in good agreement with a case study of a large mountain wave event in the region by Alexander and Teitelbaum (2011), using data from the Atmospheric InfraRed Sounder (AIRS) instrument. They did however infer longer vertical wavelengths due to the deep vertical weighting function of the AIRS instrument and the assumption of zero ground-based phase velocities. This region of longer vertical wavelength also extends further south over the Antarctic Peninsula.

The corresponding HIRDLS analysis in Fig. 4.10(d) shows typically longer λ_Z values overall, likely due to the increased sensitivity of HIRDLS to waves with long λ_Z as a result of the larger usable height range in HIRDLS profiles. Like our COSMIC analysis, Fig. 4.10(d) also shows longer mean vertical wavelengths over the southern tip of South America. However, a region of longer vertical wavelengths is also evident between 80-100°W that is not seen in our COSMIC analysis. We do not fully understand the reasons for this, but we suspect that it may be due to differing vertical wavelength sensitivities of HIRDLS and COSMIC. A full investigation into the distributions of vertical wavelengths from the HIRDLS S-Transform analysis is however beyond the scope of this study.

The results of our λ_H analysis from COSMIC profile-pairs is presented in Fig. 4.10(b). We mostly observe values of around 600-800 km, but no structured geographical pattern is evident. We suspect this distribution (or lack thereof) may be due to the viewing geometry of GPS-RO technique, more specifically the orientation of the horizontal axis joining the two profiles in each profile pair, which can vary significantly between pairs. Since the measured horizontal wavelength is the projection of the true λ_H along the axis between the two profiles, it is a upper-bound estimate heavily dependent the orientation of this horizontal axis. Even in a region where the wave field has a preferential horizontal alignment we will still recover a range of horizontal wavelength estimates due to differing orientations. HIRDLS scan-tracks are more consistently aligned ~NW-SE or ~NE-SW

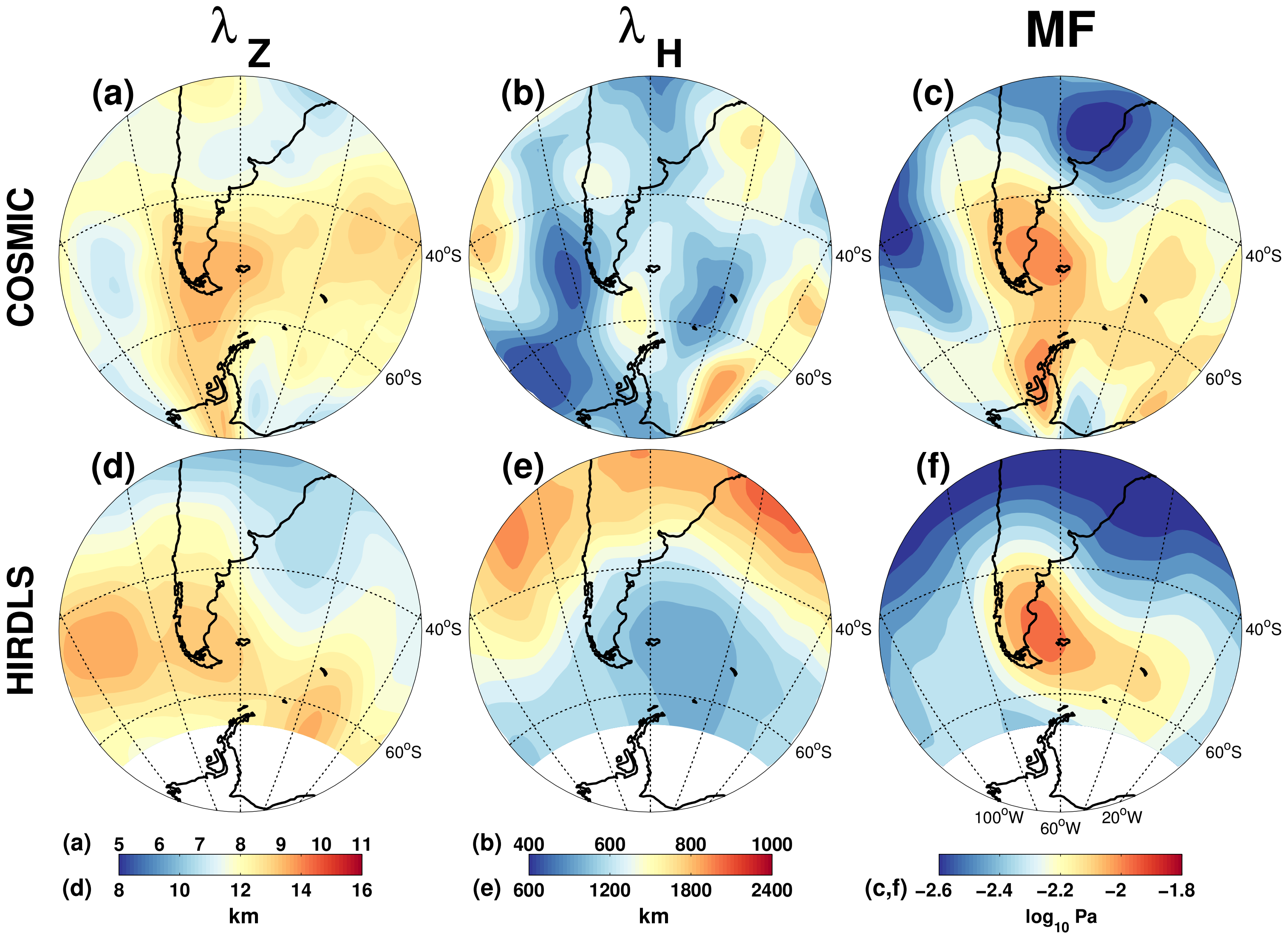

Figure 4.10: Orthographic projections of vertical wavelength λ_Z, horizontal wavelength λ_H and momentum flux (MF) for COSMIC (a,b,c) and HIRDLS (d,e,f) at 30 km (∼10 hPa) during June-August 2006.

across this region, and hence estimates of λ_H between adjacent HIRDLS profiles will be more consistent, but not necessarily more accurate. This is likely the reason that the more structured geographical distribution of λ_H shown in HIRDLS results, where shorter horizontal wavelengths are observed generally south and east of the southern tip of South America, is not observed by COSMIC.

The absolute values of our λ_H analysis are however physically reasonable and in good agreement with other studies such as Ern et al. (2004). They are however much shorter than HIRDLS estimates. Our COSMIC profile-pairs typically have smaller horizontal separations ($\sim$10 km) between profile-pairs than HIRDLS ($\sim$80 km for an ascending-descending pair, $\sim$120 km for a descending-ascending pair (e.g. Wright et al., 2015b)). This means that any absolute error in phase difference $\Delta\phi_{a,b}$ between COSMIC pairs will bias our results towards shorter λ_H than might be found in a HIRDLS profile-pair. We suspect that this may be the reason we observe lower absolute horizontal wavelength values in our COSMIC analysis than in HIRDLS. The results are not contradictory however, since both estimates represent an upper-bound. This sensitivity to errors in phase difference and their effect on λ_H estimation with regards to horizontal separation is discussed more fully in Section 4.6.

Figure 4.10(c) shows the results of our COSMIC momentum flux analysis. Two local maxima of order 10^{-2} Pa are observed over the southern tip of South America and the Antarctic Peninsula. This increased flux over the southern tip of South America is in good agreement with results from CRISTA (Ern et al., 2004) and HIRDLS (Alexander et al., 2008) and the maximum over the Antarctic Peninsula is in good agreement with results from the Vorcore superpressure balloon campaign presented in Hertzog et al. (2008). Hertzog et al. showed that most of the momentum flux in the maximum over the Antarctic Peninsula was in a westward direction, suggestive of orographic gravity waves propagating against the mean stratospheric flow. In our results, increased momentum flux is also observed to the east of the two maxima, suggestive of significant wakes of associated gravity wave flux downwind from these sources.

The HIRDLS analysis Fig. 4.10(f) shows a maximum over the southern tip of South America, consistent in location and magnitude with our COSMIC results. HIRDLS estimates of gravity wave momentum flux are slightly higher, though this could be somewhat expected since the HIRDLS analysis method used here generally resolved waves with longer vertical wavelengths than the COSMIC method. The COSMIC analysis is able to identify a secondary maximum over the Antarctic Peninsula which is not observed by HIRDLS due to the lack of measurements poleward of 62°S.

These momentum flux measurements reaffirm that the southern Andes and Antarctic

Peninsula are intense and persistent sources of gravity wave momentum flux during austral winter. Perhaps more importantly however, our results demonstrate that, given sufficient sampling density, COSMIC GPS-RO can provide physically reasonable estimates of stratospheric gravity wave momentum flux that are consistent with results from HIRDLS, CRISTA and Vorcore. The final configuration of the COSMIC constellation however restricts the number of suitable profile-pairs such that regional climatological studies of gravity wave momentum flux using our method are generally limited to the deployment phase in 2006. However, as discussed in Section 4.5, dramatically increased sampling density provided by upcoming radio occultation missions may provide an opportunity to apply this method on a global scale in coming years.

4.5 Discussion

During austral winter in the southern hemisphere, the mountains of the southern Andes and Antarctic Peninsula are a known hot spot of gravity wave fluxes (e.g. Alexander and Teitelbaum, 2007, 2011; Hoffmann et al., 2013). However, the origin of the long leeward distribution of enhanced gravity wave energy stretching eastwards far over the ocean is currently a topic for debate.

As discussed in Section 4.2.3, Sato et al. (2012) suggested that wave packets from the mountains of the southern tip of South America and northern tip of the Antarctic Peninsula could be transported significantly downwind if their horizontal wavenumber vectors are aligned at an acute angle to the mean stratospheric flow. However, using a ray-tracing analysis Sato et al. also showed that for horizontal wavelengths of 250-350 km such wave packets are rarely transported east of the prime meridian, regardless of launch angle. Hence, the distribution of increased E_p shown here eastwards of around 20°E is not likely to be explained by the downwind transportation of waves with $\lambda_H \lesssim 350\,\text{km}$. This suggests that the distribution of increased E_p eastwards of around 20°E may be the result of (1) downwind-transported mountain waves with $\lambda_H > 350\,\text{km}$; (2) locally generated non-orographic waves from tropospheric or stratospheric sources out over the ocean; or (3) some combination of these processes.

Preusse et al. (2014) used backwards ray-tracing of resolved waves in ECMWF data to show that during August 2008, waves over the southern Andes and Antarctic Peninsula overwhelmingly had lowest traceable altitude (LTA) values close to the surface, whereas waves over the southern Atlantic and Indian oceans often had average LTA values around 7-12 km. Their results are indicative of upper-tropospheric non-orographic wave sources that exist out over the oceans. Similarly, Hendricks et al. (2014) suggested that a belt of increased stratospheric gravity wave activity observed by AIRS could be attributed to non-orographic sources in winter storm tracks around the southern Atlantic and Indian Oceans. The distribution of increased stratospheric E_p in our Fig. 4.2 is morphologically reminiscent of southern hemisphere storms tracks in ECMWF ERA-40 data presented by Hoskins and Hodges (2005), which may support the suggestion by Hendricks et al.. Our Fig. 4.4 suggests that if waves from these sources significantly contribute to the region of increased stratospheric E_p over these oceans, then these waves generally have λ_Z too short ($\lesssim$ 3 km) to be resolved by COSMIC below ~20 km altitude. As these waves ascend, the mean wind speed increases and they might be refracted to longer vertical wavelengths such that they may be resolved and can contribute to the E_p in Figures 4.2 and 4.4. It should be noted however that the waves considered by Preusse et al. in ECMWF data are typically below the height region considered in this study, and the waves observed by Hendricks et al. in AIRS data are not typically visible to COSMIC.

In Section 4.2.2 we presented evidence of a southward focussing of gravity waves into the centre of the stratospheric jet. In a recent modelling study, McLandress et al. (2012) showed that zonal wind biases and vortex breakdown timing errors in a latitude band near 60°S could be greatly reduced in the Canadian Middle Atmosphere Model (CMAM) through the inclusion of non-specific orographic gravity wave drag (GWD) in the stratosphere. One hypothesis for the missing drag is unparametrized mountain waves from small islands in and around the Southern Ocean that are sub-gridscale in CMAM. A second hypothesis is the southward (northward) propagation of orographic waves from the north (south) into the southern stratospheric jet from outside the latitude band (McLandress et al., 2012; Preusse et al., 2014). Our results suggest evidence of such meridional propagation. In particular, we observe a southward focusing of waves in Fig. 4.3 into the jet around 60°W from sources further north, supporting the second hypothesis described above. It is conceivable that there exists a similar process whereby waves from the Antarctic Peninsula are focussed northwards into the jet, though we are unable to find such clear evidence for this in our results. Observational evidence of any meridional focusing is significant since many parametrization schemes used operationally in GCMs do not include such focussing phenomena (Preusse et al., 2014).

In Section 4.3 we investigated longitudinal variations in wave populations in the long leeward region of increased E_p during June-August 2006-2012. In regions immediately downwind of the southern Andes and Antarctic Peninsula we observe significantly more infrequent, large-amplitude waves than in upwind regions, while only a slight increase in the absolute number of waves is observed. Further analysis (omitted for brevity) showed that exclusion of these large amplitude waves resulted in a much more zonally uniform distribution of mean wave energy around over the Southern Ocean. This suggests that the increased E_p observed immediately downwind of the mountains in Fig. 4.7(b) is the result of increased numbers of rare, large amplitude wave events in this downwind region and not simply the result of more waves in general. As discussed in Section 4.3.2, this is consistent with the results of a super-pressure balloon and modelling study by Hertzog et al. (2012). The eastward decrease in E_p values in Fig. 4.7(b) correlates well to the eastward decrease of the frequency of occurrence of these rare, large amplitudes waves.

However, the general distributions of gravity wave amplitudes at all longitudes in the latitude band 40-65°S are broadly similar. This may be indicative of persistent, zonally uniform non-orographic source mechanisms in and around the stratospheric jet. Inertia-gravity waves, to which GPS-RO is preferentially sensitive, can often be generated at the edge of jet streams via spontaneous adjustment processes (Fritts and Alexander, 2003). Hence, a possible contribution to the long leeward region of increased E_p in Fig. 4.7(b) may be from gravity waves generated via these adjustment mechanisms.

In the context of other studies, our results therefore suggest that the long leeward region of increased E_p consists of (1) rare, large amplitude waves over 80°W-40°E from orographic sources such as the southern Andes and Antarctic Peninsula that may also have been meridionally-focussed and advected downwind; (2) a possible contribution from secondary waves generated locally in the breaking region of these primary orographic waves; (3) non-orographic waves from sources associated with winter storm tracks over the southern oceans; and (4) a zonally uniform distribution of small amplitude waves from non-orographic mechanisms such as spontaneous adjustment and jet instability around the edge of the stratospheric jet.

Finally, we described a method for the estimation of stratospheric gravity wave momentum flux from COSMIC GPS-RO. To our knowledge, there are very few studies that have successfully developed methodologies for gravity wave momentum flux estimates from GPS-RO data (e.g. Wang and Alexander, 2010; Faber et al., 2013). Our results demonstrate that, given sufficient sampling density, COSMIC GPS-RO can produce physically reasonable estimates of stratospheric gravity wave momentum flux over the southern Andes and Antarctic Peninsula that are consistent with results from CRISTA, HIRDLS and Vorcore (Ern et al., 2004; Alexander et al., 2008; Hertzog et al., 2008). It is important to note that our results have a bias towards shorter horizontal wavelength estimation for reasons discussed in Section 4.6.

The method presented here is mostly limited to the deployment phase of the COSMIC constellation only, since the number of profile-pairs that satisfy the requirements outlined in Section 4.4.1 is very low once the satellites reached their final configuration.

However, GPS-RO is an expanding technique, with new missions scheduled for launch in the next decade. The 12-satellite COSMIC-2 constellation (Cook et al., 2013) will boast more than 8000 soundings per day, measuring the occultations of satellites from the European navigation satellite system GALILEO and the Russian Global Navigation Satellite System (GLONASS), in addition to the American GPS satellite constellation. COSMIC-2 will feature two deployment phases from which large numbers of closely spaced profile-pairs can be expected. Furthermore, the number of profile-pairs available from their final configuration is likely to increase significantly and there will be increased coverage in the tropics as a result of 6 low-inclination (24°) satellites.

4.6 On the determination of λ_H from COSMIC profile-pairs

The increased number of closely-spaced profile-pairs during the deployment phase of the COSMIC constellation facilitates a focussed momentum flux study in the hot spot region. Many of these profile-pairs have very short horizontal separations $\Delta r \sim 10\,\text{km}$. The method for the estimation of λ_H described in Section 4.4.1 is inherently sensitive to error in the determination of vertical phase shift $\Delta\phi$. The short horizontal separation of these pairs may introduce a bias towards shorter horizontal wavelengths. Other than comparing our λ_H estimates to estimates from other studies as in Section 4.4.1, it is difficult to independently quantify the error and reliability of these estimates. Here we discuss the effect of short horizontal separations of COSMIC profile-pairs on the estimation of λ_H in comparison to profile-pairs from the HIRDLS mission.

4.6.1 Horizontal profile-pair separations

An estimate of horizontal wavelength λ_H can be calculated from the horizontal separation and phase difference between two adjacent profile-pairs via the relation in Equation 4.9. For a given wave field, it would be expected that in general, shorter horizontal separations between profile-pairs would result in smaller phase differences in profile-pairs.

Along the HIRDLS scan-track, vertical profiles are measured in an alternating "upscan" and "downscan" pattern. An illustration of this pattern can be seen in Wright et al. (2015a, their Fig. 4b). At an altitude of 30 km, alternating "downscan/upscan (D/U) and "upscan/downscan" (U/D) profile-pairs have horizontal separations of ~80 km and ~120 km respectively (see blue bars in our Fig. 4.12(a)).

Figure 4.11 shows histograms of gravity wave phase differences between COSMIC, HIRDLS

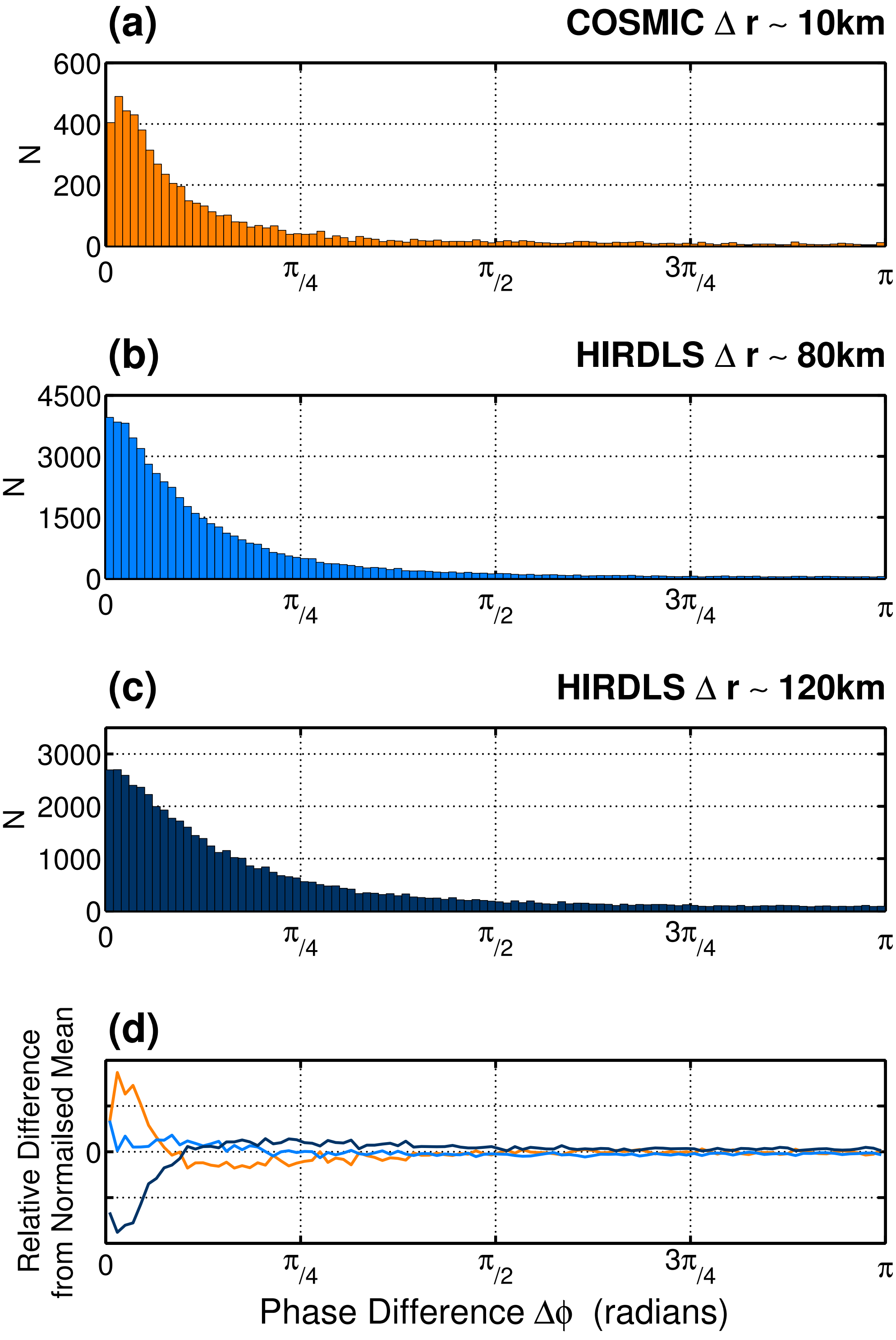

Figure 4.11: Histograms of gravity wave phase difference $\Delta\phi$ in (a) COSMIC profile-pairs, (b) HIRDLS "downscan-upscan" profile-pairs and (c) HIRDLS "upscan-downscan" profile-pairs globally during JJA 2006. Bottom panel (d) shows normalised relative difference of COSMIC (orange), HIRDLS downscan-upscan (light blue) and HIRDLS upscan-downscan (dark blue) from the mean of all three.

D/U and HIRDLS U/D profile-pairs during June-August 2006. Planetary wave features were removed from COSMIC profiles via a zonal high-pass filtering method, suppressing zonal wavenumbers $s \leqslant 6$. HIRDLS profile-pairs are processed using the method described by Wright and Gille (2013).

All three horizontal separations in Fig. 4.11 indicate a general preference towards small ($\Delta\phi < \frac{\pi}{8}$) phase differences. To investigate the relative differences between each of the distributions, we normalise each histogram such that the total number of profiles in each is equal to one. We then subtract each normalised distribution from the mean of the three to find the relative difference. The bottom panel of Fig. 4.11 indicates that COSMIC pairs with $\Delta r \sim 10\,\text{km}$ generally have more small ($\Delta\phi < \frac{\pi}{8}$) and fewer large ($\Delta\phi > \frac{\pi}{8}$) phase difference values than HIRDLS pairs. The HIRDLS U/D pairs, with the largest horizontal separation, generally have more large phase differences.

This suggests that, as might be expected, shorter horizontal separations between profile-pairs generally result in smaller phase differences. This result provides a useful sanity-check for the λ_H estimation methodology, particularly its application to COSMIC profile-pairs.

4.6.2 Biases from small phase differences

Even if the methodology is valid for horizontal separations as short as $\sim$10 km, error in the determination of $\Delta\phi$ will have a larger effect, since the method is more reliant on the determination of very small phase differences. If the absolute error in determination of $\Delta\phi$ is $\pm 0.1 \left(\sim \frac{\pi}{30}\right)$ radians, then absolute phase differences of $0 \leqslant \Delta\phi \leqslant \frac{\pi}{30}$ will be indistinguishable from each other. For a COSMIC profile-pair with $\Delta r \sim$10 km, horizontal wavelengths greater than 600 km projected along the axis joining the profile-pair would therefore be ambiguous due to this error.

The shortest theoretically resolvable horizontal wavelength from a COSMIC profile-pair is twice the horizontal separation, $2 \times \Delta r \approx 20\,\text{km}$. However, the requirement of $\lambda_H \gtrsim 270\,\text{km}$ in the line-of-sight implies the rare case where the line-of-sight is very closely aligned perpendicular to the horizontal wavenumber vector. Therefore large numbers of these very short λ_H estimates are unlikely to be physical. A cut-off of $\lambda_H \gtrsim 100\,\text{km}$ for the shortest resolvable horizontal wavelength from a COSMIC profile-pair may therefore be more realistic.

For an absolute error in phase difference of $\pm\frac{\pi}{30}$, λ_H estimates from COSMIC profile-pairs with $\Delta r \sim 10\,\text{km}$ may be accurate for $100 \lesssim \lambda_H \lesssim 600\,\text{km}$. For larger horizontal sep-

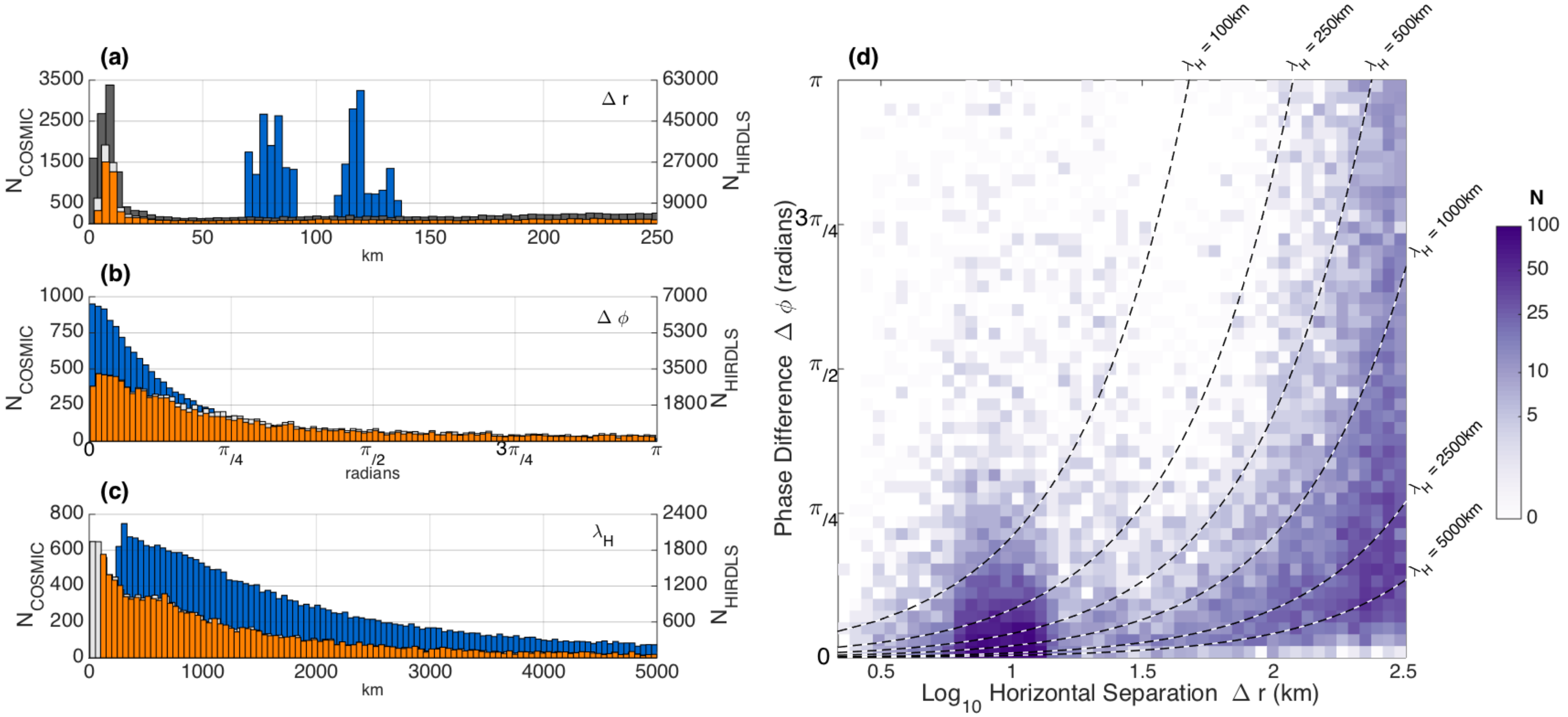

Figure 4.12: Number of COSMIC (N_{COSMIC}, orange and grey bars) and HIRDLS (N_{HIRDLS}, blue bars) profile-pairs against (a) horizontal separation Δr (b) phase difference $\Delta\phi$ and (c) projected horizontal wavelength λ_H globally for June-August 2006. Dark grey bars in (a) correspond to all available COSMIC profile-pairs. Light grey bars in (a,b,c) correspond to COSMIC profile-pairs in which a coherent wave was identified via the Wave-ID method. Orange bars in (a,b,c) correspond to COSMIC profile-pairs in which a wave was identified with $100 < \lambda_H < 5000\,\mathrm{km}$. Panel (d) shows a density plot of number of COSMIC profile-pairs N against horizontal separation and phase difference. Dashed black lines of constant λ_H are found via the relation in Eqn. 4.9.

arations and/or more accurate phase difference determinations, the upper limit is larger. Figure 4.12(d) shows a density plot of horizontal separation against phase difference for COSMIC profile-pairs in which a wave was identified via the method described in Section 4.3. Dashed black lines show lines of constant λ_H estimated via the relation in Equation 4.9. For the majority of detected waves $250 \lesssim \lambda_H \lesssim 5000\,\mathrm{km}$. A low-bias effect on the estimation of λ_H for short horizontal separations $\Delta r \sim 10\,\mathrm{km}$ due to the error in the determination $\Delta\phi$ can be seen in the bottom left corner of the panel.

In summary, we suggest that phase difference estimates from COSMIC profile-pairs from the deployment phase of the constellation are broadly in line with what we might expect when compared to HIRDLS profile-pairs. However, the typically short horizontal separations of the closely spaced COSMIC profile-pairs used in Section 4.4 are likely to introduce a low-bias in the estimation of λ_H due to error in the determination of $\Delta\phi$. We suspect that this is the reason for the differences in λ_H estimates between COSMIC and HIRDLS in Fig. 4.10(e). The estimates are not necessarily contradictory, since both represent an upper bound value, but this bias should be considered when comparing results from the two instruments using the profile-pair method described here.

4.7 Summary and Conclusions

In this study, we have used dry atmospheric temperature profiles from COSMIC GPS-RO to investigate gravity wave activity in the southern stratospheric hot spot around the southern Andes and Antarctic Peninsula. The new wavelet-based analysis technique we have presented allows quantitative identification of individual gravity waves and their properties, which we have used the determine gravity wave energies, amplitudes, momentum fluxes and short-timescale variability.

In the hot spot region, we have found clear evidence of the southward propagation of orographic gravity waves into the strong winds of the southern stratospheric jet. This phenomenon has been predicted by recent high-resolution modelling studies (e.g. Watanabe et al., 2008; Sato et al., 2009, 2012).

We also investigated the long leeward region of increased E_p stretching out over the southern oceans during austral winter. Our results suggest that this region is the result of waves from a number of overlapping orographic and non-orographic sources.

Our results, in the context of other studies, suggest that the long leeward region of increased E_p is the result of waves from a number of overlapping orographic and non-orographic wave sources. We have used the distribution of the amplitudes of individual waves to suggest that the large mean E_p values observed immediately downwind of the southern Andes and Antarctic Peninsula result from an increased number of rare, large amplitude mountain waves that have propagated downwind via the mechanism described by Sato et al. (2012). The remaining distribution is likely to be the result of waves from a variety of non-orographic sources such as storms in and around the Southern Ocean (Hendricks et al., 2014; Preusse et al., 2014) and spontaneous adjustment mechanisms around the edge of the southern stratospheric jet (Fritts and Alexander, 2003; Hei et al., 2008).

We have also described a method for the estimation of k_H from closely spaced pairs of COSMIC profiles measured during the deployment phase of the constellation in July-August 2006. We show that, given sufficient sampling density, estimations of gravity wave momentum flux in the region around the southern Andes and Antarctic Peninsula can be retrieved from COSMIC GPS-RO. These measurements use COSMIC GPS-RO to extend and complement the suite of space-borne measurement techniques available for localised study of stratospheric gravity wave momentum flux. In the coming years, the increased sampling density offered by new GPS-RO missions will allow our approach to be temporally and geographically expanded, providing further estimates of stratospheric gravity wave momentum flux on a much wider scale.

Chapter 5

A two-dimensional Stockwell transform for gravity wave analysis of AIRS measurements

In this chapter we develop a two-dimensional Stockwell transform methodology for gravity wave analysis of data from the Atmospheric Infrared Sounder (AIRS) instrument. We first derive the one-dimensional Stockwell transform (S-transform) and demonstrate its use as a tool for gravity wave analysis. We then extend the S-transform to two-dimensions and investigate the use of alternative windowing functions. Finally we use our two-dimensional Stockwell transform to measure gravity wave amplitudes, horizontal wavelengths and directions of propagation from AIRS measurements.

The key findings of this chapter are, at time of writing, accepted for final publication with the current citation:

Hindley, N. P., N. D. Smith, D. A. S. Wright, C. J. Rees, and N. J. Mitchell, 2016: A two-dimensional stockwell transform for gravity wave analysis of airs measurements. *Atmos. Chem. Phys. Disc.*, doi:10.5194/amt-2015-383

Abstract

Gravity waves (GWs) play a crucial role in the dynamics of the earth's atmosphere. These waves couple lower, middle and upper atmospheric layers by transporting and depositing energy and momentum from their sources to great heights. The accurate parametrization of GW momentum flux is of key importance to general circulation models but requires

accurate measurement of GW properties, which has proved challenging. For more than a decade, the nadir-viewing Atmospheric Infrared Sounder (AIRS) aboard NASA's Aqua satellite has made global, two-dimensional (2-D) measurements of stratospheric radiances in which GWs can be detected. However, one problem with current one-dimensional methods for GW analysis of these data is that they can introduce significant unwanted biases. Here, we present a new analysis method that resolves this problem. Our method uses a 2-D Stockwell transform (2DST) to measure GW amplitudes, horizontal wavelengths and directions of propagation using both the along-track and cross-track dimensions simultaneously. We first test our new method and demonstrate that it can accurately measure GW properties in a specified wave field. We then show that by using a new elliptical spectral window in the 2DST, in place of the traditional Gaussian, we can dramatically improve the recovery of wave amplitude over the standard approach. We then use our improved method to measure GW properties and momentum fluxes in AIRS measurements over two regions known to be intense hot spots of GW activity: (i) the Drake Passage/Antarctic Peninsula and (ii) the isolated mountainous island of South Georgia. The significance of our new 2DST method is that it provides more accurate, unbiased and better localised measurements of key GW properties compared to most current methods. The added flexibility offered by the scaling parameter and our new spectral window presented here extend the usefulness of our 2DST method to other areas of geophysical data analysis and beyond.

5.1 Introduction

Gravity waves are a vital component of the atmospheric system. These propagating mesoscale disturbances can transport energy and momentum from their source regions to great heights. They thus are a key driving mechanism in the dynamics of the middle atmosphere through drag and diffusion processes (e.g. Fritts and Alexander, 2003, and references therein).

The accurate parametrization of unresolved gravity waves in global climate models (GCMs) has proven to be a long-standing problem in the modelling community. One example of this is the "cold pole" bias (Butchart et al., 2011) suffered by nearly all GCMs, which has been suggested to be due to a deficiency of resolved and parametrized gravity wave drag near 60°S. One reason for this is that these gravity wave parametrizations continue to be poorly constrained by observations (Alexander et al., 2010). The accurate measurement of gravity wave properties is thus critical for the development of the next and current generation of climate models. In the last decade, satellite-based remote-sensing has greatly increased our capability to make gravity wave observations on a global scale, but large discrepancies between observed and modelled fluxes still remain (Geller et al., 2013). This highlights the need for more accurate global and regional gravity wave measurements.

The Atmospheric Infrared Sounder (AIRS) (Aumann et al., 2003) is a nadir-sounding spectral imager on board the Aqua satellite, launched in 2002. Part of the A-Train satellite constellation, AIRS scans the atmosphere over the range $\pm 49°$ from the nadir of the satellite in a 90 pixel ($\sim 1800\,\mathrm{km}$) wide swath, using 2378 infrared channels along a sun-synchronous polar orbit. This continuous swath is archived in granules, usually 135 pixels ($\sim 2400\,\mathrm{km}$) along-track. Stratospheric gravity waves can be detected in these granules as radiance perturbations in the 15 and 4.3 μm CO_2 emission bands (e.g. Alexander and Barnet, 2007; Hoffmann and Alexander, 2009; Hoffmann et al., 2013). AIRS measurements enable the study of stratospheric gravity waves at unprecedented horizontal resolution. In order to fully exploit these observations, accurate and easily-reproducible analysis methods for the measurement of gravity wave properties must be developed.

The Stockwell transform (S-transform) (Stockwell et al., 1996; Stockwell, 1999) is a widely-used spectral analysis technique for providing time-frequency (or distance-wavenumber) localisation of a time series (or spatial profile). This capability makes the S-transform well suited to gravity wave analysis of a variety of geophysical data (e.g. Fritts et al., 1998; Alexander et al., 2008; McDonald, 2012; Wright and Gille, 2013). The S-transform has also been used in a variety of other fields, such as the planetary (Wright, 2012), engineering (Kuyuk, 2015) and medical sciences (Goodyear et al., 2004; Yan et al., 2015).

Alexander and Barnet (2007) developed a method for measuring gravity wave amplitudes, horizontal wavelengths and directions of propagation from AIRS granules using the one-dimensional S-transform. In their method, the S-transform is computed for each cross-track row, and co-spectra between adjacent cross-track rows are used to obtain spectral information in the along-track dimension. To find the dominant waves in each granule, these cross-track co-spectra are averaged together and up to five peaks are located in each averaged spectrum. The method of Alexander and Barnet provides good first-order measurement of the properties of the (up to five) dominant wave features in a granule, but it can introduce unwanted biases as discussed further in Sect. 5.6.

Here, we present a new analysis method. AIRS radiance measurements are two-dimensional (2-D) images; thus a gravity wave analysis method using a two-dimensional Stockwell transform (2DST) is a more logical approach. In this study we present a 2DST-based method for the measurement of gravity wave amplitudes, horizontal wavelengths, and directions of propagation from AIRS measurements. Our method takes advantage of the spatial-spectral localisation capabilities of the S-transform in both dimensions simultaneously, equally and without bias.

South Georgia and the Antarctic Peninsula, together with the southern tip of South America, lie in a well-known hot spot of stratospheric gravity wave activity during austral winter, which has been extensively studied both observationally (Eckermann and Preusse, 1999; Jiang et al., 2002; Alexander and Teitelbaum, 2007; Baumgaertner and McDonald, 2007; Hertzog et al., 2008; Alexander et al., 2009; Alexander and Teitelbaum, 2011; Alexander and Grimsdell, 2013; Hindley et al., 2015) and with numerical modelling techniques (Hertzog et al., 2008; Plougonven et al., 2010; Shutts and Vosper, 2011; Hertzog et al., 2012; Sato et al., 2012; Plougonven et al., 2013) in the last decade. These mountainous regions are subjected to a strong wintertime circumpolar flow in the troposphere and stratosphere, and as a result are major orographic gravity wave sources (e.g. Hoffmann et al., 2013). Despite this, discrepancies between observed and modelled gravity wave fluxes in this region are the largest anywhere on the planet (Geller et al., 2013). This unique geography of this important region provides a “natural laboratory” in which to make clear gravity wave measurements from space, and is thus an ideal region in which to test our 2DST methodology on AIRS measurements.

In Sections 5.2 and 5.3 we introduce AIRS data and the 1-D and 2-D S-transforms. In Sect. 5.4 we apply the 2DST to a specified wave field, describing our methodology for spatial localisation of the dominant spectral components. In Sect. 5.5 we present a new alternative spectral window for use in the 2DST. In Sect. 5.6 we apply the 2DST to two selected AIRS granules measured over the Antarctic Peninsula/Drake Passage and South Georgia and discuss our results. Finally, in Sect. 5.7 we summarise the key results of this study and discuss the advantages of our 2DST method in the context of previous work.

5.2 AIRS data

Gravity waves can be detected in AIRS radiance measurements as perturbations from a background state. Here, we use AIRS Level 1B radiance measurements from the 667.77 cm^{-1} channel. These Level 1 radiances have considerably higher horizontal resolution than operational Level 2 temperature retrievals due to retrieval choices imposed on the latter (Hoffmann and Alexander, 2009). We compute brightness temperature T directly from radiance R as

$$\bar{T} = \frac{hc\nu}{k_B}\left(\ln\left(\frac{2hc^2\nu^3}{R}+1\right)\right)^{-1}, \tag{5.1}$$

where h is Planck’s constant, c is the speed of light, k_B is Boltzmann’s constant and $\nu = 667.77\,\text{cm}^{-1}$ is the wavenumber of the specified channel. Temperature perturbations T' from the local background state $\bar{T}$ are then extracted via a fourth-order polynomial fit

(Wu, 2004; Alexander and Barnet, 2007). This fit removes limb-brightening and other large-scale fluctuations. These brightness temperature perturbations are a more useful physical quantity with which to define gravity wave amplitudes, since gravity wave energies and momentum fluxes are easier defined in terms of temperature perturbations (e.g. Ern et al., 2004). Common limitations in temperature perturbation-based momentum flux calculations, such as the assumption of a monochromatic wave and observational filtering effects, do remain however and are discussed in more detail in Sect. 5.6.3.

The weighting function of the 667.77 cm^{-1} channel peaks near 3 hPa ($\sim$40 km), with a full width at half maximum of $\sim$12 km (Alexander and Barnet, 2007, also illustrated in Figure 1 of Wright et al. (2015a)). Gravity waves with vertical wavelengths shorter than 12 km are thus unlikely to be resolved and vertical wavelengths close to this limit will be strongly attenuated.

If the vertical wavelength is known, it is possible to correct for this attenuation by dividing the amplitude by an appropriate rescaling factor (Alexander and Barnet, 2007, their Fig. 4). Although methods for measuring long vertical wavelengths using multiple AIRS channels have been developed (e.g. Hoffmann and Alexander, 2009), we do not have direct measurements of vertical wavelength from our single AIRS channel, and so we do not apply such a correction to brightness temperature perturbations at this stage. The true amplitude of some waves in our initial analysis may therefore be between two and five times greater than the values shown. Later however, for the estimation of momentum flux in Sect. 5.6.3, we do apply the attenuation correction described in Alexander and Barnet (2007).

5.3 The Stockwell transform

The Stockwell Transform (S-transform) is a widely-used spectral analysis technique that can provide time-frequency (or distance-wavenumber) localisation of a time series (or profile). Developed by Robert Stockwell (Stockwell, 1999), the S-transform has been used extensively for gravity wave study (Wang et al. (2006); Alexander and Barnet (2007); Alexander and Teitelbaum (2007); Alexander et al. (2008, 2009); Wang and Alexander (2010); Hoffmann and Alexander (2009); Alexander and Teitelbaum (2011); Hoffmann et al. (2014)). The S-transform has also been used in a variety of other fields, such as planetary (Wright, 2012), engineering (Kuyuk, 2015) and medical sciences (Goodyear et al., 2004; Yan et al., 2015).

In its analytical form, the one-dimensional Stockwell transform (Stockwell et al., 1996) closely resembles a continuous wavelet transform (CWT) with a complex sinusoidal mother wavelet windowed with a scalable Gaussian window (Gibson et al., 2006). For time series data, this scalable Gaussian localises wave perturbations in the time domain through spectral localisation in the frequency domain.

5.3.1 Derivation from the Fourier transform

Here, we briefly derive the one-dimensional Stockwell transform, following the formulation Stockwell (1999).

We begin with the definition of the Fourier transform $H(t)$ for a continuous, smoothly-varying one-dimensional function of time $h(t)$

$$H(f) = \int_{-\infty}^{\infty} h(t)\mathrm{e}^{-i2\pi ft}dt \tag{5.2}$$

The Fourier transform $H(f)$ provides us with precise information as to *what* frequency components are present in $h(t)$, but not *when* in time they occurred. For many applications, this is sufficient, but gravity wave signals in one-dimensional data series often occupy a finite range, not the whole profile. Furthermore, different gravity wave signals may be present in different regions of the profile. This highlights the need for separate time-frequency localisation.

One way to achieve such localisation is to convolve $h(t)$ with an apodizing function such that each location is individually windowed, enabling the local frequency spectra to be analysed. To do this, we introduce the dimensionless variable τ, centered at $t = 0$, which describes translation along the time domain. Now we can define a normalised Gaussian window of the form

$$\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-(t-\tau)^2}{2\sigma^2}} \tag{5.3}$$

where σ is the standard deviation, which we can 'slide' along the time domain, providing a localised section of $h(t)$ at each position from which the localised frequency spectrum can be analysed. A Gaussian apodizing function is a convenient mathematical choice; it has a definite integral over an infinite range, can be compactly written analytically, and the Fourier transform of a Gaussian is another Gaussian. The S-transform $S(\tau, f)$ is thus expressed as the Fourier transform of the convolution of the function $h(t)$ and our Gaussian window:

$$S(\tau, f) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} h(t)\mathrm{e}^{\frac{-(t-\tau)^2}{2\sigma^2}}\mathrm{e}^{-i2\pi ft}dt \tag{5.4}$$

where σ is the standard deviation, representing the characteristic width of the Gaussian window. We can see from Eqn. 5.4 that the use of a single value of σ for all frequency voices in S-transform is not practical, since such a window would preferentially under-sample lower frequencies that occupy portions of $h(t)$ larger than 1σ.

To solve this problem, we scale the width of the windowing Gaussian with frequency as $\sigma = \frac{c}{|f|}$, where c is a scaling parameter usually equal to one. The S-transform can now be written in its more common form as

$$S(\tau, f) = \frac{|f|}{c\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) \mathrm{e}^{\frac{-(t-\tau)^2 f^2}{2c^2}} \mathrm{e}^{-i2\pi f t} dt \tag{5.5}$$

The scaling parameter c turns out to be very useful, as will be shown later. Setting $c < 1$ narrows the temporal Gaussian window, which broadens the corresponding spectral window. This provides increased time localisation at the expense of frequency localisation. Conversely, setting $c > 1$ broadens the temporal window and narrows the spectral window, providing increased frequency localisation at the expense of time localisation (Mansinha et al., 1997a; Fritts et al., 1998; Pinnegar and Mansinha, 2003). Taken to the extreme, a very large value for c results in such a broad temporal window that the S-transform evaluated at each location in time begins to look very similar to the Fourier transform of the entire time series, negating its usefulness. By setting different values for c, we can "tune" the localisation capabilities of the S-transform to our exact requirements for a given data set. This capability has not generally been exploited in recent gravity wave studies (e.g. Alexander and Barnet, 2007; Alexander et al., 2008; Wright and Gille, 2013). We explore the advantages of adjusting c later in this chapter.

5.3.2 Derivation from Continuous Wavelet Transform

The S-transform shares many of its time-frequency localisation capabilities with the continuous wavelet transform (CWT) (Gibson et al., 2006). In fact, we can easily arrive at the expression in Eqn. 5.5 by recalling the definition of the CWT $W(\tau, f)$ for our one-dimensional function of time $h(t)$:

$$W(\tau, f) = \int_{-\infty}^{\infty} h(t) \psi(t - \tau, f) dt \tag{5.6}$$

If we define the analysing wavelet $\psi(t - \tau, f)$ as a normalised, Gaussian-windowed sinusoid of the form

$$\psi(t - \tau, f) = \frac{|f|}{c\sqrt{2\pi}} \mathrm{e}^{\frac{-(t-\tau)^2 c^2}{2f^2}} \mathrm{e}^{-i2\pi f t} \tag{5.7}$$

and substitute it into Eqn. 5.6, we can easily see that we recover exactly the same formulation as the S-transform in Eqn. 5.5. However, since the wavelet in Eqn. 5.7 does not have zero mean, it is not strictly an admissible for the CWT. Although very similar time-frequency localisation is achieved, the complex coefficients of such a CWT will not take their usual meanings, i.e. not correlation coefficients. Furthermore, in order to computationally implement the S-transform on real physical data (i.e. data points quantised over a finite range finite) efficiently, we implement the discretized (rather than analytic) form, at which point the S-transform becomes very distinct from the concepts of the CWT.

For a smoothly-varying, continuous and one-dimensional function of time $h(t)$, the generalised analytical form of the S-transform $S(\tau, f)$ (e.g. Pinnegar and Mansinha, 2003) is given as

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t)\, \omega_g(t-\tau, f)\, e^{-i2\pi ft} dt, \tag{5.8}$$

where τ is translation in the time domain, f is frequency and $\omega_g(t-\tau, f)$ is a windowing function, scaled with frequency, that provides spatial and spectral localisation. Traditionally, $\omega_g(t-\tau, f)$ takes the form of the normalised Gaussian window

$$\omega_g(t-\tau, f) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(t-\tau)^2}{2\sigma^2}} \tag{5.9}$$

where σ is the standard deviation. A key aspect of the Gaussian window in Eqn. 5.9 is that the standard deviation is scaled for each frequency as $\sigma = \frac{c}{|f|}$, where c is a scaling parameter usually set to 1 (Mansinha et al., 1997a). This window is often referred to as the "voice Gaussian", which provides localisation of a specific frequency "voice" (Stockwell, 1999). Another key aspect of the Gaussian in Eqn. 5.9 is the normalisation factor $1/\sigma\sqrt{2\pi}$, which ensures that the integral of the window over all t is equal to unity, a requirement for any windowing function used in the S-transform. Substituting Eqn. 5.9 into Eqn. 5.8 allows us to write the S-transform more explicitly as

$$S(\tau, f) = \frac{|f|}{c\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{(t-\tau)^2 f^2}{2c^2}} e^{-i2\pi ft} dt. \tag{5.10}$$

Typically, the scaling parameter c is set to 1 (e.g. Stockwell et al., 1996; Alexander et al., 2008; Wright and Gille, 2013), but it may also be set to other values to achieve more specific time-frequency localisation requirements (e.g. Mansinha et al., 1997b; Fritts et al., 1998; Pinnegar and Mansinha, 2003). Setting $c > 1$ provides enhanced frequency localisation at the expense of time localisation, and contrarily setting $c < 1$ achieves enhanced time localisation at the expense of frequency localisation. This effect is discussed in more detail in Sect. 5.5.

5.3.3 Frequency-domain formulation

To compute the S-transform using the form in Eqn. 5.10, it seems we must compute a convolution involving the voice Gaussian and the time series for each frequency voice f, which can become quite computationally intensive. Stockwell (1999, their Sect. 3.5.1) showed that under the convolution theorem (Brigham, 1974), the time-domain convolution in Eqn. 5.10 could be written as a frequency-domain multiplication - a less computationally intensive operation - by the steps described below:

If we first recall the convolution theorem

$$p(\tau, f) * g(\tau, f) = \int_{-\infty}^{\infty} p(t, f) g(\tau - t, f) dt \tag{5.11}$$

we can see that the right hand side of this equation exactly equals the right hand side of our S-transform formulation in Eqn. 5.5 if

$$p(t, f) = h(t) e^{-i2\pi f t} \tag{5.12}$$

and

$$g(\tau - t, f) = \frac{|f|}{c\sqrt{2\pi}} e^{\frac{-(t-\tau)^2 c^2}{2f^2}} \tag{5.13}$$

Hence, using the convolution theorem, we can rewrite Eqn. 5.5 as

$$S(\tau, f) = p(\tau, f) \star g(\tau, f) \tag{5.14}$$

where

$$p(\tau, f) = h(\tau) e^{-i2\pi f \tau} \tag{5.15}$$

and

$$g(\tau, f) = \frac{|f|}{c\sqrt{2\pi}} e^{\frac{-\tau^2 f^2}{2c^2}} \tag{5.16}$$

Now let us define $B(\alpha, f)$ as the Fourier transform of the S-transform $S(\tau, f)$:

$$B(\alpha, f) = \int_{-\infty}^{\infty} S(\tau, f) e^{-i2\pi\alpha\tau} d\tau \tag{5.17}$$

where α is the corresponding frequency translation for our time translation variable τ. By also defining $P(\alpha, f)$ and $G(\alpha, f)$ as the Fourier transforms of $p(\tau, f)$ and $g(\tau, f)$ respectively, we can rewrite Eqn. 5.14 as

$$B(\alpha, f) = P(\alpha, f) G(\alpha, f) \tag{5.18}$$

which expresses the Fourier transform of the S-transform as a multiplication in the frequency domain. We express $P(\alpha, f)$ and $G(\alpha, f)$ as

$$P(\alpha, f) = \int_{-\infty}^{\infty} h(\tau)e^{-i2\pi f\tau}e^{-i2\pi\alpha\tau}d\tau = \int_{-\infty}^{\infty} h(\tau)e^{-i2\pi(\alpha+f)\tau}d\tau = H(\alpha + f) \quad (5.19)$$

and

$$G(\alpha, f) = \frac{|f|}{c\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-\tau^2 f^2}{2c^2}}e^{-i2\pi\alpha\tau}d\tau = e^{\frac{-2\pi^2c^2\alpha^2}{f^2}} \quad (5.20)$$

and substitute into Eqn. 5.18 to give

$$B(\alpha, f) = H(\alpha + f)e^{\frac{-2\pi^2c^2\alpha^2}{f^2}} \quad (5.21)$$

where $H(\alpha + f)$ is a shifted version of $H(\alpha)$, which is in turn the frequency analogue of $H(t)$. Finally, the S-transform $S(\tau, f)$ is found by taking the inverse Fourier transform of $B(\alpha, f)$, namely:

$$S(\tau, f) = \int_{-\infty}^{\infty} H(\alpha + f)e^{\frac{-2\pi^2c^2\alpha^2}{f^2}}e^{i2\pi\alpha\tau}d\alpha \quad (5.22)$$

The frequency-domain form of the voice Gaussian, denoted by $\omega_g(\alpha, f)$, is given as

$$\omega_g(\alpha, f) = e^{\frac{-2\pi^2c^2\alpha^2}{f^2}} \quad (5.23)$$

The standard deviation σ_α of this frequency-domain Gaussian window in Eqn. 5.23 scales with frequency as $\sigma_\alpha = |f|/c$. Note that this voice Gaussian is unnormalised; its peak value is equal to 1 in the frequency domain.

In this frequency-domain form, the S-transform is computed for each frequency voice f as the inverse Fourier transform of the product of $H(\alpha+f)$ and the corresponding frequency-domain voice Gaussian $\omega_g(\alpha, f)$ in Eqn. 5.23. Crucially, writing the S-transform as the frequency-domain multiplication in Eqn. 5.22 enables computationally efficient ("fast") discrete Fourier transform (DFT) algorithms and simple multiplication operations to be used. The S-transform is most commonly implemented in this manner within the atmospheric sciences.

5.3.4 The discrete S-transform

In order to implement the S-transform in Eqn. 5.22 computationally, its discrete analogue must be found.

For a one-dimensional, uniformly-sampled discrete function of time $h[k\Lambda]$ with total number of elements N, where Λ is the sampling interval time between data points and

$k = 0, 1, ...N-1$, the discrete Fourier transform is given by

$$H\left(\frac{n}{N\Lambda}\right) = \sum_{k=0}^{N-1} h[k\Lambda] e^{\frac{-i2\pi nk}{N}} \tag{5.24}$$

where $\frac{n}{N\Lambda}$ is the quantised form of frequency f for $n = 0, 1, ...N-1$. If we also discretize our time domain translation variable $\tau = j\Lambda$ where $j = 0, 1...N-1$, we can easily write the discretized form of the S-transform in Eqn. 5.5 as

$$S\left[j\Lambda, \frac{n}{N\Lambda}\right] = \sum_{k=0}^{N-1} h[k\Lambda] e^{\frac{-(k\Lambda - j\Lambda)^2 f^2}{2c^2}} e^{\frac{-i2\pi nk}{N}} \tag{5.25}$$

where a summation for all k replaces the integral for all time.

Finally, by discretizing the frequency-domain translation variable $\alpha = \frac{m}{N\Lambda}$ where $m = 0, 1, ...N-1$, we can write the discretised form the the frequency-domain form of the S-transform in Eqn. 5.22 as

$$S\left[j\Lambda, \frac{n}{N\Lambda}\right] = \sum_{m=0}^{N-1} H\left[\frac{m}{N\Lambda}\right] e^{\frac{-2\pi^2 c^2 (m-n)^2}{n^2}} e^{\frac{i2\pi mj}{N}} \tag{5.26}$$

5.3.5 Computational Implementation of the S-transform

Here, our implementation of the one-dimensional S-transform is based upon Matlab source code written by Robert Stockwell, which has become widely available online. This approach uses the computationally efficient ("fast") discrete Fourier transform (DFT) alogrithm and its inverse to implement the discrete S-transform in Eqn. 5.26. Unfortunately, Stockwell's approach also uses loops to compute the inverse FFT for each frequency voice sequentially, which is quite a slow approach for an interpretive language like Matlab, so further development in implementing the discrete S-transform without loops would be very useful.

We can see from Eqn. 5.26 that the zeroth frequency voice in the S-transform is identically equal to zero. $S(\tau, 0)$ is thus defined as independent of time, and equal to the time average of the input function $h(t)$ Stockwell (1999). This adds no extra information, so we exclude the zeroth frequency from our subsequent analysis.

In the derivation above, we have defined an input data source $h(t)$ as a function of time, and the corresponding spectral domain to be a function of frequency. This is the most common formulation when describing spectral analysis techniques. For the analysis of

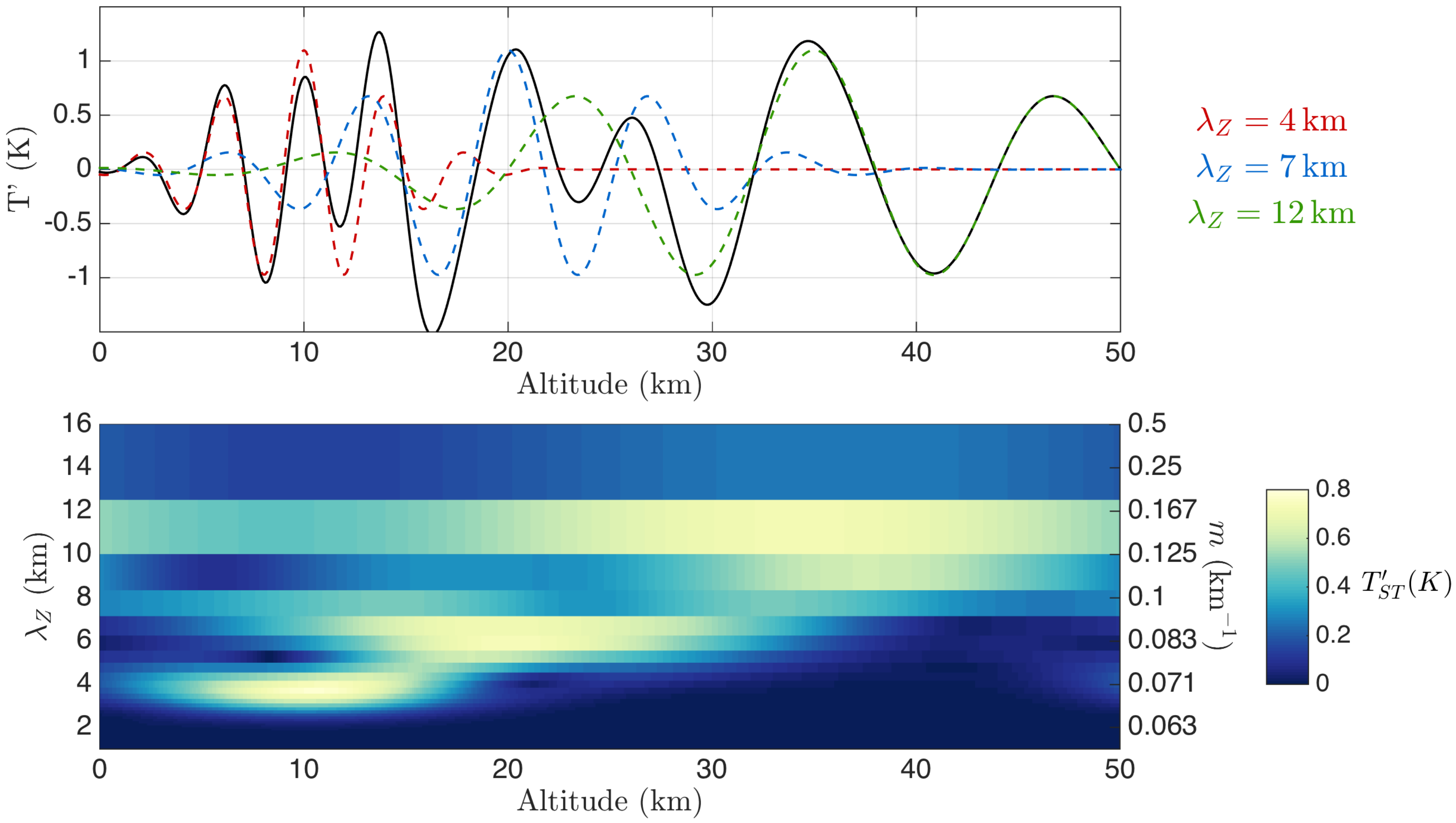

Figure 5.1: Synthetic temperature perturbation altitude profile $T'(z)$ (top panel, black solid line) and S-transform co-spectrum (bottom panel). The synthetic profile consists of three overlapping waves (red, blue and green dashed lines) with unit amplitude but different wavelengths. The waves are each localised using a Gaussian with variance equal to approximately two wavelengths.

geophysical data however, many input dats sources are profiles of some physical variable as a function of distance x. It is equivalent to write the S-transform in terms of distance and its corresponding spectral form, wavenumber k_x.

Figure 5.1 shows the one-dimensional S-transform applied to a synthetic vertical profile of temperature perturbations $T'(z)$ (solid black line), which consists of three overlapping waves with unit amplitude and different vertical wavelengths shown by the red, blue and green dashed lines.

The S-transform co-spectrum is shown in the bottom panel. Three distinct peaks are seen corresponding to the positions of the maximum amplitude of each of the synthetic waves. At long wavelengths, spectral resolution is coarse, since only wavelengths that are integer fractions of the profile length are permitted voices in the DFT. We also observe the the S-transform fails to completely recover wave amplitude, with measured values around 80% or lower of their input value. This is due in part to the non-continuous nature of the synthetic waves - we have windowed them to a specific location - but also due to the use of a Gaussian window.

As is discussed in more detail later, spectral peaks in the DFT have a characteristic width, such that spectral power is spread and shared between small groups of wavenumber voices. The Gaussian function is equal to unity at its central location, but immediately begins to decrease with increasing distance from this central location. This means that power spread over adjacent wavenumbers is reduced by the function and thus lost when the S-transform is computed by the inverse of this windowed spectrum. In section 5.5 we explore ways to solve this problem.

The S-transform has a number of desirable characteristics for geophysical data analysis. Unlike a CWT, the absolute magnitudes of the complex-valued S-transform coefficients in $S(\tau, f)$ are directly related to the true underlying amplitude of the corresponding frequency voice f at each location τ. Information regarding wave amplitude is not strictly recoverable from a CWT, since the corresponding CWT coefficients are psuedo-correlation coefficients between the signal and the analysing wavelet.

One disadvantage to using fast DFT algorithms in an S-Transform implementation is the familiar coarse wavelength resolution at low frequencies, a limitation not encountered by the CWT. Since both the S-transform and DFT algorithms are easily extended to higher dimensions however, the reduced computational expense of a DFT-based S-transform makes this a practical tool for large 2-D datasets. Retention of the wave amplitude information in the S-transform is another key advantage.

5.3.6 The two-dimensional Stockwell transform

The S-transform is easily extended to higher dimensions. For a two-dimensional image $h(x, y)$, the two-dimensional S-transform (2DST) is given by (Mansinha et al., 1997a; Stockwell, 1999)

$$S(\tau_x, \tau_y, f_x, f_y) = \frac{|f_x||f_y|}{2\pi c^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y)\, e^{-\left(\frac{(x-\tau_x)^2 f_x^2 + (y-\tau_y)^2 f_y^2}{2c^2}\right)}\, e^{-i2\pi(f_x x + f_y y)}\, dx\, dy \tag{5.27}$$

where τ_x, τ_y are translation in the x and y directions respectively. Here, f_x and f_y are simple spatial frequencies (inverse of wavelength) in the x and y directions respectively, following the notation of Stockwell (1999). For the remainder of the chapter however, we switch to using angular wavenumbers $k_x = 2\pi f_x$ and $k_y = 2\pi f_y$, since this notation is more commonly used in the atmospheric sciences. Rewriting Eqn. 5.27 in terms of angular wavenumbers k_x and k_y gives

$$S(\tau_x, \tau_y, k_x, k_y) = \frac{|k_x||k_y|}{8\pi^3 c^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y)\, e^{-\left(\frac{(x-\tau_x)^2 k_x^2 + (y-\tau_y)^2 k_y^2}{8\pi^2 c^2}\right)}\, e^{-i(k_x x + k_y y)}\, dx\, dy \tag{5.28}$$

The Gaussian windowing term in Eqn. 5.28 describes the 2-D voice Gaussian $w_g(x - \tau_x, y - \tau_y, k_x, k_y)$, where

$$w_g(x, y, k_x, k_y) = \frac{|k_x||k_y|}{8\pi^3 c^2}\, e^{-\frac{k_x^2 x^2 + k_y^2 y^2}{8\pi^2 c^2}} \tag{5.29}$$

This is the 2-D form of the 1-D Gaussian window in Eqn. 5.9. Here, the standard deviations of the 2-D Gaussian window in Eqn. 5.29 are scaled with wavenumber in the x and y directions as $2\pi c/|k_x|$ and $2\pi c/|k_y|$, where c is a scaling parameter.

As discussed in Sect. 5.3, greater computational efficiency is achieved by computing the 2DST as an operation in the wavenumber domain as

$$S(\tau_x, \tau_y, k_x, k_y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(\alpha_x + k_x, \alpha_y + k_y)\, e^{-\left(\frac{2\pi^2 c^2 \alpha_x^2}{k_x^2} + \frac{2\pi^2 c^2 \alpha_y^2}{k_y^2}\right)}\, e^{i(\alpha_x \tau_x + \alpha_y \tau_y)}\, d\alpha_x\, d\alpha_y \tag{5.30}$$

where $H(\alpha_x + k_x, \alpha_y + k_y)$ is a shifted version of $H(\alpha_x, \alpha_y)$, which is in turn the wavenumber analogue of the input image $h(x, y)$. Here wavenumbers k_x and k_y are used to scale, in α_x and α_y directions respectively, the standard deviations of the wavenumber-domain form of the 2-D voice Gaussian $W_g(\alpha_x, \alpha_y, k_x, k_y)$, which is given as

$$W_g(\alpha_x, \alpha_y, k_x, k_y) = e^{-\left(\frac{2\pi^2 c^2 \alpha_x^2}{k_x^2} + \frac{2\pi^2 c^2 \alpha_y^2}{k_y^2}\right)} \tag{5.31}$$

The 2DST is introduced and well-described by Mansinha et al. (1997a) and Mansinha et al. (1997b), who demonstrated its promise for pattern analysis. It has since been discussed and applied in a variety of fields (Liu and Wong, 2007; Kocahan et al., 2008; Liu, 2009; Barry et al., 2012) but to our knowledge it has yet to be used for geophysical data analysis in the atmospheric sciences, despite the wide use of the 1-D form. In the following section, we describe our 2DST implementation methodology for the purpose of gravity wave analysis from 2-D data.

5.4 2DST analysis of a specified wave field

To assess the capabilities of the 2DST, it is logical to first apply it to a two-dimensional specified wave field containing synthetic waves with known characteristics.

We create a specified wave field $h(x, y)$ with dimensions 100×100 km containing synthetic waves with unit amplitudes and known wavelengths. Wave amplitudes are defined as temperature perturbations T' in units of Kelvin. The synthetic waves are localised around their central locations with Gaussian functions (although note that they do overlap). We also add random ("salt and pepper") noise up to 10% of the wave amplitude.

We first compute the 2-D DFT $H(\alpha_x, \alpha_y)$ of our specified wave field $h(x, y)$. To recover an estimate of the underlying wave amplitude, we use the familiar symmetry around the zeroth frequency in the Fourier domain to recover a 2-D analogy of the analytic signal, following the approach of Stockwell (1999). A 2-D DFT contains four quadrants that contain coefficients which are in complex-conjugate pairs with the coefficients in the opposite quadrant. The sum of these pairs always yields a real signal. By setting the coefficients of two of these quadrants to zero, and doubling their opposite quadrants, we obtain a complex-valued image when we take the inverse DFT. The magnitude of this image is analogous to the underlying wave amplitude, while the complex part describes instantaneous phase. All coefficients not in a complex conjugate pair are unchanged. The full 2DST spectrum $S(\tau_x, \tau_y, k_x, k_y)$ can then be computed by taking the inverse 2-D DFT of the product of the shifted spectrum $H(\alpha_x - k_x, \alpha_y - k_y)$ and the corresponding voice Gaussian $W_g(\alpha_x, \alpha_y, k_x, k_y)$ for each wavenumber voice k_x and k_y.

For our purposes, we do not need to evaluate $S(\tau_x, \tau_y, k_x, k_y)$ for all positive and negative

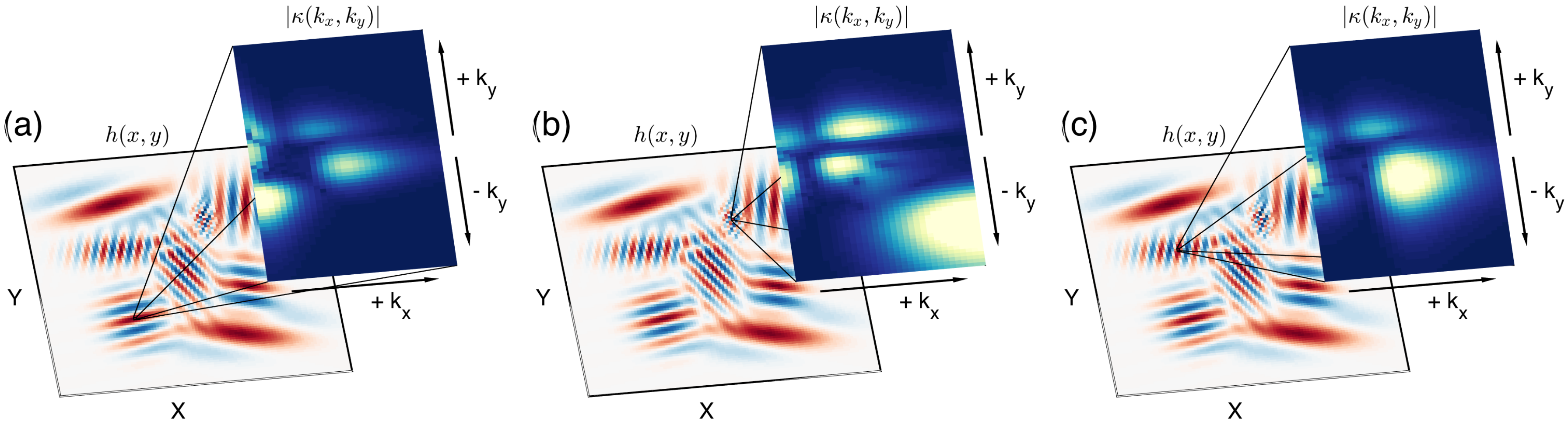

Figure 5.2: The specified wave field $h(x, y)$ (background) for which the two-dimensional Stockwell transform (2DST) has been computed. The absolute magnitudes of the localised 2DST wavenumber spectra $|\kappa(k_x, k_y)|$ (foreground) are plotted for three separate locations in panels (a), (b) and (c).

wavenumbers, since evaluating all positive and negative values of k_y and only the positive values of k_x gives us all the degrees of freedom. There is a residual 180° ambiguity in wave propagation direction which cannot be broken without additional information, which is supplied in Sect. 5.6.3.

A useful aspect of our implementation is that, like the CWT, we can compute the 2DST for any individual or range of permitted wavenumber voices by applying the appropriate wavenumber-scaled Gaussian windows. Although the permitted wavenumber voices in the spectral domain are evenly spaced, their corresponding wavelengths are limited to integer fractions (i.e. $1/n$ where $n = 0, 1, 2 \ldots N-1$) of the number of elements N in each dimension. This is an unavoidable consequence of using computationally-efficient DFT algorithms, which results in the familiar coarse spectral resolution seen at long wavelengths.

The ability to analyse an image at specific wavenumbers is a desirable aspect in geophysical data analysis, where some a priori information regarding the spectral range of wavenumbers detectable in a given dataset can be used to reduce the impact of unphysical, spurious or noisy results in 2DST analysis.

5.4.1 Measuring gravity wave properties

The 2DST $S(\tau_x, \tau_y, k_x, k_y)$ of our specified wave field is a four-dimensional object. For each location in $h(x, y)$, a two-dimensional complex-valued image of the localised spectral coefficients $\kappa(k_x, k_y)$ is evaluated.

Figure 5.2 shows a specified wave field $h(x, y)$ for which the 2DST has been computed. The absolute magnitude of the localised two-dimensional wavenumber spectrum $|\kappa(k_x, k_y)|$ is plotted for three different example locations. The coefficients of $|\kappa(k_x, k_y)|$ can be directly interpreted as the underlying amplitudes of waves with wavenumbers k_x and k_y at a given location in the specified wave field. As discussed in previous studies (e.g. Wright and Gille, 2013; Wright et al., 2015b), there are likely to be multiple peaks in $|\kappa(k_x, k_y)|$ corresponding to overlapping waves at the same location in $h(x, y)$. Indeed, in Fig. 5.2(b) we examine a location in the specified wave field where a small, high-wavenumber wave is located at the intersection of four lower-wavenumber waves. The localised spectrum computed by the 2DST shown in the foreground represents this feature well. The maximum spectral response is located in a peak at high k_x and k_y wavenumbers, with four smaller spectral peaks at lower wavenumbers with lower spectral responses.

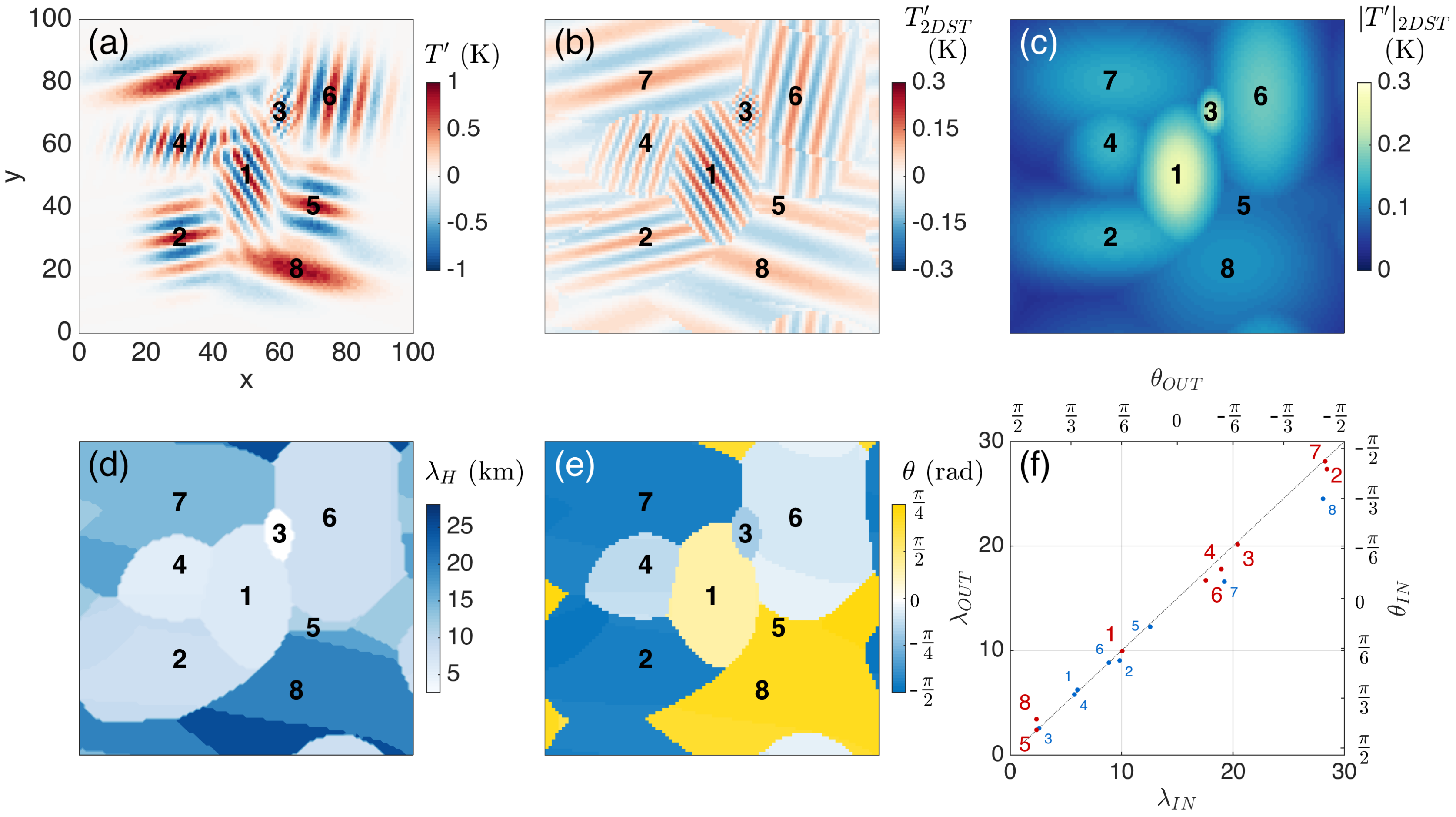

Figure 5.3: The specified wave field (a), containing synthetic waves numbered 1 to 8, for which our two-dimensional Stockwell transform analysis has been performed. The "reconstructed" wave field T'_{2DST}, underlying wave amplitudes $|T'|_{2DST}$, horizontal wavelengths λ_H and directions θ (measured anticlockwise from the positive x direction) are shown in panels (b), (c), (d), and (e) respectively. Distances and wavelengths have units of kilometres and amplitudes have units of Kelvin. Panel (f) compares input and measured wavelengths (blue numbered dots) and input and measured propagation angles (red numbered dots) for the eight synthetic waves. The dashed grey line in (f) shows 1:1 correspondence.

A four-dimensional complex-valued function can be difficult to visualise. A more useful product might be a series of two-dimensional images, the same size as the input image, that contain the characteristics of the dominant wave at each location. In the implementation presented here, we neglect overlapping waves and identify a single dominant wave for each location in $h(x, y)$.

For each such location, we record the complex coefficient of $\kappa(k_x, k_y)$ located at the spectral peak of $|\kappa(k_x, k_y)|$. This yields one complex-valued image $\xi(\tau_x, \tau_y)$, with the same dimensions as the specified wave field $h(x, y)$, which contains the amplitude and phase of the dominant wave at each location.

The location of the spectral peak in $|\kappa(k_x, k_y)|$ also gives us the wavenumbers k_x and k_y to which this peak coefficient corresponds. Hence, we can produce two further images $K_x(\tau_x, \tau_y)$ and $K_y(\tau_x, \tau_y)$ which contain the dominant wavenumbers at each location in the specified wave field to which the coefficients of $\xi(\tau_x, \tau_y)$ correspond.

Thus, in the three images $\xi(\tau_x, \tau_y)$, $K_x(\tau_x, \tau_y)$ and $K_y(\tau_x, \tau_y)$, we can measure the amplitudes, phases, wavelengths and propagation directions of the dominant wave features at each location in our specified wave field.

Figure 5.3(a) shows our specified wave field $h(x, y)$. The central locations of the eight synthetic waves with unit amplitudes and known wavelengths are numbered 1 to 8.

By taking the real part of the complex-valued image $\xi(\tau_x, \tau_y)$ containing the dominant coefficients, we can recover a "reconstruction" of the specified wave field, which is shown in Fig. 5.3(b). This is made possible by the approach described above in Sect. 5.4.1.

The 2DST identifies the different spectral regimes of the specified wave field very well, but the reconstructed wave amplitudes are reduced by comparison to their original values.

We suspect the main reason for the reduced amplitudes relates to the "spreading" of spectral power in the transform. Here, as is often the case for gravity waves in the real world, our simulated waves form small wave packets, where wave amplitude decreases around a central location. Such wave packets are usually represented in the spectral domain as some combination of wavenumber voices, in addition to the dominant wavenumber of each of the packets, in order to accurately describe their spatial properties. This means that the spectral power of a single, non-infinite wave packet can be spread across multiple wavenumber voices. Spectral leakage can further contribute to this effect.

The Gaussian window in the 2DST is equal to one at its central location, but immediately falls away with increasing radius. This means that any spectral power contained in adjacent wavenumber voices, which is required to fully reconstruct the wave, is reduced.

When the inverse DFT is computed, the recovered wave amplitude at this location is thus often diminished.

A further reason for the diminished amplitude recovery in Fig. 5.3(b) is due to wave undersampling. This undersampling effect is worse for longer wavelengths, since fewer wave cycles are present in the same-sized region of the image. The wave undersampling limitations of the S-transform are well-understood in one dimension (Wright, 2010; Wright et al., 2015b).

Figure 5.3(c) shows the absolute magnitude of the complex-valued image $|\xi(\tau_x, \tau_y)|$, which corresponds to the full underlying amplitude of the dominant wave at each location. This output is useful for defining regions of the specified wave field that do or do not contain clear and obvious wave features (McDonald, 2012).

The horizontal wavelength $\lambda_H(\tau_x, \tau_y) = (K_x(\tau_x, \tau_y)^2 + K_y(\tau_x, \tau_y)^2)^{-1/2}$ of the dominant wave at each location is shown in Fig. 5.3(d). Again, the different regimes of each wave in the specified wave field are clearly distinguished.

The direction of wave propagation $\theta(\tau_x, \tau_y)$, measured anticlockwise from the x-axis, is found as $\tan^{-1}\frac{K_x}{K_y}$, and plotted as Fig. 5.3(e). Note that $\theta(\tau_x, \tau_y)$ is subject to a $\pm\pi$ radian ambiguity, which is reconciled with *a priori* information in our AIRS analysis in Sect. 5.6.3.

To assess the effectiveness of our spectral analysis of the specified wave field, we compare the known wavelengths and propagation angles of the synthetic waves in the test image with the 2DST-measured wavelengths and propagation angles in Fig. 5.3(f). For each wave numbered 1 to 8, blue dots show the input wavelength λ_{IN} against measured wavelength λ_{OUT}, indicating that the 2DST measures the horizontal wavelengths and propagation angles in the test image very well. Generally, shorter wavelengths are well-resolved but longer wavelengths are slightly underestimated. This may be due to the coarse spectral resolution of DFT-based methods for waves with wavelengths that are a large fraction of the image size, since such waves can be more susceptible to spectral leakage problems.

5.5 An alternative spectral window

The use of the Gaussian window in the S-transform has some convenient mathematical advantages; it is analytically simple and has a definite integral over an infinite range. However, when it is used for 2-D S-transform analysis an unfortunate side effect of the Gaussian window is the poor recovery of wave amplitude, discussed in the previous section. Although a Gaussian is traditionally used, any suitable apodizing function may be used, so long as its spatial integral is equal to unity (Stockwell, 2007). For example, Pinnegar and Mansinha (2003) used an asymmetric hyperbolic time-domain window for enhanced measurement of the onset times of one-dimensional time series components.

In this section, we introduce a new spectral windowing function for the 2DST. This new function takes the shape of an ellipse in the wavenumber domain, and a first-order Bessel function of the first kind $\mathbf{J}_1(z)$ function with a scaled $1/z$ envelope in the spatial domain (for definition of z see Eqn. 5.34 below). For this reason we refer to this window as the Elliptic-Bessel (E-B) window. We find that when AIRS measurements are analysed with the 2DST using this new Elliptic-Bessel window in place of the traditional Gaussian, the measurement of gravity wave amplitudes is greatly improved. Spectral resolution is also improved slightly, without adversely compromising spatial resolution.

5.5.1 The Elliptic-Bessel window

As discussed in Sect. 5.4.1, the spectral peaks in a DFT spectrum have a characteristic width, where the spectral power is spread in a broad peak around the central wavenumber. This spectral power is slightly reduced when a Gaussian window is applied, due to the immediate decrease in the Gaussian function around the central location. The effect can be mitigated, but not fully reconciled, by decreasing the scaling parameter c, which broadens the Gaussian window in the wavenumber domain. However, this decreases the width of the spatial window, which increases the effect of wave undersampling for low wavenumbers.

One solution to this problem is to use a window that is an ellipse in the wavenumber domain. Here we introduce an Elliptic-Bessel window W_{eb}, defined in the wavenumber domain as the ellipse

$$W_{eb}(\alpha_x, \alpha_y, k_x, k_y) = \begin{cases} 0 & \text{for} \quad \left(\frac{\alpha_x}{a}\right)^2 + \left(\frac{\alpha_y}{b}\right)^2 \geqslant 1 \\ 1 & \text{for} \quad \left(\frac{\alpha_x}{a}\right)^2 + \left(\frac{\alpha_y}{b}\right)^2 < 1 \end{cases} \tag{5.32}$$

where $a = |k_x|/2\pi c$ and $b = |k_y|/2\pi c$ are the widths in the α_x and α_y directions. We see that the semi-major and semi-minor axes of this "voice ellipse" scale with angular wavenumbers and are equal to the standard deviations of the equivalent voice Gaussian window in Eqn. 5.31.

A key feature of this new window is that, in the wavenumber domain, it does not immediately decrease with displacement from the central location, but rather has a scalable

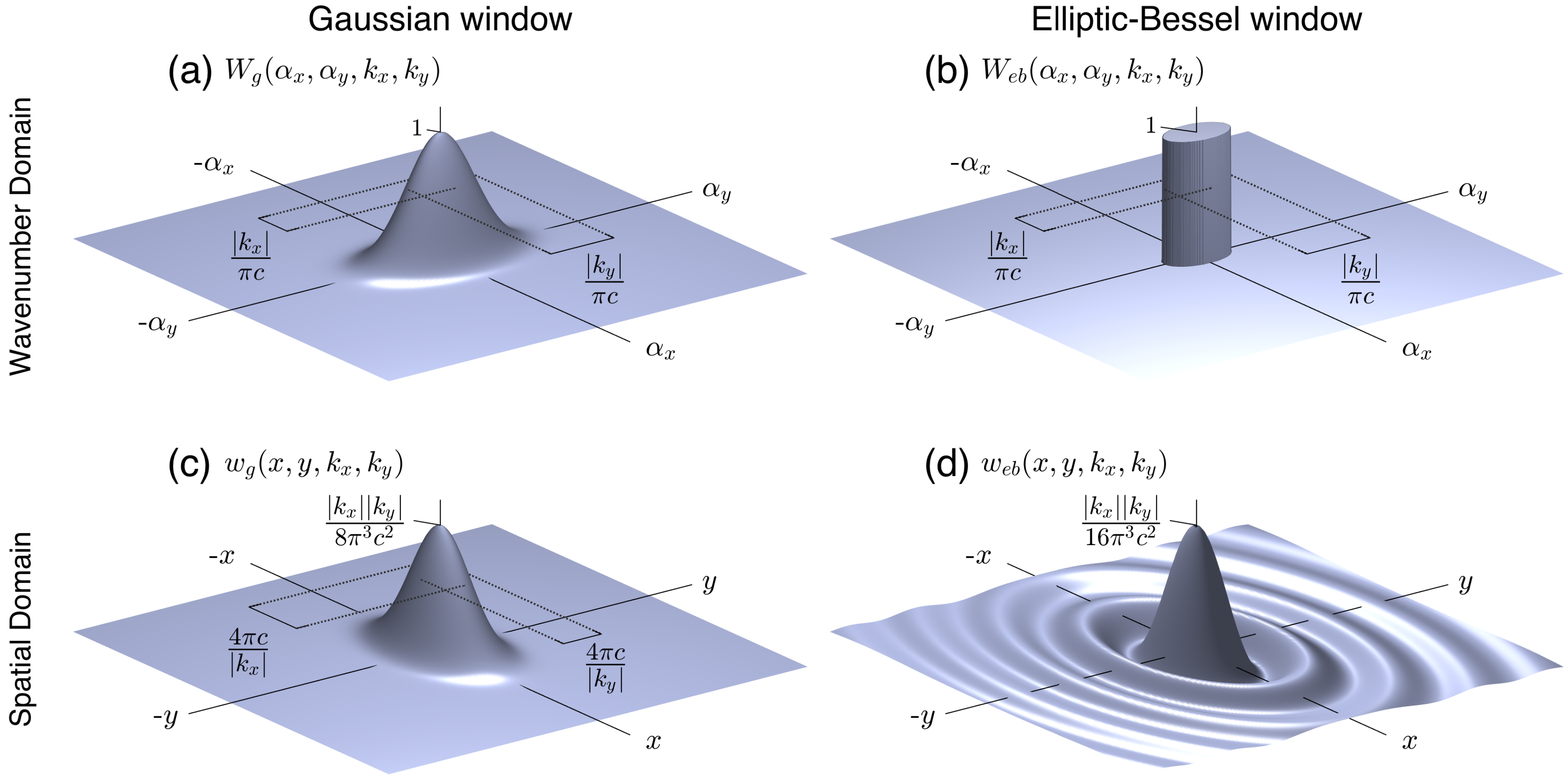

Figure 5.4: Illustrative surface plots of the wavenumber-domain (top row) and spatial-domain (bottom row) forms of the traditional Gaussian (a,c) and new Elliptic-Bessel (b,d) windowing functions used in the two-dimensional Stockwell transform for arbitrary wavenumbers k_x and k_y and scaling parameter c. The semi-major and semi-minor axes of the Ellitic-Bessel window in (b) are equal to the corresponding standard deviations of the Gaussian window in (a), where both windows have a central value equal to unity. In the spatial domain, the Elliptic-Bessel window in (d) has a central value equal to half that of the Gaussian window in (c), both of which have spatial integrals equal to unity. For details, see text.

elliptical region within which the function is equal to unity. Thus, the window captures a much greater extent of the targeted spectral peak at k_x and k_y, which can greatly improve wave amplitude recovery compared to the traditional Gaussian window. The width of W_{eb} can also be more carefully adjusted using the scaling parameter c. This window can then be used in place of the Gaussian windowing term in Eqn. 5.30.

One requirement for any apodizing window used in the Stockwell transform is that the spatial integral of the function must be equal to unity. This is so that the spatial integral of the Stockwell transform is equal to the Fourier transform $H(k_x, k_y)$ (Mansinha et al., 1997a), namely

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\tau_x, \tau_y, k_x, k_y)\, d\tau_x\, d\tau_y = H(k_x, k_y) \tag{5.33}$$

which has the useful result of making the 2DST fully invertible (Stockwell, 1999, 2007).

The normalisation term $|k_x||k_y|/8\pi^3c^2$ in the traditionally-used Gaussian window in Eqn. 5.29 ensures that the Gaussian window satisfies this requirement. To check that the Elliptic-Bessel function is admissible as an apodizing function, we must first find its spatial form, then check that its spatial integral is also equal to unity.

The Elliptic-Bessel window $W_{eb}(\alpha_x, \alpha_y, k_x, k_y)$ is easily defined in the wavenumber domain as an ellipse, but its spatial form, which we denote as $w_{eb}(x, y, k_x, k_y)$ is given as

$$w_{eb}(x, y, k_x, k_y) = \frac{|k_x||k_y|}{8\pi^3c^2}\,\frac{\mathbf{J}_1(z)}{z} \tag{5.34}$$

where $\mathbf{J}_1$ is the first-order Bessel function of the first kind (Abramowitz and Stegun, 1964) and

$$z = \frac{1}{2\pi c}\sqrt{k_x^2x^2 + k_y^2y^2}.$$

A short derivation of the function in Eqn. 5.34 is provided in Appendix 5.8.1. Fortunately, the spatial integral of Eqn. 5.34 is indeed equal to unity, proof of which is presented in Appendix 5.8.2. This confirms that the Elliptic-Bessel window is admissible as an apodizing function in the 2DST, and validates its use in this study and beyond.

To recap our notation in this study, we have described two windowing functions for the 2DST: the traditional Gaussian and the new Elliptic-Bessel windows, which we denote in the spatial domain as $w_g(x, y, k_x, k_y)$ and $w_{eb}(x, y, k_x, k_y)$ respectively, and in the wavenumber domain as $W_g(\alpha_x, \alpha_y, k_x, k_y)$ and $W_{eb}(\alpha_x, \alpha_y, k_x, k_y)$ respectively, where the W_g and W_{eb} are the Fourier transforms of w_g and w_{eb}.

Figure 5.4 shows three-dimensional surface plots of the spatial and wavenumber domain forms of the traditional 2-D Gaussian and new Elliptic-Bessel windows used in the 2DST here.

The surfaces in Figures 5.4(a) and 5.4(c) show the wavenumber-domain and spatial-domain forms of the Gaussian window, for arbitrary wavenumbers k_x and k_y. As discussed above, the maximum value of the Gaussian is equal to unity in the wavenumber domain, but equal to $|k_x||k_y|/8\pi^3c^2$ in the spatial domain such that its spatial integral is equal to unity. This is a requirement of any windowing function in the S-transform. The standard deviations of the w_g and W_g scale with wavenumbers k_x and k_y as described in Sect. 5.3.6 (Eqns. 5.29 and 5.31), providing the voice Gaussian.

Likewise, Figures 5.4(b) and 5.4(d) show the wavenumber-domain (W_{eb}) and spatial-domain (w_{eb}) forms of the Elliptic-Bessel window, for the same arbitrary wavenumbers k_x and k_y as used for the Gaussian windows in panels (a) and (c). The semi-major and semi-minor axes of the elliptic region in 5.4(b) are scaled with wavenumbers k_x and k_y, and are equal to the standard deviations of the equivalent Gaussian in 5.4(a), providing the voice ellipse.

The spatial-domain form of the Elliptic-Bessel window w_{eb}, described by a Bessel-shaped function within an envelope, is shown in Fig. 5.4(d) and described by Eqn. 5.34. The maximum value of w_{eb} is $|k_x||k_y|/16\pi^3c^2$, which ensures that its spatial integral is equal to unity. This is equal to half of the maximum value of the equivalent Gaussian in Fig. 5.4(c), since the terms involving z in Eqn. 5.34 tend to $1/2$ as $x \to 0$ and $y \to 0$. The width of the central region of w_{eb} is very slightly wider than the equivalent Gaussian, resulting in slightly coarser spatial resolution. However, the ability to "tune" the 2DST with the scaling parameter c ensures that this effect can be compensated by a reasonable trade-off.

In the next section we show that the use of the Elliptic-Bessel window in the 2DST, in place of the traditional Gaussian window, significantly improves wave amplitude recovery. This is very useful for our analysis of AIRS data in Sect. 5.6.

5.5.2 Invertibility

A very convenient aspect of the S-transform is its invertibility. Since we have shown here that both the traditional Gaussian and new Elliptic-Bessel windows have spatial integrals equal to unity, the 2DST can be completely inverted to recover the original 2-D image, whichever of these windows or real non-zero positive values of the scaling parameter c are used. Note that a traditional 1-D or 2-D CWT does not necessarily have this capability. The fact that we can achieve such flexibility in spatial-spectral resolutions by swapping windows or by adjusting c, yet still retain the capability of inversion, further highlights the strength of the 2DST as a tool for spatial-spectral analysis of geophysical data.

Unfortunately, to take full advantage of DFT algorithms and the inversion capability of the 2DST for AIRS data, we must compute the 2DST using all permitted wavenumber voices in both dimensions. This requires nearly 12 000 inverse DFT calculations for each AIRS granule using the traditional voice-by-voice implementation described here, the computational load of which could be quite impractical for large-scale studies. Interpolating AIRS measurements to a coarser resolution with fewer pixels could be one solution to reduce computational cost, but this will obviously undersample short horizontal wavelengths in the data. Faster methods for computing the S-transform have been developed (Brown et al., 2010) which may increase practicality in the future. Other steps, such as avoiding programming loops and ensuring that any 2-D objects to be transformed have dimensions that are powers of two, may also reduce relative computational expense.

5.5.3 The effect of window choices on AIRS granules

Figure 5.5 shows an AIRS granule over the Southern Andes measured on 24th May 2008, analysed using the 2DST with three different windowing approaches.

In Figs. 5.5(b) and 5.5(e), we use a Gaussian windowing function with the scaling parameter c set to one. This is the window usually used in 1DST implementations. We see that, as discussed above, this choice of window is only able to recover the very general, long-horizontal wavelength features of the granule, with poor spatial localisation and significantly reduced amplitude. This is due to a large proportion of the spectral response being lost by the windowing Gaussian when applied to two dimensions.

We can reduce the impact of this by decreasing the scaling parameter c, which broadens (narrows) the spectral (spatial) window. This provides improved amplitude recovery and improved spatial localisation at the expense of spectral localisation (Fritts et al., 1998). Since we only select a single dominant spectral peak for each location on the granule, this is acceptable for our purposes. The "reconstructed" perturbations and horizontal wavelengths (Figs. 5.5(c) and (f)) are now much more representative of the wave features in the granule.

One problem remains, however. By decreasing c, we narrow our spatial window. In regions where wave amplitudes are low, such as the bottom-left corner of Fig. 5.5(a), this narrow Gaussian window starts to undersample long wavelengths, such that only very short wavelengths are attributed to the region. The Elliptic-Bessel window used in Figs. 5.5(d) and (g) performs better at recovering the underlying larger-scale structure of the granule, without defaulting to the small-scale noisy variations. Amplitude recovery at all wavelengths is also improved over either of the Gaussian approaches.

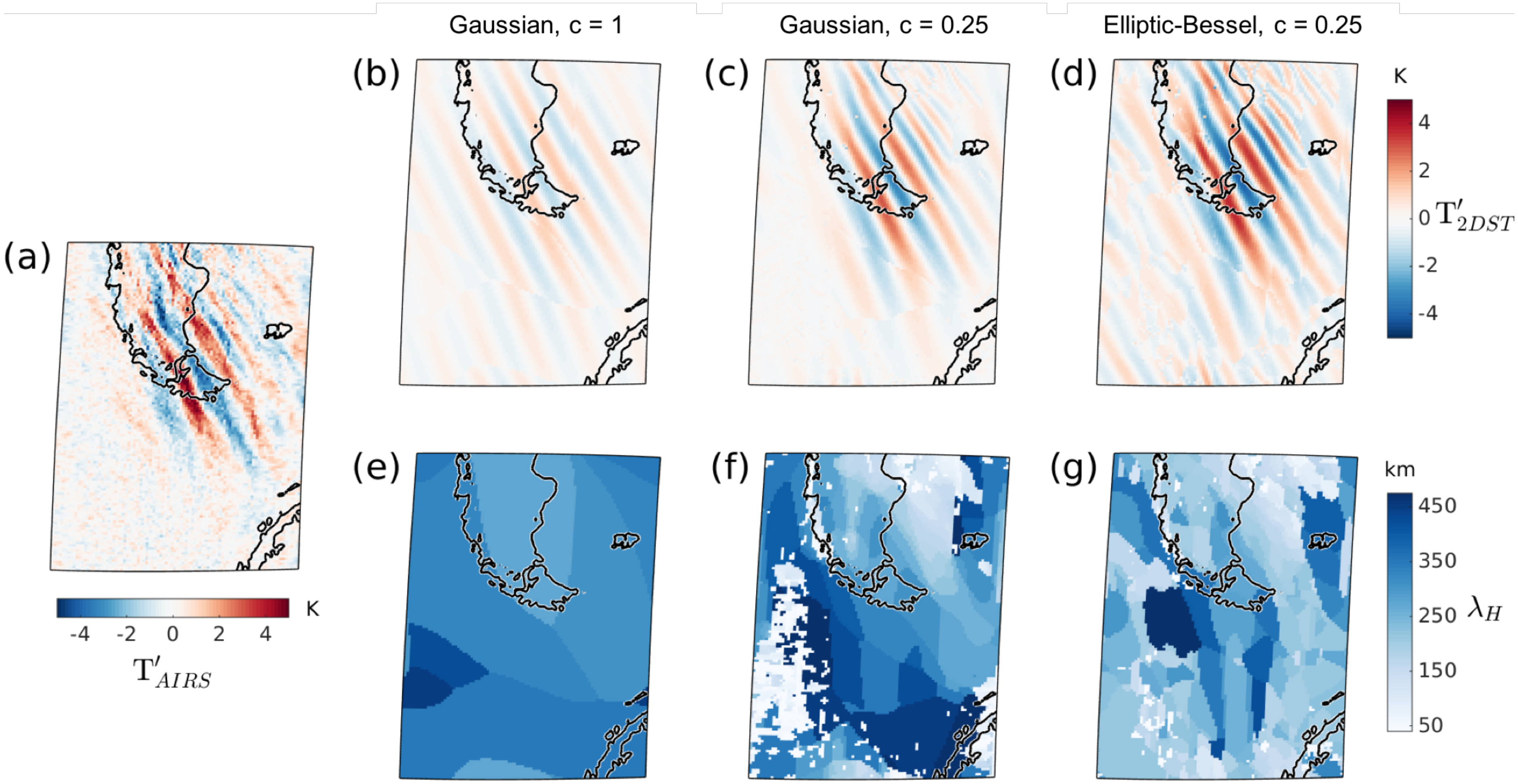

Figure 5.5: Orthographic projection of AIRS brightness temperature perturbations T'_{AIRS} (a) from a granule over the Southern Andes at 0530 UTC on 24th May 2008, with "reconstructed" temperature perturbations T'_{2DST} and horizontal wavelengths λ_H computed using the 2DST with three different windowing approaches: (b,e) a Gaussian window with scaling parameter $c = 1$, (c,f) a Gaussian window with $c = 0.25$, and (d,g) the Elliptic-Bessel window with $c = 0.25$.

In the general case, these low-amplitude, small-scale variations are unlikely to be due to gravity waves with vertical wavelengths visible to AIRS, so their recovery is something we try to avoid. Furthermore, such wavelengths are very close to or at the Nyquist limit for these data. Our confidence in their measurement is thus very low, yet the momentum fluxes they transport can dominate. We discuss this further in Sect. 5.6.4.

For the windowing functions considered, it is clear from Fig. 5.5 that the scaling parameter c has a significant effect in determining the spatial-spectral localisation capabilities of the 2DST. The Elliptic-Bessel windowing function, with a scaling parameter of $c = 0.25$, was selected for our AIRS analysis in the next section. This choice provided the best trade-off between spatial and spectral localisation of different wave regimes in AIRS measurements.

As discussed in Sect. 5.4.1, the 2-D images $\xi(\tau_x, \tau_y)$, $K_x(\tau_x, \tau_y)$ and $K_y(\tau_x, \tau_y)$ contain the dominant measured wave amplitudes and wavelengths at each location on the granule. These images are computed on a pixel-by-pixel basis, selecting a single monochromatic wave with the largest amplitude in the localised spectrum for each pixel.

As a result, the reconstructed images shown in Fig. 5.5(b-d), computed by taking the real part of the complex image $\xi(\tau_x, \tau_y)$, will never be perfect representations of the input data, but provide a “best guess” of the dominant features of the granule.

Since we have shown that the 2DST is fully invertible for both the Gaussian and Elliptic-Bessel windowing approaches (Sections 5.5 and Appendix 5.8), a complete reconstruction of the input image is of course producible by taking the “inverse” of the full four-dimensional 2DST object, but here we desire 2-D “maps” of wave properties, so a best guess method is used.

A possible quantitative metric to assess the first-order effectiveness of our 2DST analysis in Fig. 5.5 could be to compare the variance of the input image with the variances of each of the reconstructions. However, since the reconstructions are computed as a best guess method on a pixel-by-pixel basis, their total variance is not readily related to the total variance of the input image and thus may not be meaningful as a comparison. Furthermore, such use of the image variance would only be appropriate if the distribution of perturbations was unimodal and ideally Gaussian, which is not the case for an image of a sinusoidal wave. In practice however, we generally expect the variance of the reconstruction not to exceed the variance of the input image, since wave amplitudes computed on a pixel-by-pixel basis from a localised spectrum will usually be underestimated for the reasons given in Sect. 5.4.1.

It is not impossible that in some rare cases the total variance of the reconstruction could exceed the total variance of the input image, for example due to the spatial extent of a wave feature being slightly over-estimated. If the localised spectrum for one pixel is affected a larger amplitude wave feature in one of its neighbouring pixels, this can result in subtle artificial "borders" between different wave regimes in the reconstructions. This is not a limitation of the 2DST itself, but arises in the somewhat forced extraction of localised gravity-wave parameters contained in the 4-D Stockwell transform object $S(\tau_x, \tau_y, k_x, k_y)$ in order to produce the 2-D image. This effect should be carefully considered in future work to ensure wave properties are not over-represented.

5.6 AIRS gravity wave analysis using the 2DST

In this section, we use our 2DST-based method to perform gravity wave analysis on two-dimensional granules of AIRS radiance measurements, comparing our analysis to that of previous studies. We use the 2DST to measure gravity wave amplitudes, horizontal wavelengths, and directions of propagation. We then use ECMWF-derived wind speeds and the assumption of an orographic wave source to infer vertical wavelengths and make estimates of gravity wave momentum flux (the vertical flux of horizontal pseudomomentum) by closely following the method of Alexander et al. (2009).

5.6.1 AIRS granule selection and pre-processing

The first AIRS granule selected for our study is granule 32 of 6th September 2003, over South Georgia. The second granule is a 135-pixel swath over the intersection between granules 39 and 40 on 2nd August 2010, located over the Antarctic Peninsula and Drake Passage.

Alexander et al. (2009) and Hoffmann et al. (2014) performed an analysis of these AIRS granules over South Georgia and the Antarctic Peninsula respectively. Their studies measured wave amplitudes, horizontal wavelengths and wave propagation directions using a one-dimensional S-transform method, as described by Alexander and Barnet (2007). In their method, the one-dimensional S-transform is computed for each cross-track row. Then, covariance spectra are computed between pairs of adjacent cross-track rows to measure phase shifts in the along-track direction, from which along-track wavelengths can be inferred. To find the dominant wave features in a granule, these co-spectra are averaged together and up to five spectral peaks are found in this averaged spectrum. This approach can provide computationally fast, first-order gravity wave analysis of AIRS granules, and it has been used in numerous other studies (e.g. Alexander and Teitelbaum, 2011; Alexander and Grimsdell, 2013; Wright et al., 2015a). We thus select these two granules of AIRS measurements over South Georgia and the Antarctic Peninsula for ease of comparison between the 1-D S-transform method used by Alexander et al. (2009) and Hoffmann et al. (2014) and the new 2DST analysis method presented here.

One limitation of this method is the 1-D S-transform method the phase difference measurements required to recover along-track wavenumbers can introduce a strong cross-track bias in resolved features, since the S-transform is only computed in the cross-track direction. In addition, waves which occupy only small regions of the granule in the along-track direction may also be under-represented in the averaged co-spectrum. Furthermore, selecting no more than five dominant waves in the averaged co-spectrum implicitly limits the maximum number of available along-track wavenumber voices to no more than five for each location on the entire granule. The use of a two-dimensional Stockwell transform is a logical solution to each of these problems. With the increased convenience of computational power since the study of Alexander and Barnet (2007), the 2DST now represents a more practical alternative to the one-dimensional method.

Before implementing the 2DST, each granule of brightness temperature perturbations is interpolated onto a regularly-spaced grid with approximately 17.7 km and 20.3 km separating adjacent pixels in the along-track and cross-track directions respectively. In the centre of the AIRS swath, the resolution of this regularly-spaced grid closely matches the spatial resolution of AIRS, so very little if any information is lost. Towards the edge of the swath, this grid is finer than the spatial resolution than AIRS, but the grid points will not exactly match the location of the AIRS footprints. A useful graphic of typical AIRS footprints can be found in (Hoffmann et al., 2014, their Fig. 2).

As a result of using DFT algorithms, the maximum numbers of permitted wavenumber voices available in the along-track and cross-track directions are limited to $N_{AT} - 1$ and $N_{XT} - 1$, where N_{AT} and N_{XT} are the number of pixels in the along-track and cross-track directions (135 and 90 respectively). These wavenumber voices have corresponding wavelengths that are integer fractions of the total along-track and cross-track dimensions of the granule. Here, we compute the 2DST for wavelengths greater than around 40 km. This is just over twice the Nyquist-sampling distance between AIRS pixels after interpolation onto our regular grid. The zeroth frequencies are omitted. Increased along-track spectral resolution at low wavenumbers can be obtained by applying the 2DST to two or more adjacent granules, thus increasing the number of along-track pixels.

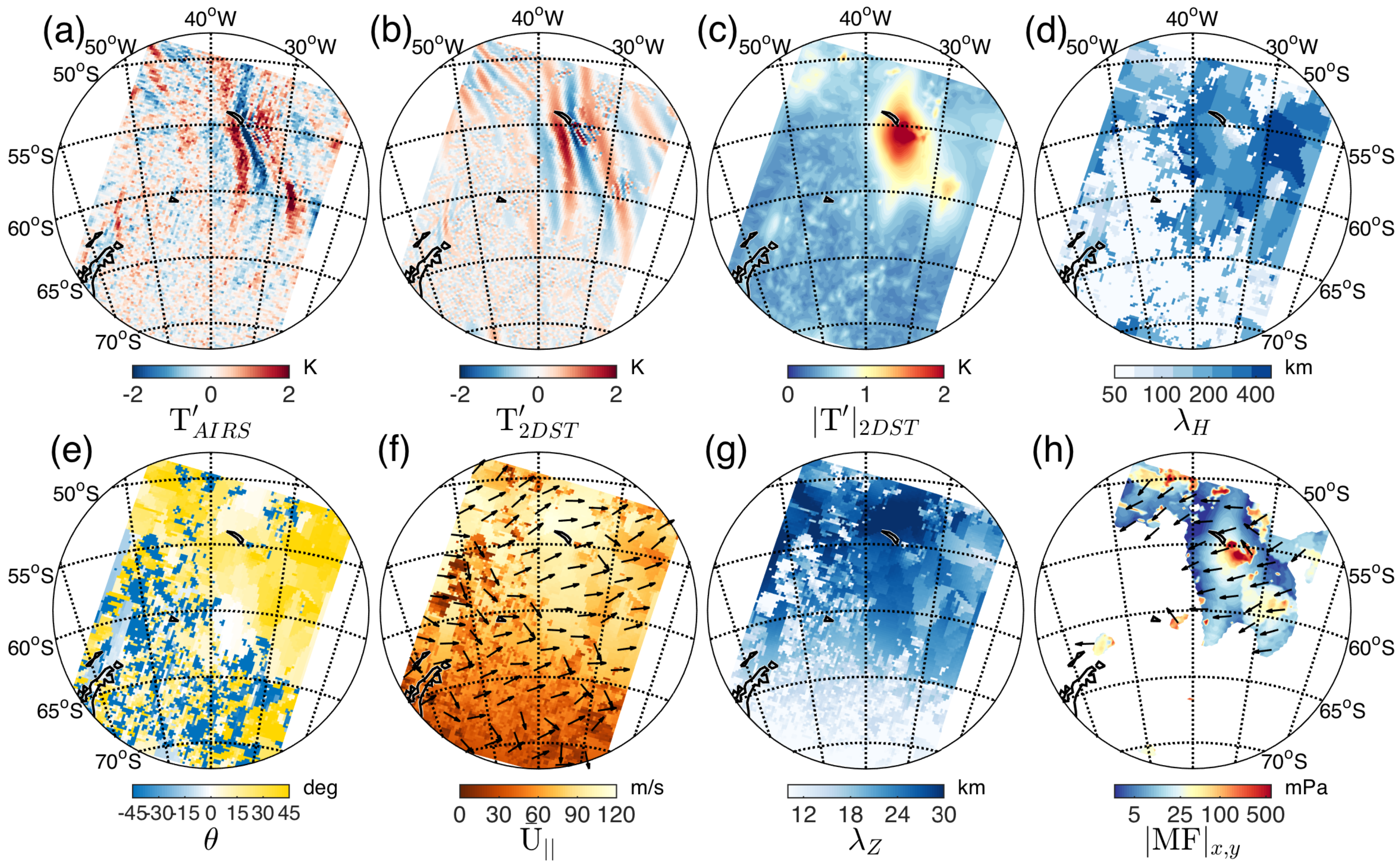

Figure 5.6: Orthographic projections of a granule of AIRS brightness temperature measurements T'_{AIRS} (a) over South Georgia at around 0300 UTC on 6th September 2003 and selected outputs (b–e) of our 2DST analysis using the Elliptic-Bessel window. This granule was also analysed by Alexander et al. (2009) (their Fig. 3) using a one-dimensional S-Transform method. The 2DST outputs shown here are reconstructed brightness temperature perturbations T'_{2DST} (b), underlying wave amplitudes $|T'|_{2DST}$ (c), horizontal wavelengths λ_H (d) and wave propagation directions θ (e) in degrees anticlockwise from east. Also shown are mean wind speed parallel to the horizontal wavenumber vectors $\bar{U}_{||}$ (f) from ECMWF operational analyses at $z \approx 40$ km, vertical wavelengths λ_Z (g) and the magnitude of the horizontal component of vertical momentum flux $|MF|_{x,y}$ (h). Black arrows in (f) and (h) show the horizontal direction of $\bar{U}_{||}$ and $|MF|_{x,y}$ respectively. For details, see text.

5.6.2 AIRS gravity wave properties measured by the 2DST

The results of our 2DST analysis of the selected AIRS granules over South Georgia and the Antarctic Peninsula are shown in Figs. 5.6 and 5.7 respectively.

In both Figures, panel (a) shows the brightness temperature perturbation measurements calculated as described in Sect. 5.2. Note that the colour scale is chosen so as to make wave perturbations clearer by eye, but at some locations it is saturated.

Clear wave-like perturbations are observed in both granules directly over and to the east of the mountain ranges. As in previous work, such clear wave-like perturbations are attributable to gravity waves with a high degree of certainty.

Reconstructed 2DST temperature perturbations T'_{2DST} are shown in panel (b). These are found by taking the real part of the complex 2DST object $\xi(\tau_x, \tau_y)$ as described in Sect. 5.4.1.

The image T'_{2DST} shows the dominant wave features in the granule reconstructed using only the pre-defined range of permitted wavenumber voices in the 2DST. Since we only consider the coefficients of the dominant wavenumber at each location, this reconstruction cannot be perfect, but it provides a visual inspection of how well the 2DST outputs represent the dominant wave characteristics of the granule. The T'_{2DST} image can be used to "fine-tune" the 2DST by changing the windowing function, by adjusting the scaling parameter c, or by redefining the range of frequency voices until the desired outcome is achieved. Such fine-tuning flexibility cannot be so easily achieved using the 1DST method.

Generally, the agreement between reconstructed wave features in Figure 5.6(b) and AIRS measurements in Figures 5.6(a) is very good, but some uncertainties remain. As discussed in Sect. 5.5.3, there is some discrepancy regarding the spatial extent of some wave features, such as a small positive wave crest located just south west of South Georgia which appears to be located slightly east, with an apparently slightly over-estimated amplitude, than is observed in the AIRS measurements. Conversely, at 63°S 28°W, a positive bow-shaped wave crest is observed in the AIRS measurements but is under-estimated in the reconstruction. As mentioned in Sect. 5.5.3, these small misrepresentations are not a limitation of the 2DST itself, but rather the forced extraction of gravity wave parameters from the 4-D S-transform object in order to create the 2-D reconstruction, where only one single wave feature with the largest localised spectral amplitude is assigned at each location. The overall agreement is still very good, but future work to improve the extraction of gravity wave parameters from the 4-D S-transform object may help to resolve some of these discrepancies.

Panels 5.6(c) and 5.7(c) show full underlying wave amplitudes $|T'|_{2DST}$ for each granule. This is found by taking the absolute magnitude of the complex 2DST object $\xi(\tau_x, \tau_y)$ as described in Sect. 5.4.1. This property provides us with a useful metric with which to define regions of the granule which do or do not contain wave-like perturbations, such that we can limit spurious detections (e.g. McDonald, 2012). In Figs. 5.6(h) and 5.7(c–h), we exclude regions of each granule where the underlying wave amplitude is more than one standard deviation below the mean underlying wave amplitude of the granule. In Figs. 5.6(c–g), we do not exclude such regions for discussion purposes, so as to provide an example of the data we would otherwise omit.

Panels 5.6(d) and 5.7(d) show absolute horizontal wavelengths $\lambda_H = 2\pi(k_{AT}^2 + k_{XT}^2)^{-1/2}$, where k_{AT} and k_{XT} are the along-track and cross-track angular wavenumbers respectively. We can see that these horizontal wavelengths clearly define different regimes of the dominant wave features of the granules, as in the test case in Sect. 5.4.1, though the AIRS data are more complex. In the South Georgia granule in Fig. 5.6(d), we see that the island lies within a wave field where long horizontal wavelengths are dominant around and to the east of the island over the ocean, with their wavenumber vectors aligned roughly parallel to the direction of the mean flow. This is characteristic of a wing-shaped mountain wave field (Alexander and Grimsdell, 2013), and is in good agreement with visual inspection of the granule itself.

In panels 5.6(e) and 5.7(e), we show the orientation of the horizontal wavenumber vector measured anticlockwise from east. θ is calculated by first projecting the along-track and cross-track wavenumber vectors k_x and k_y into their zonal and meridional components k and l using the azimuths of the along-track and cross-track directions at each location on the granule, then taking $\theta = \tan^{-1}(\frac{l}{k})$. Note that θ only describes the orientation and not the true horizontal direction of propagation of the wavenumber vectors, which retain a $\pm 180°$ ambiguity that we break below.

In the South Georgia granule (Fig. 5.6), we see that our 2DST measurements in the southern region of the granule are largely dominated by small-scale, low-amplitude, short horizontal wavelength features with random directions of propagation. Most of these features are likely to be due to noise and not attributable to coherent wave structures. By using a threshold amplitude, such regions are effectively removed, leaving well-defined regions with clear wavelike perturbations. The contribution of small-scale features that remain after this step is discussed further in Sect. 5.6.4.

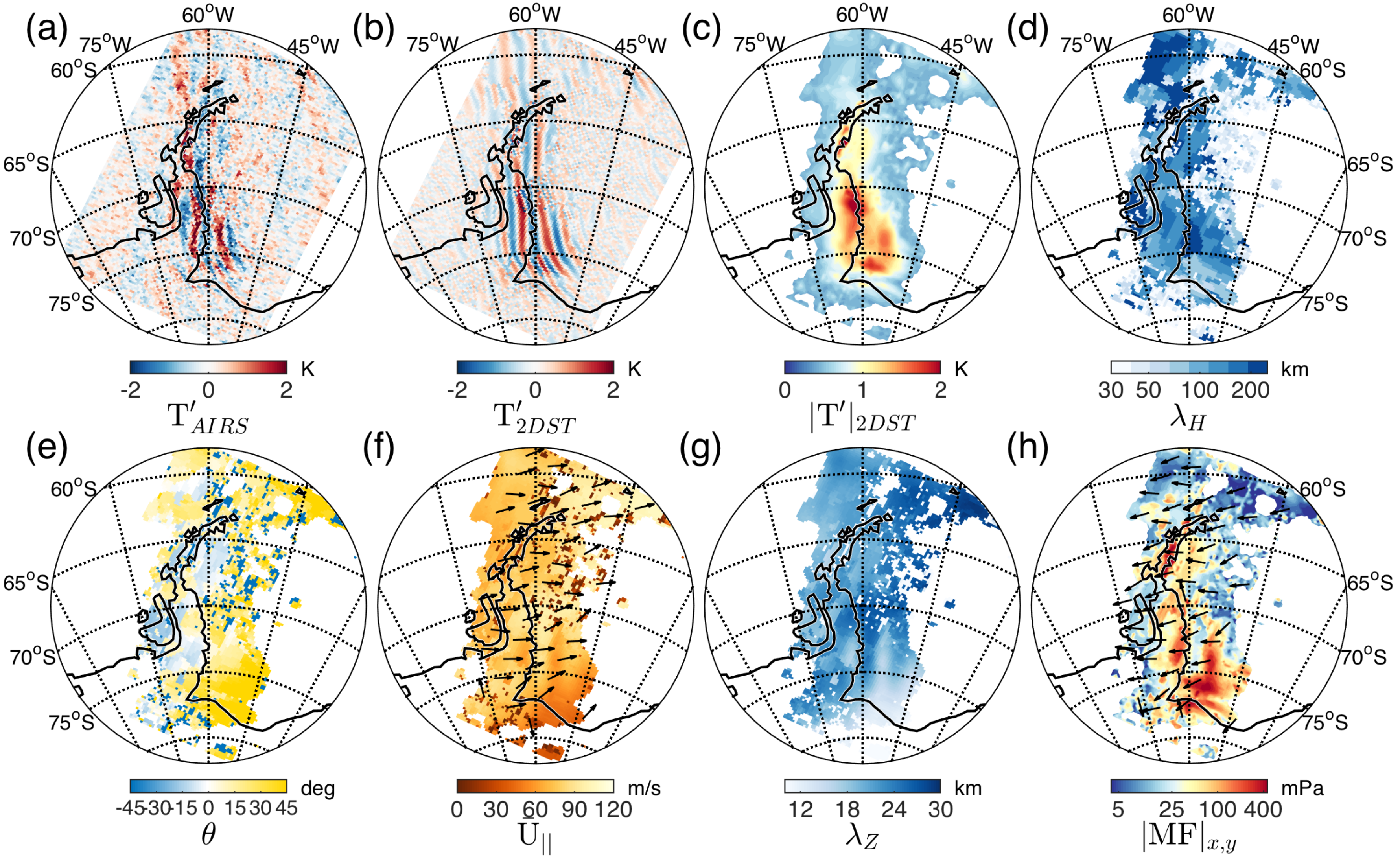

Figure 5.7: As Fig. 5.6, but for AIRS measurements over the Antarctic Peninsula around 0400 UTC on 2nd August 2010. These measurements were also analysed by Hoffmann et al. (2014, their Fig. 8) using a one-dimensional S-Transform method. Here, regions in panels (c–h) where the underlying wave amplitude $|T'|_{2DST}$ is less than one standard deviation below the mean are coloured white.

5.6.3 Momentum fluxes

Here we make estimates of gravity wave momentum flux for the dominant wave-like features measured by the 2DST in our selected granules, following the method of Alexander et al. (2009).

Ern et al. (2004) showed that the zonal and meridional components of gravity wave momentum flux MF_x and MF_y can be given by

$$(\mathrm{MF}_x, \mathrm{MF}_y) = \frac{\rho}{2}\left(\frac{g}{N}\right)^2\left(\frac{T'_a}{\bar{T}}\right)^2\left(\frac{k}{m}, \frac{l}{m}\right) \tag{5.35}$$

where ρ is density at a height of 40 km, g is the acceleration due to gravity, N is the buoyancy frequency, T'_a is the attenuation-scaled full underlying wave amplitude, $\bar{T}$ is the background temperature, and k, l and m are wavenumbers in the zonal, meridional and vertical directions respectively. Eqn. 5.35 assumes the mid-frequency approximation (Fritts and Alexander, 2003, their Sect. 2.1), which is valid for a large portion of the observable gravity wave spectrum. For these waves, $m^2 \ll (k^2 + l^2)$ and Coriolis effects are neglible, as is the case for nearly all waves measured here due to the size of the AIRS beam footprint (Hoffmann et al., 2014) and the vertical weighting function of the 667.77 cm^{-1} channel (Alexander and Barnet, 2007). When compared with momentum fluxes calculated without making the mid-frequency approximation, Ern et al. (2004) found discrepancies typically not exceeding around 10%.

We also have a ±180° ambiguity in direction of propagation, which we break by assuming the waves in our granules always propagate against the mean flow. One last variable in Eqn. 5.35 is vertical wavenumber $m = 2\pi/\lambda_Z$. Under the assumption that the waves are upwardly propagating mountain waves with ground-based phase velocity equal to zero, λ_Z is given as

$$\lambda_Z \approx \frac{2\pi \bar{U}_{||}}{N} \tag{5.36}$$

where $\bar{U}_{||}$ is the component of the mean wind speed parallel to the wave's horizontal wavenumber vector (Eckermann and Preusse, 1999). $\bar{U}_{||}$ is found by projecting the mean wind vector $\bar{U}$ in the direction of the wave propagation angle θ shown in Figs. 5.6(e) and 5.7(e).

Figures 5.6(f) and 5.7(f) show $\bar{U}_{||}$ coincident with each granule at an altitude of 40 km from ECMWF operational analyses, projected onto each granule's regular grid. Orange contours and black arrows show the magnitude and direction of $\bar{U}_{||}$ respectively. Vertical wavelength λ_Z is shown in Figs. 5.6(g) and 5.7(g).

Towards the south-eastern corner of both granules, mean wind speeds become quite weak. As they fall below around 40 ms^{-1}, vertical wavelengths start to drop below the vertical resolution limit of the AIRS channel. The wave field may continue into this region, but the vertical wavelengths may be too short to be resolved such that wave amplitudes are attenuated to below the ambient noise level. This is particularly clear in the Antarctic Peninsula granule, where detectable wave fronts abruptly terminate just as $\bar{U}_{||}$ begins to fall below 30 ms^{-1}.

No further wave amplitude attenuation corrections were applied to regions where vertical wavelengths are inferred to be below 12 km, since this is beyond the resolution limit of the weighting function of the 667.77 $^{-1}$cm channel. It should be noted however that although theoretically valid, the attenuation correction applied here can in some cases increase wave amplitudes by perhaps an order of magnitude or more, despite being significantly susceptible to errors in wind speed, wave direction or both. This correction should be applied with caution, and the effects of its inclusion are discussed below and in Sect. 5.6.4.

In the South Georgia granule, peak momentum flux values of more than 500 mPa are associated with a small region of large amplitude and short horizontal wavelength wave features, located toward the south-eastern tip of the island. In the Antarctic Peninsula granule, momentum fluxes of a few hundred milliPascals are generally co-located with the clearly visible wave structures in the raw brightness temperature perturbations just downwind of the peninsula.

The key strength of the results presented here is the much-improved spatial-spectral localisation and resolution capabilities provided by full two-dimensional treatment of the AIRS data. Confidence in the accuracy of subsequent measured quantities in our 2DST-based analysis is thus greatly improved over previous 1-D S-transform-based methods. Understandably, the former is more computationally intensive than the latter, and this should be considered if datasets are large or computational resources are limited.

Although the magnitude, direction and distribution of momentum fluxes in both granules are broadly in line with previous AIRS gravity wave studies in the region (e.g Alexander and Teitelbaum, 2007; Alexander et al., 2009; Alexander and Teitelbaum, 2011), momentum flux magnitudes are much higher than those observed in studies using limb-sounder datasets (e.g. Ern et al., 2004; Alexander et al., 2008; Ern et al., 2011; Geller et al., 2013; Hindley et al., 2015; Wright et al., 2015a), which generally do not exceed a few tens of mPa in this region. While we would not expect the magnitude of these fluxes to be the same as those observed by limb-sounders (1. nadir-sounding instruments are generally more sensitive to waves with longer vertical wavelengths and hence higher momentum fluxes due to the deep vertical weighting function (Alexander and Barnet, 2007); 2. momentum flux estimates from limb-sounders are typically lower-bound estimates due to the projection of horizontal wavelengths (Ern et al., 2004); and 3. wavelength-dependent wave amplitude attentuation corrections are not generally applied to limb-sounder results) we note that our results and those of other AIRS gravity wave studies in this region, which use a correction factor for wave amplitude attenuation based upon a vertical wavelength estimation, are substantially higher. This effect and its implications are discussed in the next section.

5.6.4 Small-scale perturbations and the attenuation correction

In the implementation of any spectral image processing, it is important to strike a balance between accurate measurement of the desired properties and the spurious interpretation of noise. One of the advantages of AIRS measurements is the high horizontal resolution of the data. With the close exception of the Infrared Atmospheric Sounding Interferometer (IASI, e.g. Clerbaux et al., 2009), currently no other spaceborne instrument can measure stratospheric gravity waves with comparable horizontal resolution to AIRS. Therefore, accurate measurement of resolved waves with horizontal wavelengths close to the AIRS resolution limit are of great importance. Such short horizontal wavelength waves, if reliably resolved, will generally carry higher momentum fluxes via Eqn. 5.35. However, as we approach the resolution limits of AIRS measurements and our spectral methods, our confidence in the accuracy of our measurement of such waves decreases.

In Fig. 5.6, our 2DST analysis resolves a very small region of short horizontal wavelengths over the south-eastern tip of South Georgia. Wave perturbations in this small region are just a few pixels across, but their uncorrected brightness temperature perturbations are large, with peaks of order 5 to 6 K. If these perturbations were located almost anywhere else on the granule we would likely attribute them to retrieval noise.

However, gravity waves with very short horizontal wavelengths, tightly packed in a region immediately downwind of a mountainous island, are in good agreement with mountain wave theory. Examples of such waves can be found in, for example, the modelling studies of Shutts and Vosper (2011) and Alexander and Teitelbaum (2011). If these waves are indeed real, their accurate measurement is of great importance.

An added complication is introduced as a result of the mountain wave assumption used. Since these waves are only a few pixels across, their directions of propagation are difficult to define, introducing a random element. The component of the mean wind parallel to the horizontal wavenumber vector can thus be very low, which decreases the vertical

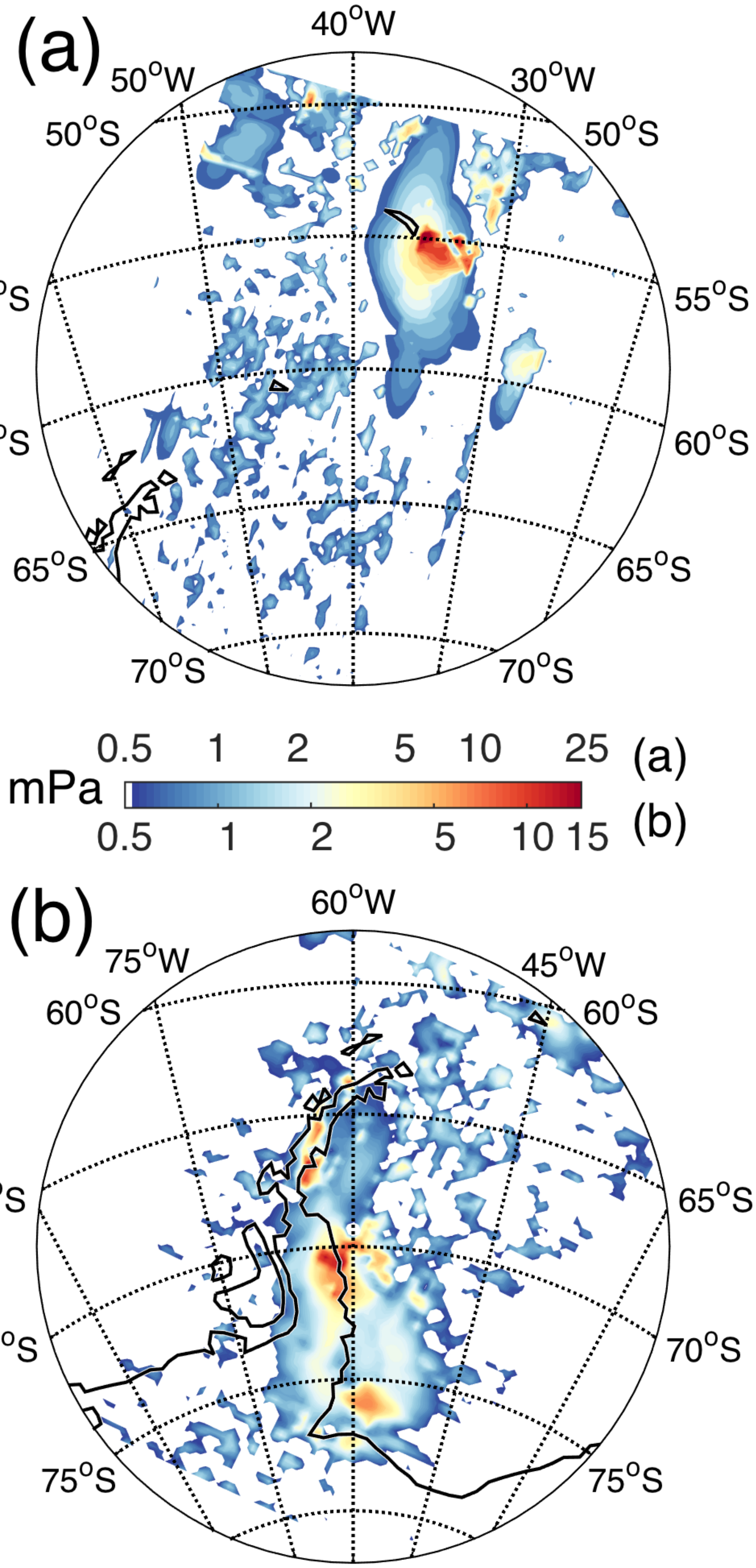

Figure 5.8: As Figures 5.6h and 5.7h but for estimates of the magnitude of gravity wave momentum flux over South Georgia (a) and the Antarctic Peninsula (b) calculated without correcting for attenuation of wave amplitude as discussed in Sect. 5.6. Momentum fluxes less than 0.5 mPa are coloured white.

wavelength estimate, which in turn increases the attenuation correction applied to the observed temperature perturbations. This attenuation correction can increase temperature perturbations by 400% or more which, since momentum flux is proportional to the square of wave amplitude, can increase our estimate of momentum flux to extremely large values.

Thus, by applying a correction factor for wave amplitude attentuation, very small-scale low-confidence perturbations can yield extremely high momentum fluxes which can dominate the momentum budget of the entire granule if the mountain wave assumption is used, correctly or otherwise. This is evident in the South Georgia granule studied by Alexander et al. (2009) and shown in our Fig. 5.6, where peak momentum fluxes are localised over only a few large-amplitude pixels just over the south-eastern tip of the island. In Fig. 5.6(h), the flux peaks at almost 1000 mPa for one pixel in this small region, though the colour scale is saturated. Whether these fluxes are real or not, they can nevertheless be extremely large and should be approached with caution. Further investigation into their authenticity is of immediate importance if GCMs are to incorporate their fluxes into parametrization schemes.

Figures 5.8(a) and 5.8(b) show gravity-wave momentum fluxes over the South Georgia and Antarctic Peninsula granules as shown in Figures 5.6 and 5.7, only with fluxes calculated without scaling the temperature perturbations for attenuation as described in Sect. 5.6. The highest momentum flux values in Figures 5.8(a) and 5.8(b) are around 70 and 30 mPa respectively, located within small regions in the wave field where the colour scale is saturated. As is expected, these values are much lower than those calculated using the attenuation-corrected wave amplitudes.

If these extremely high fluxes correspond to real waves, then their measurement is of crucial importance. However, if such perturbations are simply instrument noise and their fluxes are spurious, then the biases and errors introduced by their inclusion in broader studies could be very large.

Without *a priori* knowledge of the wave environment, which is most readily gained by visual inspection of the AIRS measurements, it would be unwise to include the fluxes from these small-scale perturbations in any automated analysis. Indeed, Wright et al. (2015b) suggest that when AIRS granules are pre-smoothed with a boxcar of width 3 pixels, resolved momentum fluxes calculated using the method of Alexander et al. (2009) can be reduced by a order of magnitude. This suggests that these small-scale, high-momentum flux features are reasonably common and can impact larger-scale momentum flux estimates, either realistically or spuriously.

Further work investigating this problem is encouraged. Pre-smoothing granules so as to exclude these perturbations (Wright et al., 2015b, 2016) is one solution. Excluding even more wavenumber voices corresponding to short horizontal wavelengths from our 2DST analysis is another solution. However, in cases such as the South Georgia granule presented in Fig. 5.6, where such perturbations may well be physical, this exclusion can reduce peak momentum fluxes by orders of magnitude, introducing a systematic low-bias and thus further uncertainty.

5.6.5 Implications of small-scale waves for general circulation models

In Sect. 5.6.4 we touched upon the potential significance of the detection of small-horizontal scale gravity waves and the high momentum fluxes they can transport into the middle atmosphere. Despite their small scale, these waves can have exhibit significant drag effects on mean flow, potentially an order of magnitude more per unit area than the larger-scale inertia-gravity waves that we studied using COSMIC GPS-RO data in Chapter 4.

The impact and representation of these small-horizontal scale waves in General Circulation Models (GCMs) is currently an active area of research. As mentioned in Sect. 5.1, it has been hypothesised (McLandress et al., 2012) that unresolved gravity-wave drag, possibly from small islands like South Georgia (Alexander and Grimsdell, 2013), may be a large contributing factor in the "cold-pole" bias (Butchart et al., 2011), a problem suffered by nearly all GCMs and a serious impediment to model progress.

As a consequence of the Courant-Friedrichs-Lewy (CFL) condition, in order for a propagating, time-varying atmospheric wave to be correctly simulated in a GCM, the horizontal and vertical grid sizes of the model (dH and dZ respectively) divided by the model time step dt must be smaller than the horizontal and vertical phase speeds of the wave (c_H and c_Z respectively) as

$$c_H \leqslant \frac{dH}{dt} \quad \text{and} \quad c_Z \leqslant \frac{dZ}{dt} \tag{5.37}$$

In other words, wave motion arising in a generic numerical model by the evaluation of a physical state in each grid cell has the fundamental limitation that a given wave perturbation must either remain in a grid cell after one model time step or move to an adjacent cell - and no further. So, if the spatial characteristics of the wave imply that its phase speed would transport its perturbations further than one grid cell over one time step, such propagation is not permitted as grid cells can only dynamically interact with adjacent cells, and the wave disappears.

In a high-resolution modelling study, Sato et al. (2012) used the T213L256 "Kanto" GCM developed by Watanabe et al. (2008) to present large-scale gravity wave activity in the southern hemisphere at a then-unprecedented global vertical resolution. The Kanto GCM has a horizontal triangularly-truncated spectral resolution of T213, which corresponds to latitude and longitude grid spacing dH of 0.5625°, giving a characteristic horizontal grid spacing of around $\sim$60 km at the equator, with improved zonal horizontal resolution towards the poles. In the studies by Watanabe et al. and Sato et al., the model was run with a time step dt of 30 seconds.

In the horizontal, in order for the CFL condition in Eqn. 5.37 to be satisfied in the Kanto model, gravity wave horizontal phase speeds cannot exceed $dH/dt \approx 1000\,\text{ms}^{-1}$, which is of course supersonic and not possible for atmospheric waves, so this condition is always satisfied in the horizontal.

In the vertical however, a key aspect of the Kanto model is that it contains 256 vertical layers, with a spacing dZ of $\sim$ 300 m throughout the middle atmosphere. While the closely-spaced vertical layers allow for excellent vertical resolution, the 30 second time step means that upwardly-propagating gravity waves with vertical phase speeds greater than $dZ/dt \approx 10\,\text{ms}^{-1}$ will break the CFL condition. If we consider the kinds of waves we observe in Fig. 5.6, with modest horizontal and vertical wavelength estimates of $\lambda_H \sim 100$ km and at least $\lambda_Z \sim 12$ km, we can use relations in Chapter 2 to estimate a vertical phase speeds of at least $\sim$10 ms^{-1} for such waves. Any shorter horizontal or longer vertical wavelengths (which we almost certainly observe in Fig. 5.6(g)), will break the CFL condition and cannot be accurately represented in the vertical in the Kanto model.

At time of writing, the Integrated Forecast System (IFS) used operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF) runs on a T1279 (horizontal grid spacing of $\sim$30 km or better) with time step of 7 seconds. The IFS is run on 137 vertical layers, with separation dZ of $\sim$1200 m at an altitude of around $\sim$40 km. While this is coarser than the Kanto model in the vertical, the resolution is more than sufficient to resolve the kinds of waves we consider in Fig. 5.6. For gravity waves to break the CFL condition in a model with a gridscale like the IFS, vertical phase speeds in the middle atmosphere must exceed $dZ/dt \approx 1200/7 \approx 170\,\text{ms}^{-1}$, a limit which is not encountered by any of the middle-atmosphere waves we have considered so far.

This means that, while the constantly improving horizontal and vertical resolutions of GCMs provide advantageous resolution of smaller and smaller-scale processes, model time steps must also be kept as short as possible in order to ensure waves with relatively short (long) horizontal (vertical) scales satisfy the CFL condition and can be accurately modelled. As increased computational power allows for higher and higher model resolutions, the importance of this consideration is set to develop in the future.

5.7 Summary and Conclusions

In this study, we have applied the two-dimensional Stockwell transform (2DST) to granules of AIRS measurements, extracting gravity wave amplitudes, wavelengths and directions of propagation. Our 2DST method builds upon the work of Alexander and Barnet (2007), who used the one-dimensional Stockwell transform for the same purpose. Their method can introduce a strong cross-track bias problem, which we solve by using a full two-dimensional Stockwell transform.

We first define our 2DST implementation and test it on a specified wave field containing synthetic waves with known amplitudes, wavelengths and directions of propagation. We find that the 2DST provides very good spatial representation of the dominant spectral components of the specified wave field, accurately measuring wavelengths and orientations of all the synthetic waves.

Due to the spread of spectral power in the spectral domain and wave undersampling in the spatial domain, we find that localised wave amplitudes as measured by the 2DST are reduced by more than a factor of two when the typical Gaussian windowing function is used in the Stockwell transform. We compensate for this by decreasing the scaling parameter c and by replacing the Gaussian window with a new alternative Elliptic-Bessel window, which we test on a granule of AIRS measurements over the southern Andes. We find that this new spectral window provides a better balance between spatial-spectral localisation and the accurate measurement of wave amplitudes. Wave amplitude recovery is thus improved to around 80% to 90% of input values. We also demonstrate that this new window is indeed valid for use in the 2DST by showing its spatial integral is equal to unity.

Next, we measure gravity wave amplitudes, horizontal wavelengths and directions of propagation in two granules of AIRS measurements over South Georgia and the Drake Passage/Antarctic Peninsula region. Our 2DST method significantly improves two-dimensional representation of the dominant spectral features of the granules over previous 1DST methods. These spectral features are directly measured in both dimensions simultaneously for each location of the granule, without the introduction of potential biases caused by the use of averaged co-spectra. This is a clear advantage over previous methods.

Another key advantage of our 2DST method is the ability to visually inspect the quality of our spectral analysis. By taking the real parts of the dominant localised spectral coefficients at each location, a reconstruction of the granule can be created. This can be used to fine-tune the adjustable parameters, and provide a useful sanity check on the performance of the 2DST. Future work may involve comparing this output to the original data via a variance argument or similar, such that we can obtain a quantitative measure of the quality of the 2DST analysis for quality control purposes in larger scale studies.

To conclude, our new 2DST-based gravity wave analysis method for AIRS data makes significant improvements over current methods in several key areas, and we would advocate its use in future work.

5.8 Admissibility of the Elliptic-Bessel window in the Stockwell Transform

In Sect. 5.5 we introduced the Elliptic-Bessel window as new apodizing function for the 2-D Stockwell transform (2DST). One requirement for any apodizing function for use in the Stockwell transform is that its spatial sum must be equal to unity. If this condition is satisfied, the spatial sum of the 2DST is equal to the 2-D Fourier transform, making the 2DST fully invertible.

In this appendix we demonstrate that the Elliptic-Bessel window is admissible as an apodizing function in the S-transform. To do this, we must first find the spatial analogue of the wavenumber-domain ellipse we defined in Eqn. 5.32. We must then take the spatial integral of this function to demonstrate that it is equal to unity.

5.8.1 The Elliptic-Bessel window in the spatial domain

The Elliptic-Bessel window is defined in the wavenumber (α_x, α_y) domain as

$$W_{eb}(\alpha_x, \alpha_y, k_x, k_y) = \left\{ \begin{array}{llcl} 0 & \text{for} \quad \left(\frac{\alpha_x}{a}\right)^2 + \left(\frac{\alpha_y}{b}\right)^2 & \geqslant & 1 \\ 1 & \text{for} \quad \left(\frac{\alpha_x}{a}\right)^2 + \left(\frac{\alpha_y}{b}\right)^2 & < & 1 \end{array} \right. \tag{5.38}$$

where $a = |k_x|/2\pi c$ and $b = |k_y|/2\pi c$ are the half-widths of the ellipse in the α_x and α_y directions (see Fig. 5.4(b)). The spatial-domain form of the Elliptic-Bessel window, denoted here by $w_{eb}(x, y, k_x, k_y)$, is found by taking the inverse 2-D Fourier transform of Eqn. 5.38 as

$$\begin{aligned} w_{eb}(x, y, k_x, k_y) &= \mathscr{F}_x^{-1} \mathscr{F}_y^{-1} \left[W_{eb}(\alpha_x, \alpha_y, k_x, k_y) \right] \\ &= \frac{1}{4\pi^2} \int_{\infty}^{\infty} \int_{\infty}^{\infty} W_{eb}(\alpha_x, \alpha_y, k_x, k_y) e^{i(\alpha_x x + \alpha_y y)} d\alpha_x \, d\alpha_y \end{aligned} \tag{5.39}$$

Since $W_{eb}(\alpha_x, \alpha_y, k_x, k_y) = 1$ within the ellipse and zero everywhere else, and has double symmetry, we can change the limits of integration to be the boundaries of the ellipse, expressing the total integral as a sum of four equal quadrants

$$w_{eb}(x, y, k_x, k_y) = \frac{4}{4\pi^2} \int_0^{\sqrt{b^2 - \frac{\alpha_x^2 b^2}{a^2}}} \int_0^{a} e^{i(\alpha_x x + \alpha_y y)} d\alpha_x \, d\alpha_y \tag{5.40}$$

We then recognise that the exponential term in the transform above can be replaced with sine and cosine functions as

$$\begin{aligned} e^{i(\alpha_x x + \alpha_y y)} &= (\cos(\alpha_x x) + i \sin(\alpha_x x)) (\cos(\alpha_y y) + i \sin(\alpha_y y)) \\ &= \cos(\alpha_x x) \cos(\alpha_y y) + i \sin(\alpha_x x) \cos(\alpha_y y) \\ &\qquad + i \sin(\alpha_y y) \cos(\alpha_x x) - \sin(\alpha_x x) \sin(\alpha_y y) \end{aligned} \tag{5.41}$$

We can omit the last three terms in 5.41 since, due to the symmetry of the sine function around $(0, 0)$, each term will eventually sum to zero. We can then rewrite Eqn. 5.40 as

$$w_{eb}(x, y, k_x, k_y) = \frac{4}{4\pi^2} \int_0^{\sqrt{b^2 - \frac{\alpha_x^2 b^2}{a^2}}} \int_0^{a} \cos(\alpha_x x) \cos(\alpha_y y) d\alpha_x \, d\alpha_y \tag{5.42}$$

This integral can be further simplified if we switch to polar coordinates using the substitutions $\alpha_x = ar\cos(\phi)$ and $\alpha_y = br\sin(\phi)$ after which the expression in 5.42 becomes

$$w_{eb}(x, y, k_x, k_y) = \frac{ab}{4\pi^2} \int_0^{2\pi} \int_0^{1} \cos(arx\cos(\phi)) \cos(bry\sin(\phi)) \, r \, dr \, d\phi \tag{5.43}$$

Next we substitute $\mathcal{A} = arx$ and $\mathcal{B} = bry$ and, using multiple angle formulae, rewrite 5.43 as

$$\begin{aligned} w_{eb}(x, y, k_x, k_y) &= \frac{ab}{4\pi^2} \int_0^{1} r \int_0^{2\pi} \cos(\mathcal{A}\cos(\phi)) \cos(\mathcal{B}\sin(\phi)) \quad d\phi \, dr \\ &= \frac{ab}{8\pi^2} \int_0^{1} r \int_0^{2\pi} [\cos(\mathcal{A}\cos\phi + \mathcal{B}\sin\phi) \\ &\qquad + \cos(\mathcal{A}\cos\phi - \mathcal{B}\sin\phi)] \quad d\phi \, dr \\ &= \frac{ab}{8\pi^2} \int_0^{1} r \int_0^{2\pi} [\cos(\sqrt{\mathcal{A}^2 + \mathcal{B}^2}\cos(\phi - \Lambda) \\ &\qquad + \cos(\sqrt{\mathcal{A}^2 + \mathcal{B}^2}\cos(\phi + \Lambda)] \quad d\phi \, dr \end{aligned} \tag{5.44}$$

where $\Lambda = \tan^{-1}(\mathcal{B}/\mathcal{A})$. Here, Λ is simply an arbitrary phase due to the periodicity of the cosine function when integrated over 0 to 2π, so the integrals of both terms in the square brackets in Eqn. 5.44 will be equal. Hence we can simply add these terms such that we have

$$w_{eb}(x,y,k_x,k_y) = \frac{ab}{8\pi^2}\int_0^1 r \int_0^{2\pi} 2 \quad \cos\left(\sqrt{\mathcal{A}^2+\mathcal{B}^2}\cos\phi\right) \quad d\phi\, dr \tag{5.45}$$

Next we recall the integral definition of the zeroth-order Bessel function of the first kind $\mathbf{J}_0(x)$ (Abramowitz and Stegun, 1964) given as

$$\mathbf{J}_0(x) = \frac{1}{2\pi}\int_0^{2\pi} \cos\left(x\cos\phi\right) \quad d\phi \tag{5.46}$$

and substitute into Eqn. 5.45 and reintroduce our substitutions of $\mathcal{A} = arx$ and $\mathcal{B} = bry$ to give

$$w_{eb}(x,y,k_x,k_y) = \frac{ab}{2\pi}\int_0^1 \mathbf{J}_0\left(r\sqrt{a^2x^2+b^2y^2}\right) r\, dr \tag{5.47}$$

We now use a new substitution that $\xi = r\sqrt{a^2x^2+b^2y^2}$ and rewrite Eqn. 5.47 as

$$w_{eb}(x,y,k_x,k_y) = \frac{ab}{2\pi}\int_0^{\sqrt{a^2x^2+b^2y^2}} \frac{\xi \quad \mathbf{J}_0(\xi)}{a^2x^2+b^2y^2} \quad d\xi \tag{5.48}$$

Next we use the standard result (e.g. Abramowitz and Stegun, 1964) that

$$\int_{x_1}^{x_2} x\,\mathbf{J}_0(x) \quad dx \quad = \quad x\,\mathbf{J}_1(x)\,\Big|_{x_1}^{x_2} \tag{5.49}$$

to rewrite Eqn. 5.48 as

$$\begin{aligned} w_{eb}(x,y,k_x,k_y) &= \frac{ab}{2\pi}\frac{\xi \quad \mathbf{J}_1(\xi)}{(a^2x^2+b^2y^2)}\,\Bigg|_0^{\sqrt{a^2x^2+b^2y^2}} \\ &= \frac{ab}{2\pi}\frac{\mathbf{J}_1\left(\sqrt{a^2x^2+b^2y^2}\right)}{\sqrt{a^2x^2+b^2y^2}} \end{aligned} \tag{5.50}$$

Finally, recalling that $a = |k_x|/2\pi c$ and $b = |k_y|/2\pi c$ are the half-widths of the original ellipse in 5.38, we now can write the analytical expression for the spatial form of the Elliptic-Bessel window as

$$w_{eb}(x,y,k_x,k_y) \quad = \quad \frac{|k_x||k_y|}{8\pi^3c^2}\,\frac{\mathbf{J}_1(z)}{z} \tag{5.51}$$

where

$$z \quad = \quad \frac{1}{2\pi c}\sqrt{k_x^2x^2+k_y^2y^2}$$

This spatial-domain form of the Elliptic-Bessel window in Eqn. 5.51 is plotted in Fig. 5.4(d).

Equation 5.51 describes a $\mathbf{J}_1(z)$ function within a scaled $1/z$ envelope. Because of this, the terms involving z in 5.51 converge to $1/2$ as $x \to 0$ and $y \to 0$, such that the central region of the function has peak value of $|k_x||k_y| \,/\, 16\pi^3c^2$, as shown in Fig. 5.4(d). Interestingly, this value is equal to half the peak value of the equivalent Gaussian window shown in Fig. 5.4(c). The central peak of the Elliptic-Bessel window is also, for each frequency voice, slightly broader than that of the equivalent voice Gaussian.

5.8.2 Spatial integral of the Elliptic-Bessel window

Now that we have found an analytical expression for the spatial-domain form of the Elliptic-Bessel window (Eqn. 5.51), we can proceed to check that it is admissible as an apodizing function in the 2-D Stockwell transform; namely that its spatial sum is equal to unity (e.g. Pinnegar and Mansinha, 2003). The spatial sum of Eqn. 5.51, denoted here by $\mathbb{I}$, can be written as

$$\mathbb{I} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{|k_x||k_y|}{8\pi^3c^2}\,\frac{\mathbf{J}_1(z)}{z}\quad dx\,dy \tag{5.52}$$

This integral can be simplified if we reintroduce our substitutions $a = |k_x| \,/\, 2\pi c$ and $b = |k_y| \,/\, 2\pi c$ and switch to polar coordinates, using the substitutions $x = \frac{\Lambda\cos(\varphi)}{a}$ and $y = \frac{\Lambda\sin(\varphi)}{b}$ to give

$$\begin{aligned}\mathbb{I} &= \frac{ab}{2\pi}\int_0^{2\pi}\int_0^{\infty} \frac{\mathbf{J}_1(\Lambda)}{\Lambda}\frac{\Lambda}{ab}\quad d\Lambda\,d\varphi \\ &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} \mathbf{J}_1(\Lambda)\quad d\Lambda\,d\varphi \\ &= \frac{2\pi}{2\pi}\int_0^{\infty} \mathbf{J}_1(\Lambda)\quad d\Lambda \end{aligned}\tag{5.53}$$

Using the standard result (e.g. Abramowitz and Stegun, 1964) that

$$\int_{x_1}^{x_2} \mathbf{J}_1(x)\quad dx \quad = \quad -\mathbf{J}_0(x)\,\Big|_{x_1}^{x_2} \tag{5.54}$$

we see that Eqn. 5.53 becomes

$$\begin{aligned}\mathbb{I} &= -\mathbf{J}_0(\Lambda)\,\Big|_0^{\infty} \\ &= (0) - (-1) \\ &= 1\end{aligned}\tag{5.55}$$

as required. This result confirms that the spatial sum of $w_{eb}(x, y, k_x, k_y)$ is indeed equal to unity, thus the Elliptic-Bessel window is admissible as an apodizing window for the 2-D Stockwell transform.

5.8.3 Admissibility of other windows

In this appendix so far, we have found a useful analytical expression for spatial form of the Elliptic-Bessel window presented in this study. We have then shown that its spatial integral is equal to unity and it is thus admissible as an apodizing function in the 2-D Stockwell Transform. In other cases, a quick test may be performed on candidate S-transform windowing functions to check if this spatial integral is unity.

If we take the spatial integral $\mathbb{I}$ of the spatial-domain form of a candidate windowing function $w(x, y, k_x, k_y)$, namely

$$\mathbb{I} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x, y, k_x, k_y)\, dx\, dy \tag{5.56}$$

and introduce the factor $e^{-i(k_x x + k_y y)}$, noting that when $k_x = k_y = 0$ this factor is equal to unity, then $\mathbb{I}$ can be written as

$$\begin{aligned} \mathbb{I} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x, y, k_x, k_y) e^{-i(k_x x + k_y y)}\, dx\, dy \,\Big|_{k_x = k_y = 0} && (5.57) \\ &= W(\alpha_x, \alpha_y, k_x, k_y) \,\Big|_{k_x = k_y = 0} && (5.58) \end{aligned}$$

where $W(\alpha_x, \alpha_y, k_x, k_y)$ is the wavenumber domain form of the candidate window and the notation $|_{k_x = k_y = 0}$ denotes that the function is evaluated at $k_x = k_y = 0$. This means that if the value of $W(\alpha_x, \alpha_y, k_x, k_y)$ evaluated at $k_x = 0$ and $k_y = 0$ is equal to unity, then its spatial integral will also be equal to unity. If it is not, then the candidate window is not admissible for use the Stockwell transform. Figures 5.4(a) and 5.4(b) show that both the Gaussian window and the Elliptic-Bessel window are equal to unity at $k_x = k_y = 0$, and thus satisfy this requirement. This short test may be helpful in the design of proposed alternative S-transform windowing functions in the future.

Chapter 6

Summary

6.1 Summary

In this thesis, we have used COSMIC GPS-RO measurements to investigate the nature of gravity waves in and around the southern stratospheric polar vortex and gravity wave hot spot of the southern Andes and Antarctic Peninsula. We have developed innovative new one-dimensional and two-dimensional analysis methods for the detection and measurement of gravity waves in both one-dimensional COSMIC and two-dimensional AIRS data.

We began in Chapter 1 with a brief introduction to the atmosphere of the earth. We discuss the general characteristics of the middle atmosphere and the importance of gravity waves in general circulation.

In Chapter 2 we gave an in-depth description of atmospheric gravity waves. We derived their mathematical dispersion, discussed orographic and non-orographic wave sources, wave breaking mechanisms and gave a mathematical treatment of wave-borne energy and momentum transport.

In Chapter 3 we discussed the instruments and datasets used for the studies in this thesis. We gave a brief overview of satellite remote sensing of the earth's atmosphere, with a discussion of the main types of satellite viewing geometry and the subset of the gravity wave spectrum detectable by each type. We described the GPS radio occultation technique, the COSMIC satellite constellation, the retrieval of atmospheric temperature measurements and the vertical and horizontal resolution limits of the technique. We then gave an overview of the AIRS and HIRDLS instruments, data from which is used in later sections.

In Chapter 4 we conducted an investigation into the nature of gravity waves in and around the southern stratospheric gravity wave hot spot and southern polar vortex using COSMIC GPS-RO data. We found evidence of meridional propagation of waves into the polar vortex from sources far to the north. We then developed a quantitative wavelet-based identification methodology for COSMIC data and used it to investigate gravity wave intermittency over the mountains of the southern Andes and Antarctic Peninsula and around the polar vortex. We then made estimates of gravity wave horizontal wavelengths and momentum flux over the hot spot using pairs of closely-spaced and closely-timed COSMIC profiles.

In Chapter 5 we developed a new two-dimensional gravity wave analysis methodology for AIRS measurements based on a two-dimensional Stockwell transform. We tested our method on a specified wave field and showed that we could accurately measure horizontal wavelengths and directions of propagation. We showed that, by using an alternative elliptical window in our two-dimensional Stockwell transform, we could dramatically improve the recovery of wave amplitude. We then applied our method to AIRS measurements, measuring gravity wave horizontal wavelengths, amplitudes, directions of propagation and momentum fluxes over South Georgia and the Antarctic Peninsula, where we found that our new method made significant improvements over current methods.

In Chapter 7 we discuss parallel and future studies that have been or could be conducted by the author in collaboration with colleagues. We describe parts I and II of the ambitious multi-instrument study of gravity wave vertical and horizontal wavelengths, energies and momentum fluxes over Tierra del Fuego. We describe the innovative three-dimensional gravity wave measurement technique made possible by the combination of AIRS and Aura-MLS measurements. We also discuss the potential extension of this method to a global scale. Finally, we propose a short study with a view to constraining the precise line-of-sight horizontal resolution of GPS-RO using 2DST-analysed AIRS measurements and co-located COSMIC profiles.

Chapter 7

Future Studies

In this section we discuss proposals for some future studies building upon or extending the work presented in the preceding chapters of this thesis.

7.1 Multi-instrument gravity-wave measurements over Tierra del Fuego and the Drake Passage, Part 2: Gravity wave momentum fluxes, acceleration, intermittency and spectral subsetting

C. J. Wright, M. Ern, D. C. Fritts, D. Janches, N. P. Hindley, N. J. Mitchell, and co-authors.
In preparation for *Atmospheric Measurement Techniques*, expected submission 2016.

This author's contribution to the second volume of this two-part study (first part: Wright et al. (2015a)) is the measurement of stratospheric gravity wave momentum fluxes from COSMIC GPS-RO and the inter-dataset comparison of fluxes from each of the seven instruments. A preliminary abstract for this study is provided below:

Abstract

Gravity waves in the terrestrial atmosphere are of vital geophysical importance, underlying a vast range of effects and processes. They can be studied with a broad range of observational datasets, including limb- and nadir-sounding satellites, radars, and radiosondes. However, these instruments, even when observing the same volume of atmosphere, often measure very different spectral regions of the gravity wave spectrum. Understanding

these differences is vital to advancing our understanding of the true underlying gravity wave spectrum, and thus accurately modelling the effects of these waves on the atmospheric flow. Here, we investigate a range of such instruments over a region centred on the SAAMER meteor radar at 53.8°S , 67.8°W . This region is at the centre of the largest known source of gravity waves in the world, the Andes/Antarctic Peninsula orographic hotspot. In the first part of this study (Wright et al. (2015a), hereafter referred to as "Part I"), we investigated gravity wave potential energies (GWPE) and vertical wavelengths (λ_Z) over this region using data from AIRS, COSMIC, HIRDLS, MLS-Aura, SAAMER, SABER and radiosondes. Here, we extend that analysis to study gravity wave momentum fluxes and horizontal wavelengths. We then characterise the short-timescale variability, or intermittency, both GWPE and GWMF in our datasets. Finally, we investigate the effects on inter-dataset comparability of simple spectral subsetting, in order to partition to first order the differences between our datasets in terms of observational filter effects and other instrumental differences.

7.2 Global measurements of three-dimensional gravity-wave properties from combined AIRS and MLS observations

Wright et al. (2016) demonstrated a successful proof-of-concept combination of measurements from AIRS and the Microwave Limb-Sounder (MLS) instrument for the simultaneous measurement of three-dimensional gravity wave properties from space. Since MLS (on board the Aura satellite) follows AIRS (on board the Aqua satellite) around the same orbital track, there have been more than a million co-locations of AIRS granules and MLS measurements during the satellites' lifetimes globally, with more co-located measurements being made every day (see Fig. 7.1). This presents an unprecedented opportunity to study stratospheric gravity wave momentum fluxes and, crucially, the direction of these fluxes on a global scale. As yet, no observational study has successfully measured the three-dimensional properties and fluxes globally without very large assumptions and biases. Using the two-dimensional Stockwell transform method developed in Chapter 5, this technique could enable the unprecedented study of gravity-wave dynamics on a global scale, making possible the study of not only their sources but also their driving effects on planetary-scale phenomena such as the Quasi-Biennial Oscillation (QBO) and/or the Brewer-Dobson circulation.

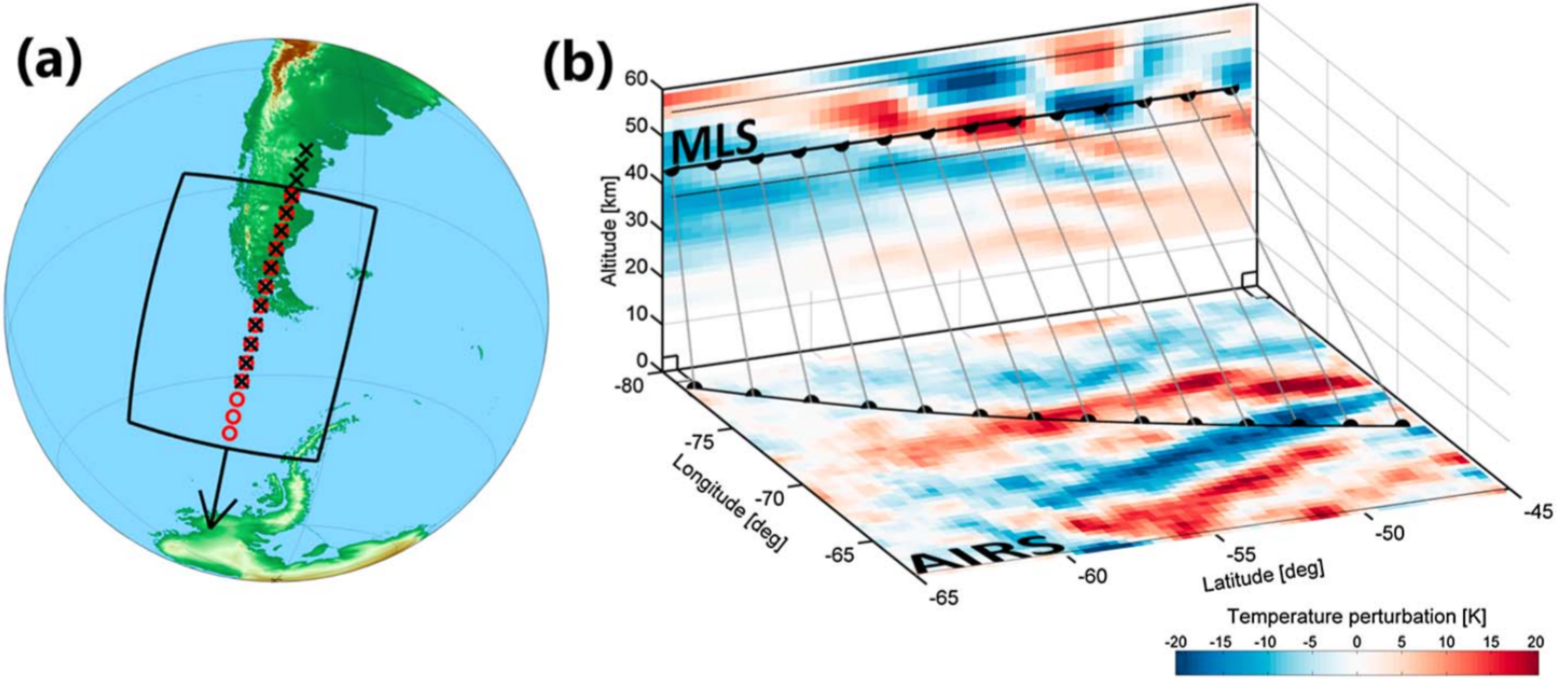

Figure 7.1: (a) Overlap between AIRS and MLS measurements for granule AIRS.2008.05.06.057. Black arrow indicates direction of travel, black solid lines indicate the outline of the AIRS granule, and black crosses (red circles) indicate temporally coincident (time-shifted spatially coincident) MLS profiles. (b) Temperature perturbations T' from the same pairing, with MLS shown on the vertical plane and AIRS on the horizontal plane. Semicircles indicate coincident points (MLS profile locations at 42 km altitude) in both planes. From Wright et al. (2016).

7.3 Global estimates of gravity-wave momentum flux from closely-spaced pairs of COSMIC GPS-RO profiles

In chapter 4 we used closely-spaced pairs of COSMIC temperature profiles to make estimates of stratospheric gravity wave momentum fluxes over the southern Andes/Drake Passage/Antarctic Peninsula region. Our results in section 4.4 demonstrate that, given sufficient sampling density, COSMIC GPS-RO can produce physically reasonable estimates of momentum flux over the southern Andes and Antarctic Peninsula that are consistent with results from CRISTA, HIRDLS and Vorcore (Ern et al., 2004; Alexander et al., 2008; Hertzog et al., 2008).

The study in chapter 4, although useful, is limited to a regional case study. The logical step forward is thus to expand the method in section 4.4 to a global scale. Here we describe a preliminary approach and some initial findings of a study applying the COSMIC closely-spaced profile-pairs method on a global scale, with a discussion of the expected benefits of the upcoming COSMIC-2 satellite constellation and aspects that could be further explored in future work.

7.3.1 Removal of planetary wave artefacts

One obstacle in the application of our method to a global scale is the contribution of temperature perturbations due to planetary-scale waves in COSMIC measurements.

Planetary waves are large-scale oscillations that are coherent around a circle of latitude, with horizontal wavelengths comparable to, or integer fractions of, the Earths circumference. At high latitudes, as were considered in Chapter 4, planetary wave amplitudes are generally low, with relatively long vertical wavelengths that are usually filtered out by the processing in Sect. 4.2.

In the tropics however, one particular species of planetary-scale wave – Kelvin waves – can have vertical wavelengths as short as 4–5 km, well within the detectable range of gravity wave vertical wavelengths in COSMIC data (see Fig. 7.2(a)). In an isolated COSMIC profile, these Kelvin wave perturbations are indistinguishable from gravity waves.

However, Kelvin waves are zonal wavenumber 1, i.e. they have horizontal wavelengths equal to the circumference of the earth. Although gravity waves around the equator can reach very low intrinsic frequencies, and hence extremely long wavelengths, as the Coriolis parameter $f \to 0$ in Equation 2.23, they are generally not so large as comparable

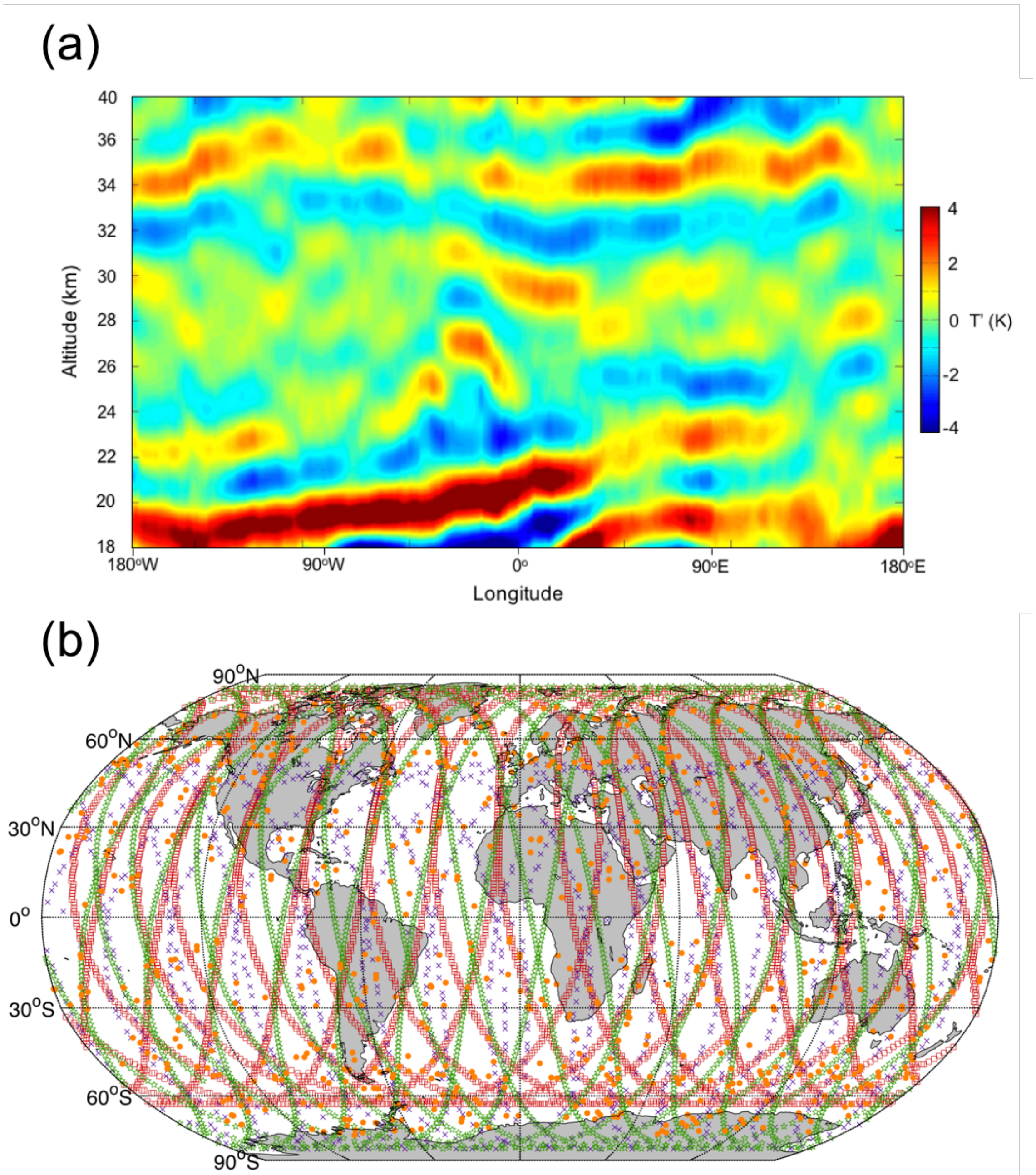

Figure 7.2: 3-day mean of stratospheric COSMIC temperature perturbations T' due to Kelvin waves and gravity waves (a) in a latitude band centred on the equator between ±5° N/S, and daily sampling denisty of COSMIC (orange dots), HIRDLS (red squares), SABER (purple crosses, southern yaw phase) and Aura-MLS (green diamonds).

to be the circumference of the earth. This difference allows us to separate Kelvin wave perturbations from gravity waves in the tropics. But, since the sampling density of the COSMIC constellation is quite poor in the tropics, extra measurements are required to do this.

In a preliminary study, we combine measurements from COSMIC and HIRDLS with those two other limb-sounding instruments: the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) instrument aboard the NASA TIMED satellite and the Microwave Limb-Sounder (MLS) aboard the Aura satellite. The daily sampling densities of these four instruments, shown in Fig. 7.2(b), allow for a comprehensive daily atmospheric "background" state in the stratosphere to be measured, which includes Kelvin waves.

For each day, each dataset is first interpolated to 1 km altitude resolution, then all four are binned and averaged into $5°$ latitude bands. At each height level, each band is zonally high-pass filtered with an approximate cut-off around 45–60° longitude. Thus, large planetary-scale perturbations due to Kelvin waves are heavily suppressed, such that the only perturbations with horizontal scales $\lesssim$ 45–60° longitude ($\lambda_H \lesssim$ 5000–6000 km at the equator) remain, which are much more likely to be attributable to gravity waves. Subtracting these perturbations from the combined measurements acts as a low-pass filter, giving an underlying background atmospheric temperature state, and a measure of local Kelvin wave perturbations on a daily basis with a $5°$ latitude resolution.

Then, for each closely-spaced profile pair of COSMIC profiles measured during the deployment phase of the COSMIC constellation, this Kelvin-wave background state can be subtracted, before the profile is vertically high-pass filtered as described in Sect. 4.2. With these steps, Kelvin wave perturbations are essentially removed (or very heavily suppressed) in COSMIC profiles, and we can now proceed to follow the steps described in Sect. 4.4 to make estimates of gravity wave momentum fluxes on a global scale.

7.3.2 Global COSMIC gravity wave momentum fluxes

Some preliminary results of our global momentum flux analysis are shown in Fig. 7.3. The general form of these results in promising, particularly during July-September 2006 in Fig. 7.3(a). We see the familiar hot spots of the southern Andes and Antarctic Peninsula, and an associated leeward region of increased flux. We also see three regions of enhanced gravity wave momentum flux associated with deep tropical convection over central America, northern and central Africa and south-east Asia. These features are fairly well-known and our results are in good agreement with observations by, for exam-

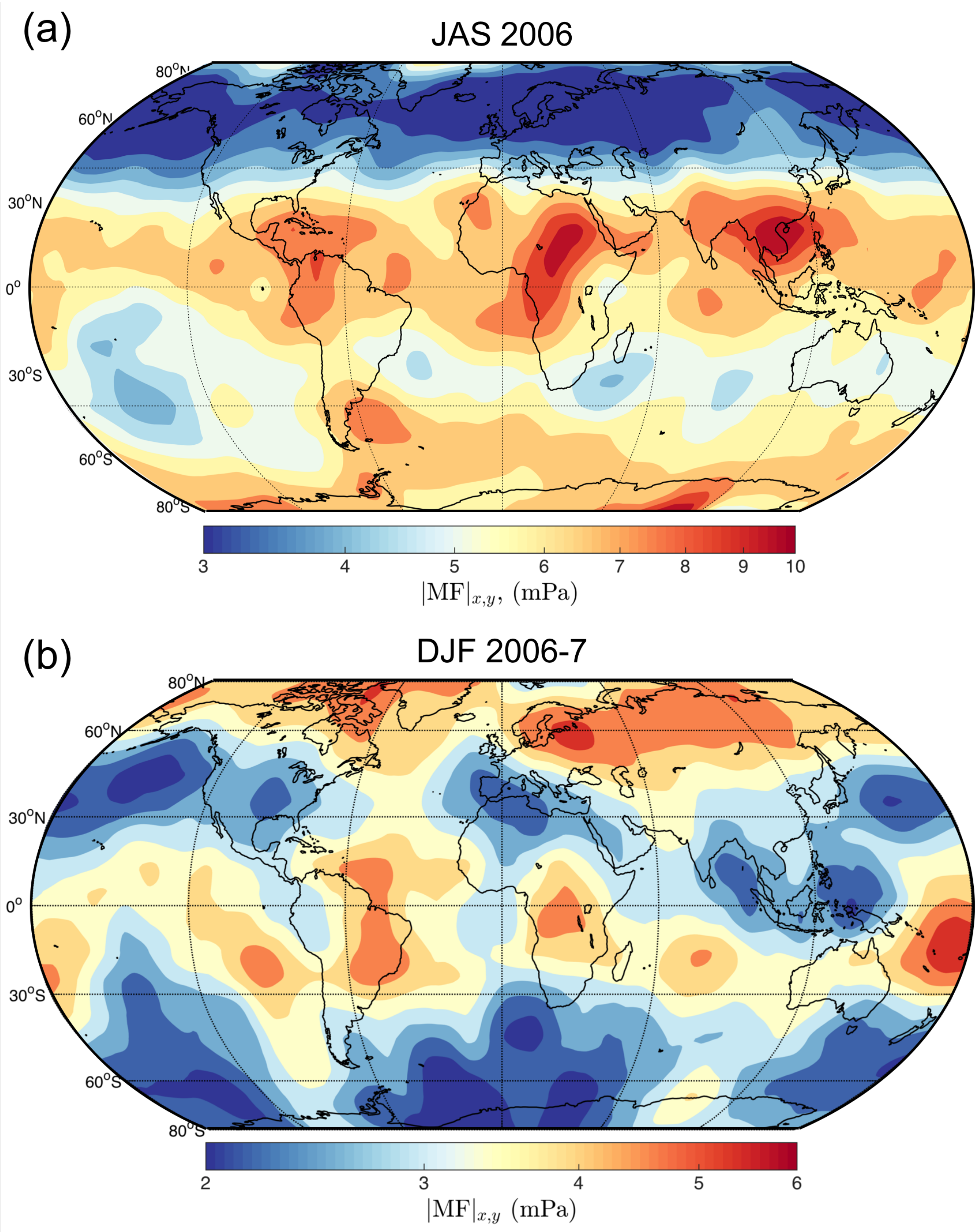

Figure 7.3: Global estimates of the magnitude of stratospheric gravity-wave momentum flux $|MF|_{x,y}$ from closely-spaced pairs of COSMIC GPS-RO profiles at around 30 km altitude for June to September 2006 (a) and December 2006 to February 2007 (b).

ple, the Upper Atmosphere Research Satellite Microwave Limb-Sounder (UARS-MLS) (Wu and Waters, 1996), CRISTA (Ern et al., 2004), HIRDLS (Yan et al., 2010) and AIRS (Hoffmann et al., 2013).

During December 2006 to February 2007 in Fig. 7.3(b), we see enhancements over South America and southern Africa and also over Scandinavia in the northern hemisphere. These features are in reasonably good agreement with results by Wang and Alexander (2010) and Faber et al. (2013), who also used COSMIC to make estimates of gravity wave momentum fluxes.

However, there are a number of features in Fig. 7.3(a) and (b) which are not so easily explained, and there is generally significant noise and ambiguity in the plots. These may be caused by incomplete removal of planetary wave features, or as a result of the 3-month time averaging required to acquire useful statistics in opposing seasons. In either case, increasing the sampling density of COSMIC GPS-RO would greatly improve the reliability, usefulness and temporal resolution of the gravity wave momentum flux technique presented here.

7.3.3 The upcoming COSMIC-2 GNSS-RO constellation

Fortunately, radio occultation is an expanding technique. The COSMIC-2 constellation (Cook et al., 2013) is scheduled for completion in the coming decade and will deploy 12 low-earth orbiting satellites into two orbital planes inclined at 72° and 24°, providing increased sampling density in equatorial regions over the first COSMIC constellation. Furthermore, the COSMIC-2 constellation will track and measure occultations of two more Global Navigation Satellite System (GNSS) constellations in addition to the American GPS system: the Russian Globalnaya Navigazionnaya Sputnikovaya Sistema (GLONASS) and European Galileo constellations.

COSMIC-2 will feature two deployment phases from which large numbers of closely-spaced profile-pairs can be expected, but the duration of this phase is of course limited. With such an increase in the number of soundings per day in the final constellation arrangement, a significant number of closely-spaced profile pairs may still arise, particularly with the inclusion of two more GNSS constellations.

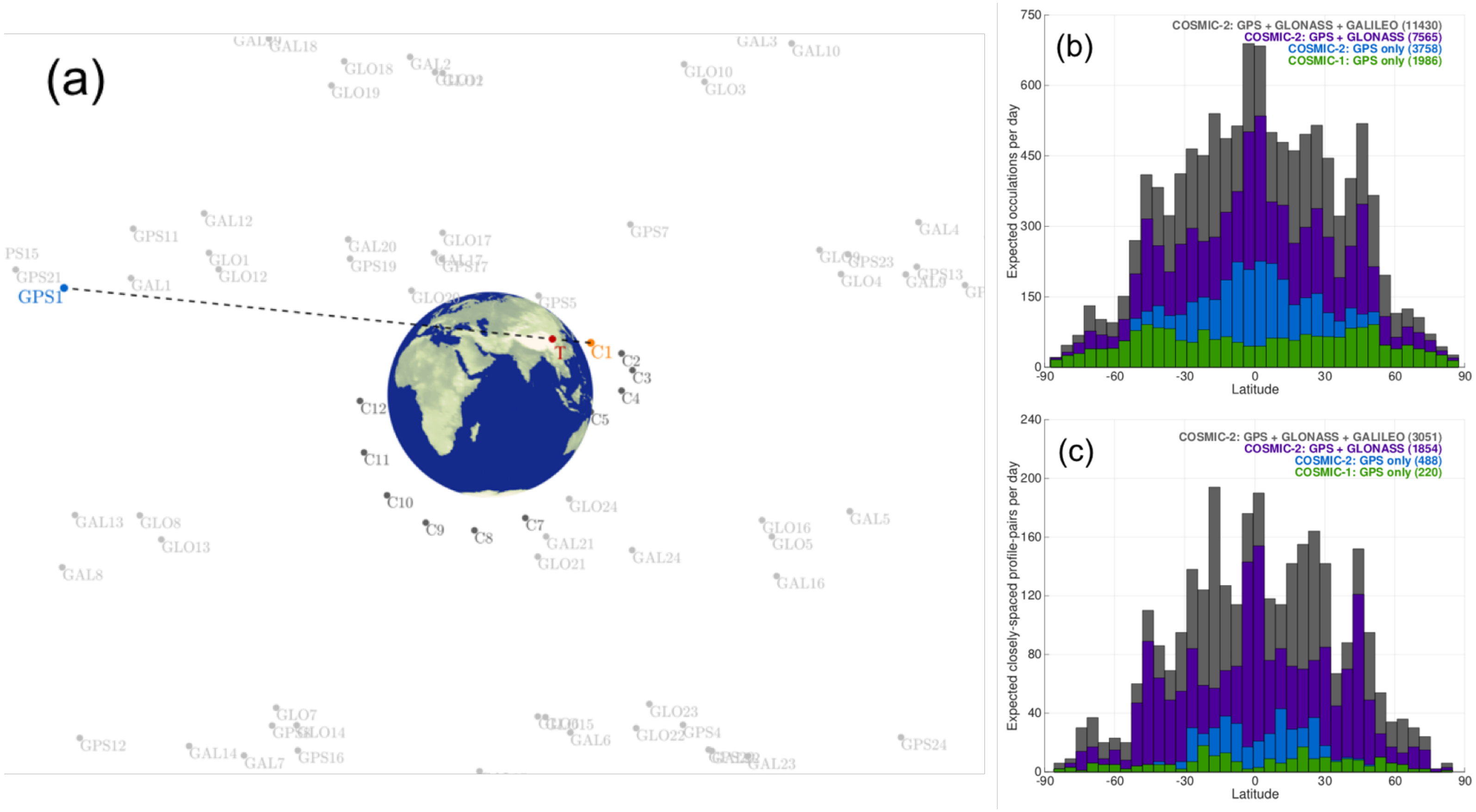

Figure 7.4: Snapshot of a numerical simulation of the COSMIC-2 (C1-12), GPS (GPS1-24), GLONASS (GLO1-24) and Galileo (GAL1-24) satellite constellations (a), with histograms of simulated daily occultations (b) and closely-spaced profile-pairs (c) against latitude provided by the COSMIC-2 constellation when using some or all of these GNSS constellations. Green bars in (b) and (c) also show the number of simulated occultations and profile-pairs against latitude for the current COSMIC-1 and GPS constellations. Note that the numbers of measured profiles in (b) and (c) have been reduced by a factor of 2 to simulate tracking losses and other errors. The black dashed line in (a) shows the approximate ray path of a GPS signal from an occulting GPS satellite as measured by a low-inclination COSMIC-2 satellite, with tangent point T in the earth's atmosphere.

7.3.4 A numerical GNSS-RO simulation

To investigate the numbers and distributions of occultations and closely-spaced profile pairs that might be expected from the COSMIC-2 constellation, a simple numerical simulation was created. Assuming circular orbits and nominal functionality of the COSMIC-2 and GNSS constellations, the simulation computes a to-scale model of the orbital dynamics of 84 satellites in total (Fig. 7.4), calculating their positions every 60 seconds for 30 days.

In the simulation, the 12 identical COSMIC-2 satellites are arranged at an altitude of 800 km into two orbital planes at 72° and 24° each containing 6 satellites, with a spacing of 30° between adjacent satellites in each plane.

Nominal deployment constellations are chosen for the simulated GPS, GLONASS and Galileo GNSS constellations, consisting of 24, 24 and 30 satellites in total arranged at orbital altitudes of 20 180 km, 19 140 km and 23 222 km respectively in 6, 3, and 3 orbital planes each containing 4, 8 and 8 satellites. These planes are inclined at 55°, 64.8° and 56° respectively, with 72°, 45° and 45° separations between adjacent satellites in each plane. Note that these nominally simulated configurations are subject to change compared to the real constellations. For example, the GPS constellation had a further 3 satellites added in 2011, and the Galileo constellation has yet to reach its full capacity. We also do not include the potential addition of the Chinese Beidou constellation currently under construction, which may or may not be used for radio occultation in the future. Nevertheless, this simulation provides a good preliminary estimate of global GNSS capabilities in the next decade or so.

In order to detect an occultation in our simulation, we track the projected latitude and longitude locations of each satellite on the earth's surface. From geometry, we know that when the great-circle distance (minor arc) on the earth's surface between the locations of any two COSMIC-2 and GNSS satellites equals a specific distance (11 741 km for GPS, 11 371 km for GLONASS and 11 598 km for Galileo), we know that the line of sight between the two satellites lies at a tangent to the earth's surface (black dashed line in Fig. 7.4(a)) and the GNSS satellite is occulting. Based on the success rates of Liou et al. (2007) and the exclusion of anomalous profiles we make the conservative estimate that, uniformly, only 50% of these simulated occultation events are likely to result in a robust, reliable useful atmospheric measurement profile at the tangent point T in Fig. 7.4(a).

Figure 7.4(b) shows a histogram of the mean number of COSMIC-2 occultation profiles against latitude produced by the simulation on a typical day. Two observations are immediately clear: the inclusion of 6 low-inclination COSMIC-2 satellites dramatically increases the number of profiles made in the tropics over the high-inclination COSMIC-1 constellation, even with only the GPS constellation (blue bars); and an enormous number of profiles could be measured if all three GNSS constellations could be used in their final deployment and operational (dark grey bars).

Figure 7.4(c) shows a histogram of the mean number of daily closely-spaced profile pairs against latitude from the simulation, with all satellites in their final deployment configuration. Note that, as with the COSMIC constellation, the deployment phase of the COSMIC-2 constellation will have extended periods where the satellites are physically close together, dramatically increasing the numbers of closely-spaced profile pairs during this time. As we can see from Fig. 7.4(c) however, very high numbers of profile pairs may also be available from the final deployment configuration of COSMIC-2 if the occultations of GLONASS and Galileo satellites are also measured. These high numbers of profile pairs may facilitate more in-depth study of global stratospheric gravity wave momentum fluxes from COSMIC-2 in the future.

7.4 What is the horizontal line-of-sight resolution of GPS radio occultation?

GPS Radio Occultation (GPS-RO) is a widely-used atmospheric measurement technique, boasting high vertical resolution and sub-kelvin temperature accuracy. In Chapter 3 we saw that the theoretical line-of-sight horizontal resolution of GPS-RO could be described as the horizontal width ΔL of an atmospheric layer along the ray path with height ΔZ (Kursinski et al., 1997) as

$$\Delta L = 2(2R\Delta Z)^{\frac{1}{2}}$$

where $R_T \sim 6400\,\text{km}$ is the radius of the tangent point. However, the vertical resolution of COSMIC GPS-RO is variable over different altitudes, depending on factors such as whether a geometric optics (GO, $\Delta Z \approx 1.4\,\text{km}$) or full-spectrum inversion (FSI, $\Delta Z \approx 0.2 - 0.3\,\text{km}$) retrieval method is used (Tsuda et al., 2011). For a range of vertical resolutions, we find

$$\Delta L = 2(2R\Delta Z)^{\frac{1}{2}} \approx \begin{cases} 270\,\text{km} & \text{for} \quad \Delta Z = 1.4\,\text{km} \\ 160\,\text{km} & \text{for} \quad \Delta Z = 0.5\,\text{km} \\ 101\,\text{km} & \text{for} \quad \Delta Z = 0.2\,\text{km} \end{cases} . \tag{7.1}$$

For a single bending angle measurement at some point in a vertical GPS-RO profile, the horizontal resolution can be defined by an alternative method. Distance x along the ray path through the atmospheric limb, centred at the tangent point, can be approximately written as

$$\begin{aligned}
x^2 &= (R_T + z)^2 - R_T^2 \\
x^2 &= R_T^2 + 2R_T z + z^2 - R_T^2 \\
x^2 &\approx 2R_T z \qquad (7.2)
\end{aligned}$$

where R_T is the radius of the tangent point and z is the altitude from the tangent point, if we assume $z^2 \ll 2R_T z$. If bending angle α is assumed to vary proportionally with refractivity n, which is assumed to vary proportionally with density ρ such that

$$\alpha \propto n \propto \rho(z) = \rho_T e^{\frac{-z}{H}} \qquad (7.3)$$

where ρ_T is density at the tangent point then, assuming spherical symmetry, atmospheric density as a function of distance x along the ray path is a Gaussian of the form

$$\rho(x) = \rho_T e^{\frac{-x^2}{2R_T H}}. \qquad (7.4)$$

Under aperture diffraction theory, two objects within one FWHM of each other cannot be individually distinguished. This implies, for a range of atmospheric scale heights, horizontal resolutions of

$$\Delta L = 2\sqrt{2\ln 2}\sqrt{R_T H} \approx \begin{cases} 461.4\,\mathrm{km} & \text{for} \quad \Delta Z = 1.4\,\mathrm{km} \\ 498.4\,\mathrm{km} & \text{for} \quad \Delta Z = 0.5\,\mathrm{km} \\ 532.8\,\mathrm{km} & \text{for} \quad \Delta Z = 0.2\,\mathrm{km} \end{cases} \qquad (7.5)$$

where σ is the variance of the Gaussian in Eqn. 7.4 above. Alternatively still, if we consider the region along the ray path where 50% of the net bending took place (Kursinski et al., 2000), i.e. containing 50% of the atmospheric mass along the ray path, given by $2 \times 0.69\sigma$, we find

$$\Delta L \approx 1.38\sqrt{R_T H} \approx \begin{cases} 270.4\,\mathrm{km} & \text{for} \quad H = 6\,\mathrm{km} \\ 292.1\,\mathrm{km} & \text{for} \quad H = 7\,\mathrm{km} \\ 312.3\,\mathrm{km} & \text{for} \quad H = 8\,\mathrm{km} \end{cases} \qquad (7.6)$$

Although there is broad agreement in the scale of these horizontal resolutions, clearly large disparity persists. Ionospheric corrections, or assumptions inherent in the temperature retrieval of COSMIC may further affect these values. For gravity wave study, where understanding the precise spectral range of waves to which we are sensitive is key, this is an area in which we should seek to make improvements. Furthermore, all of these limits are calculated theoretically and, to our knowledge, this aspect of GPS-RO has not been tested observationally.

To improve our understanding the line-of-sight resolution of GPS-RO, we can combine COSMIC measurements with co-located measurements from AIRS and, if required, IASI (Hoffmann et al., 2014). An obvious location for this comparison is the southern Andes and Antarctic Peninsula, where mountain wave scales, structures and directions are reasonably well-understood. Crucially, in Figures 5.6 and 5.7 we can see that waves in this region can have horizontal wavelengths both larger and smaller the limits calculated above.

We would first identify a clear wave-like structure in an AIRS granule. We would then measure its horizontal wavelength and direction of propagation using our 2DST method developed in Chapter 5. Next we can then take a co-located COSMIC profile and project this horizontal wavelength into the line-of-sight of the occultation ray path. We can then check if the COSMIC temperature perturbation profile contained a wave using our Wave-ID method developed in Chapter 4 or by inspection, and perform checks to see if this wave is likely to be the same wave structure resolved by AIRS. We can perform simple checks such as ensuring $\lambda_Z > 12\,\text{km}$ (if we use AIRS channel 75) in the COSMIC profile, or make predictions of what the vertical wavelength might be by assuming a mountain wave in AIRS and using Eqn. 5.36.

If the wave is detected, and its vertical wavelength is admissible, then we know that the projected horizontal wavelength of the wave resolved by AIRS is also resolvable by COSMIC. If it is not, we know that the horizontal resolution of COSMIC is greater than this wavelength. By repeating this process for a number of case-studies by means of an automated routine, we can begin to more carefully define the line-of-sight horizontal resolution of GPS-RO observationally for the real atmosphere, information which may prove to be very useful going forward.

7.5 Gravity wave analysis of airglow imagery using the two-dimensional Stockwell transform

In Chapter 5 we described a new application of the two-dimensional Stockwell transform (2DST) for gravity wave analysis of AIRS data. A logical next study might be to apply the 2DST to the purpose of gravity wave analysis of airglow imagery.

Airglow is the emission of (usually visible to near-infrared) radiation from the atmo-

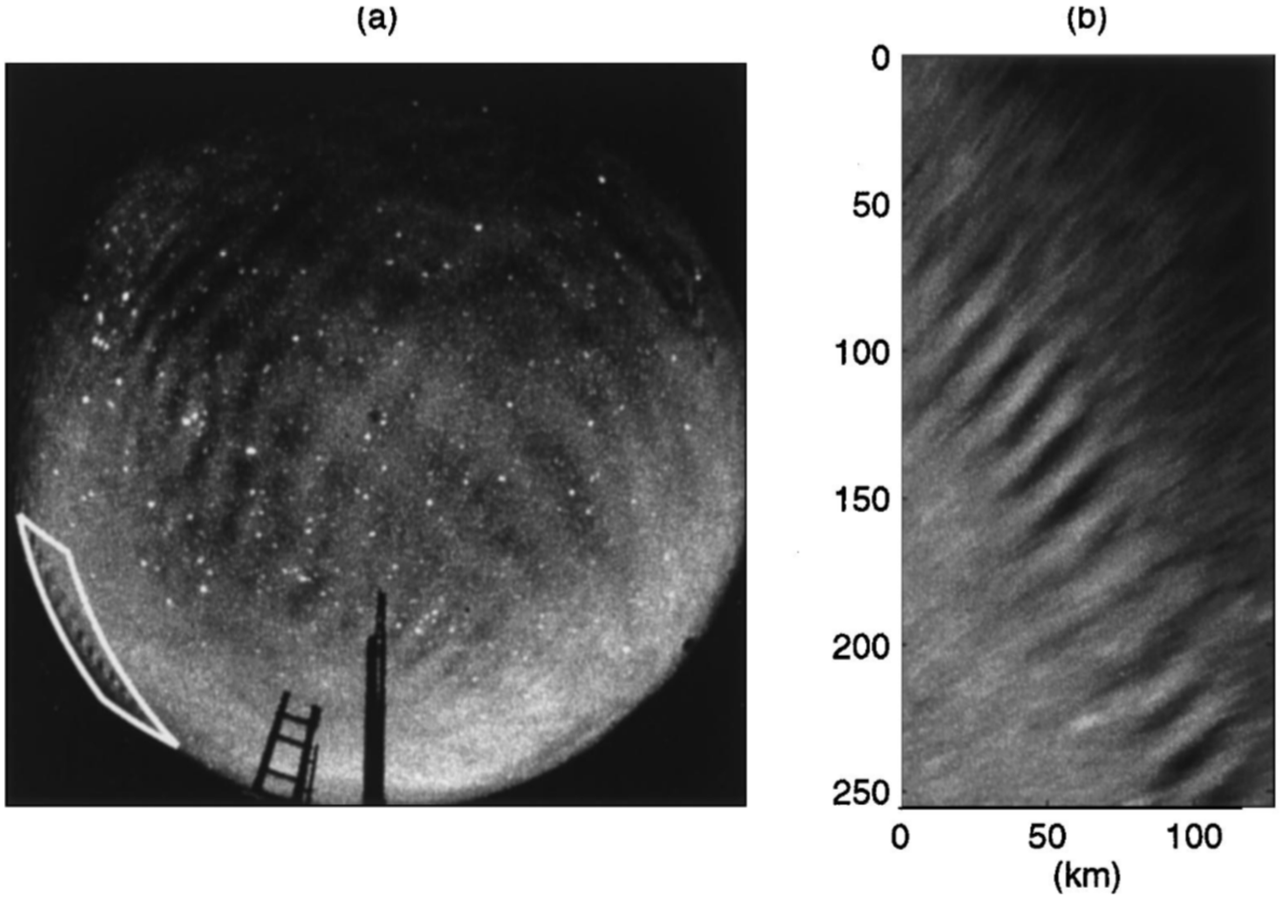

Figure 7.5: 512×512-pixel resolution all-sky camera image (a) taken with a 90-second exposure at 10:19 UTC on 22nd October 1993 from Haleakala, Hawaii, in which mesospheric gravity waves can be seen. Panel (b) shows an unwarped image of the region in the white border in (a), showing small-scale waves at very low elevation. From Garcia et al. (1997).

sphere itself, and can be measured by ground-based nadir-viewing all-sky cameras at numerous locations across the globe. One location of interest might the Rothera base on the Antarctic Peninsula, where one such all-sky camera is based, in a known region of intense gravity-wave activity. Airglow measurements are warped 2-D images of the local mesosphere, where gravity wave perturbations can be visible as huge waves across the sky (Fig. 7.5).

The 2DST could be a useful tool for detecting and measuring these waves in these 2-D data. Coincident measurements from AIRS, viewing in the stratosphere, could also be used to investigate whether large coherent wave fields could extend all the way from the troposphere up through the stratosphere and into the mesosphere, potentially providing useful information on troposphere-mesosphere coupling.

References

Abramowitz, M. and I. Stegun, 1964: *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied mathematics series, Dover Publications.

Alexander, M. J., 1998: Interpretations of observed climatological patterns in stratospheric gravity wave variance. *J. Geophys. Res.*, **103**, 8627–8640, doi:10.1029/97JD03325.

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., S. D. Eckermann, D. Broutman, and J. Ma, 2009: Momentum flux estimates for South Georgia Island mountain waves in the stratosphere observed via satellite. *Geophy. Res. Lett.*, **36**, L12816, doi:10.1029/2009GL038587.

Alexander, M. J., M. Geller, C. McLandress, S. Polavarapu, P. Preusse, F. Sassi, K. Sato, S. Eckermann, M. Ern, A. Hertzog, Y. Kawatani, M. Pulido, T. A. Shaw, M. Sigmond, R. Vincent, and S. Watanabe, 2010: Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models. *Quart. J. Roy. Meteor. Soc.*, **136**, 1103–1124, doi:10.1002/qj.637.

Alexander, M. J., J. Gille, C. Cavanaugh, M. Coffey, C. Craig, T. Eden, G. Francis, C. Halvorson, J. Hannigan, R. Khosravi, D. Kinnison, H. Lee, S. Massie, B. Nardi, J. Barnett, C. Hepplewhite, a. Lambert, and V. Dean, 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, 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.*, **118**, 11589–11599, doi:10.1002/2013JD020526.

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**, n/a–n/a, doi:10.1029/2006JD008368.

— 2011: Three-dimensional properties of Andes mountain waves observed by satellite: A case study. *J. Geophys. Res.*, **116**, n/a–n/a, doi:10.1029/2011JD016151.

Alexander, P. and A. de la Torre, 2011: A method to improve the determination of wave perturbations close to the tropopause by using a digital filter. *Atmos. Meas. Tech.*, **4**, 1777–1784, doi:10.5194/amt-4-1777-2011.

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**, doi:10.1029/2009JD011851, d17103.

Alexander, S. P., T. Tsuda, Y. Kawatani, and M. Takahashi, 2008: Global distribution of atmospheric waves in the equatorial upper troposphere and lower stratosphere: COSMIC observations of wave mean flow interactions. *J. Geophys. Res.*, **113**, n/a–n/a, doi:10.1029/2008JD010039, d24115.

Andrews, D. G., 2000: *An Introduction to Atmospheric Physics*. Cambridge University Press.

Andrews, D. G., F. W. Taylor, and M. E. McIntyre, 1987: The influence of atmospheric waves on the general circulation of the middle atmosphere. *Philosophical Transactions of the Royal Society A - Mathematical Physical and Engineering Sciences*, **323**, 693–705, doi:10.1098/rsta.1987.0115.

Anthes, R. A., P. A. Bernhardt, Y. Chen, L. Cucurull, K. F. Dymond, D. Ector, S. B. Healy, S. P. Ho, D. C. Hunt, Y. H. Kuo, H. Liu, K. Manning, C. Mccormick, T. K. Meehan, W. J. Randel, C. Rocken, W. S. Schreiner, S. V. Sokolovskiy, S. Syndergaard, D. C. Thompson, K. E. Trenberth, T. K. Wee, N. L. Yen, and Z. Zeng, 2008: The COSMOC/FORMOSAT-3 - Mission early results. *Bulletin of the American Meteorological Society*, **89**, 313+, doi:10.1175/BAMS-89-3-313.

Aumann, H., M. Chahine, C. Gautier, M. Goldberg, E. Kalnay, L. McMillin, H. Revercomb, P. Rosenkranz, W. Smith, D. Staelin, L. Strow, and J. Susskind, 2003: Airs/amsu/hsb on the aqua mission: design, science objectives, data products, and processing systems. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 253–264, doi:10.1109/TGRS.2002.808356.

Bacmeister, J. T. and M. R. Schoeberl, 1989: Breakdown of Vertically Propagating Two-Dimensional Gravity Waves Forced by Orography. *J. Atmos. Sci.*, **46**, 2109–2134.

Baldwin, M. P., L. J. Gray, T. J. Dunkerton, K. Hamilton, P. H. Haynes, W. J. Randel, J. R. Holton, M. J. Alexander, I. Hirota, T. Horinouchi, D. B. A. Jones, J. S. Kinnersley, C. Marquardt, K. Sato, and M. Takahashi, 2001: The quasi-biennial oscillation. *Reviews of Geophysics*, **39**, 179–229, doi:10.1029/1999RG000073.

Barry, R. L., S. C. Strother, and J. C. Gore, 2012: Complex and magnitude-only preprocessing of 2D and 3D BOLD fMRI data at 7T. *Magnetic Resonance in Medicine*, **67**, 867–871, doi:10.1002/mrm.23072.

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.

Beldon, C. L. and N. J. Mitchell, 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**, 875–884, doi:10.1016/j.jastp.2009.03.009.

Beyerle, G., T. Schmidt, G. Michalak, S. Heise, J. Wickert, and C. Reigber, 2005: GPS radio occultation with GRACE: Atmospheric profiling utilizing the zero difference technique. *Geophy. Res. Lett.*, **32**, doi:10.1029/2005GL023109.

Brigham, E. O., 1974: *The Fast Fourier Transform*. Prentice-Hall Inc.

Brown, R. A., M. L. Lauzon, and R. Frayne, 2010: A General Description of Linear Time-Frequency Transforms and Formulation of a Fast, Invertible Transform That Samples the Continuous S-Transform Spectrum Nonredundantly. *IEEE Trans. Sig. Proc.*, **58**, 281–290, doi:10.1109/TSP.2009.2028972.

Butchart, N., A. J. Charlton-Perez, I. Cionni, S. C. Hardiman, P. H. Haynes, K. Krger, P. J. Kushner, P. A. Newman, S. M. Osprey, J. Perlwitz, M. Sigmond, L. Wang, H. Akiyoshi, J. Austin, S. Bekki, A. Baumgaertner, P. Braesicke, C. Brhl, M. Chipperfield, M. Dameris, S. Dhomse, V. Eyring, R. Garcia, H. Garny, P. Jckel, J.-F. Lamarque, M. Marchand, M. Michou, O. Morgenstern, T. Nakamura, S. Pawson, D. Plummer, J. Pyle, E. Rozanov, J. Scinocca, T. G. Shepherd, K. Shibata, D. Smale, H. Teyssdre, W. Tian, D. Waugh, and Y. Yamashita, 2011: Multimodel climate and variability of the stratosphere. *J. Geophys. Res.*, **116**, n/a–n/a, doi:10.1029/2010JD014995, d05102.

Cavcar, M., 2000: The international standard atmosphere (isa). Anadolu University, 26470, Eskisehir, Turkey.

Clerbaux, C., A. Boynard, L. Clarisse, M. George, J. Hadji-Lazaro, H. Herbin, D. Hurtmans, M. Pommier, A. Razavi, S. Turquety, C. Wespes, and P. F. Coheur, 2009: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmos. Chem. Phys.*, **9**, 6041–6054.

Cook, K., C.-J. Fong, M. Wenkel, P. Wilczynski, N. Yen, and G. Chang, 2013: Formosat-7/cosmic-2 gnss radio occultation constellation mission for global weather monitoring, 1–8. doi:10.1109/AERO.2013.6497317.

Derber, J.: 2014, The impact and importance of GPS-RO in NWP. *Fifth EUMETSAT ROM-SAF user workshop on Applications of GPS Radio Occultation Measurements*, European Centre for Medium-Range Weather Forecasts, June 16–18, 2014, Reading, UK.

Dunkerton, T. J., 1997: The role of gravity waves in the quasi-biennial oscillation. *J. Geophys. Res.*, **102**, 26053–26076, doi:10.1029/96JD02999.

Durran, D. R.: 2003, Lee Waves and Mountain Waves. *Encyclopedia of the Atmospheric Sciences*, Holton, J. R. and Pyle, J. and Curry, J. A., ed., 1161–1169.

Eckermann, S. D. and P. Preusse, 1999: Global Measurements of Stratospheric Mountain Waves from Space. *Science*, **286**, 1534–1537, doi:10.1126/science.286.5444.1534.

Egger, J. and K. P. Hoinka, 2005: Downward control from the lower stratosphere? *J. Atmos. Sci.*, **62**, 3808–3817, doi:10.1175/JAS3569.1.

Ern, M. and P. Preusse, 2012: Gravity wave momentum flux spectra observed from satellite in the summertime subtropics: Implications for global modeling. *Geophy. Res. Lett.*, **39**, doi:10.1029/2012GL052659.

Ern, M., P. Preusse, M. J. Alexander, and C. D. Warner, 2004: Absolute values of gravity wave momentum flux derived from satellite data. *J. Geophys. Res.*, **109**, D20103, doi:10.1029/2004JD004752.

Ern, M., P. Preusse, J. C. Gille, C. L. Hepplewhite, M. G. Mlynczak, J. M. Russell, and M. Riese, 2011: Implications for atmospheric dynamics derived from global observations of gravity wave momentum flux in stratosphere and mesosphere. *J. Geophys. Res.*, **116**, D19107, doi:10.1029/2011JD015821.

Ern, M., P. Preusse, S. Kalisch, M. Kaufmann, and M. Riese, 2013: Role of gravity waves in the forcing of quasi two-day waves in the mesosphere: An observational study. *J. Geophys. Res.*, **118**, 3467–3485, doi:10.1029/2012JD018208.

Faber, A., P. Llamedo, T. Schmidt, A. de la Torre, and J. Wickert, 2013: On the determination of gravity wave momentum flux from GPS radio occultation data. *Atmos. Meas. Tech.*, **6**, 3169–3180, doi:10.5194/amt-6-3169-2013.

Fjeldbo, G. and V. R. Eshleman, 1965: Bistatic radar-occultation method for study of planetary atmospheres. *JOURNAL OF GEOPHYSICAL RESEARCH*, **70**, 3217–&, doi:10.1029/JZ070i013p03217.

— 1969: Atmosphere of venus as studied with the mariner 5 dual radio-frequency occultation experiment. *Radio Science*, **4**, 879–897, doi:10.1029/RS004i010p00879.

Fjeldbo, G., A. J. Kliore, and V. R. Eshlemen, 1971: Neutral Atmosphere of Venus as studied with Mariner-V radio occultation experiments. *The Astronomical Journal*, **76**, 123–140, doi:10.1086/111096.

Fong, C., W. Shiau, C. Lin, T. Kuo, C. Chu, S. Yang, N. L. Yen, S. Chen, Y. Kuo, Y. Liou, and S. Chi, 2008: Constellation Deployment for the FORMOSAT-3/COSMIC Mission. *IEEE Transactions on Geoscience and Remote Sensing*, **46**, 3367–3379, doi:10.1109/TGRS.2008.2005202.

France, J. A., V. L. Harvey, M. J. Alexander, C. E. Randall, and J. C. Gille, 2012: High Resolution Dynamics Limb Sounder observations of the gravity wave-driven elevated stratopause in 2006. *J. Geophys. Res.*, **117**, doi:10.1029/2012JD017958.

Fritts, D. C. and M. J. Alexander, 2003: Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics*, **41**, 1003, doi:10.1029/2001RG000106.

Fritts, D. C. and P. K. Rastogi, 1985: Convective and dynamical instabilities due to gravity wave motions in the lower and middle atmosphere: Theory and observations. *Radio Science*, **20**, 1247–1277, doi:10.1029/RS020i006p01247.

Fritts, D. C., D. M. Riggin, B. B. Balsley, and R. G. Stockwell, 1998: Recent results with an mf radar at mcmurdo, antarctica: Characteristics and variability of motions near 12-hour period in the mesosphere. *Geophy. Res. Lett.*, **25**, 297–300, doi:10.1029/97GL03702.

Garcia, F. J., M. J. Taylor, and M. C. Kelley, 1997: Two-dimensional spectral analysis of mesospheric airglow image data. *Applied Optics*, **36**, 7374–7385.

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.

Garcia, R. R. and W. J. Randel, 2008: Acceleration of the Brewer-Dobson circulation due to increases in greenhouse gases. *J. Atmos. Sci.*, **65**, 2731–2739, doi:10.1175/2008JAS2712.1.

Garcia, R. R. and S. Solomon, 1983: A numerical model of the zonally averaged dynamical and chemical structure of the middle atmosphere. *J. Geophys. Res.*, **88**, 1379–1400, doi:10.1029/JC088iC02p01379.

Geller, M., M. J. Alexander, P. Love, J. Bacmeister, M. Ern, A. Hertzog, E. Manzini, P. Preusse, K. Sato, A. Scaife, and T. Zhou, 2013: A Comparison between Gravity Wave Momentum Fluxes in Observations and Climate Models. *Journal of Climate*, **26**, 6383–6405, doi:10.1175/JCLI-D-12-00545.1.

Gibson, P. C., M. P. Lamoureux, and G. F. Margrave, 2006: Letter to the editor: Stockwell and wavelet transforms. *Journal of Fourier Analysis and Applications*, **12**, 713–721, doi:10.1007/s00041-006-6087-9.

Gille, J., J. Barnett, and Co-authors, 2008: High Resolution Dynamics Limb Sounder: Experiment overview, recovery, and validation of initial temperature data. *J. Geophys. Res.*, **113**, doi:10.1029/2007JD008824.

Gille, J., J. Barnett, J. Whitney, M. Dials, D. Woodard, W. Rudolf, A. Lambert, and W. Mankin: 2003, The high resolution dynamics limb sounder (HIRDLS) experiment on aura. *Infrared Spaceborne Remote Sensing XI*, Strojnik, M., ed., SPIE, volume 5152 of *Proceedings of SPIE-The International Society for Optical Engineering*, 162–171, Conference on Infrared Spaceborne Remote Sensing XI, SAN DIEGO, CA, AUG 06-08, 2003.

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.

Gong, J., J. Yue, and D. L. Wu, 2015: Global survey of concentric gravity waves in airs images and ecmwf analysis. *J. Geophys. Res.*, **120**, 2210–2228, doi:10.1002/2014JD022527.

Goodyear, B. G., H. M. Zhu, R. A. Brown, and J. R. Mitchell, 2004: Removal of phase artifacts from fMRI data using a Stockwell transform filter improves brain activity detection. *Magnetic Resonance in Medicine*, **51**, 16–21, doi:10.1002/mrm.10681.

Guest, F. M., M. J. Reeder, C. J. Marks, and D. J. Karoly, 2000: Inertia-gravity waves observed in the lower stratosphere over Macquarie Island. *J. Atmos. Sci.*, **57**, 737–752, doi:10.1175/1520-0469.

Hajj, G., E. R. Kursinski, and L. J. Romans, 2002: A technical description of atmospheric sounding by GPS occultation. *J. Atmos. Sci.*, **64**, 451–469.

Hajj, G. A., C. O. Ao, B. A. Iijima, D. Kuang, E. R. Kursinski, A. J. Mannucci, T. K. Meehan, L. J. Romans, M. D. Juarez, and T. P. Yunck, 2004: CHAMP and SAC-C atmospheric occultation results and intercomparisons. *J. Geophys. Res.*, **109**, doi:10.1029/2003JD003909.

Hajj, G. A. and L. J. Romans, 1998: Ionospheric electron density profiles obtained with the global positioning system: Results from the GPS/MET experiment. *Radio Science*, **33**, 175–190, doi:10.1029/97RS03183.

Hei, H., T. Tsuda, and T. Hirooka, 2008: Characteristics of atmospheric gravity wave activity in the polar regions revealed by GPS radio occultation data with CHAMP. *J. Geophys. Res.*, **113**, D04107, doi:10.1029/2007JD008938.

Hendricks, E., J. Doyle, S. D. Eckermann, Q. Jiang, and P. Reinecke, 2014: What Is the Source of the Stratospheric Gravity Wave Belt in Austral Winter? *J. Atmos. Sci.*, **71**, 1583–1592, doi:10.1175/JAS-D-13-0332.1.

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.

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**, 3056–3070, doi:10.1175/2008JAS2710.1.

Hindley, N. P., N. D. Smith, D. A. S. Wright, C. J. Rees, and N. J. Mitchell, 2016: A two-dimensional stockwell transform for gravity wave analysis of airs measurements. *Atmos. Chem. Phys. Disc.*, doi:10.5194/amt-2015-383.

Hindley, N. P., C. J. Wright, N. D. Smith, and N. J. Mitchell, 2015: The southern stratospheric gravity wave hot spot: individual waves and their momentum fluxes measured by cosmic gps-ro. *Atmos. Chem. Phys.*, **15**, 7797–7818, doi:10.5194/acp-15-7797-2015.

Hines, C., 1974: *The upper atmosphere in motion: a selection of papers with annotation*. Geophysical Monograph, American Geophysical Union.

Hines, C. O., 1960: Internal atmospheric gravity waves at ionospheric heights. *Canadian Journal of Physics*, **38**, 1441–1481.

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, L., M. J. Alexander, C. Clerbaux, A. W. Grimsdell, C. I. Meyer, T. Roessler, and B. Tournier, 2014: Intercomparison of stratospheric gravity wave observations with AIRS and IASI. *Atmos. Meas. Tech.*, **7**, 4517–4537, doi:10.5194/amt-7-4517-2014.

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.*, **118**, 416–434, doi:10.1029/2012JD018658.

Holton, J. R. and M. J. Alexander, 1999: Gravity waves in the mesosphere generated by tropospheric convention. *Tellus A*, **51**, 45–58, doi:10.1034/j.1600-0870.1999.t01-1-00005.x.

Höpfner, M., N. Larsen, R. Spang, B. P. Luo, J. Ma, S. H. Svendsen, S. D. Eckermann, B. Knudsen, P. Massoli, F. Cairo, G. Stiller, T. v. Clarmann, and H. Fischer, 2006: Mipas detects antarctic stratospheric belt of nat pscs caused by mountain waves. *Atmos. Chem. Phys.*, **6**, 1221–1230, doi:10.5194/acp-6-1221-2006.

Hoskins, B. and K. Hodges, 2005: A new perspective on Southern Hemisphere storm tracks. *Journal of Climate*, **18**, 4108–4129, doi:10.1175/JCLI3570.1.

Houghton, J. T., 1977: *The Physics of Atmospheres*. Cambridge University Press.

Jiang, J. H., B. Wang, K. Goya, K. Hocke, S. D. Eckermann, J. Ma, D. L. Wu, and W. G. Read, 2004: Geographical distribution and interseasonal variability of tropical deep convection: Uars mls observations and analyses. *J. Geophys. Res.*, **109**, n/a–n/a, doi:10.1029/2003JD003756, d03111.

Jiang, J. H., D. L. Wu, and S. D. Eckermann, 2002: Upper Atmosphere Research Satellite (UARS) MLS observation of mountain waves over the Andes. *J. Geophys. Res.*, **107**, doi:10.1029/2002JD002091.

Kliore, A., G. S. Levy, D. L. Cain, V. R. Eshleman, G. Fjeldbo, and F. D. Drake, 1965: Mariner 4 measurements near Mars - Intitial results - Occultation experiment - Results of 1st direct measurement of Mars atmosphere and ionosphere. *Science*, **149**, doi:10.1126/science.149.3689.1243.

Kocahan, O., E. Coskun, and S. Oezder: 2008, Profile Measurement of Objects by Using Stockwell and Continuous Wavelet Transforms. *2008 IEEE 16TH Signal Processing, Communication and Applications Conference, Vols. 1 and 2*, IEEE, 804–807, IEEE

16th Signal Processing and Communications Applications Conference, Aydin, Turkey, APR 20-22, 2008.

Kursinski, E. R., G. A. Hajj, W. I. Bertiger, S. S. Leroy, T. K. Meehan, L. J. Romans, J. T. Schofield, D. J. McCleese, W. G. Melbourne, C. L. Thornton, T. P. Yunck, J. R. Eyre, and R. N. Nagatani, 1996: Initial results of radio occultation observations of earth's atmosphere using the global positioning system. *Science*, **271**, 1107–1110, doi:10.1126/science.271.5252.1107.

Kursinski, E. R., G. A. Hajj, S. S. Leroy, and B. Herman, 2000: The GPS radio occultation technique. *Terrestrial Atmospheric and Oceanic Sciences*, **11**, 53–114.

Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.*, **102**, 23429–23465, doi:10.1029/97JD01569.

Kuyuk, H. S., 2015: On the use of Stockwell transform in structural dynamic analysis. *Sadhana-Academy Proceedings in Engineering Sciences*, **40**, 295–306, doi:10.1007/s12046-014-0301-2.

Liou, Y. A., A. G. Pavelyev, A. A. Pavelyev, N. Yen, C. Y. Huang, and C. J. Fong, 2007: FORMOSAT-3/COSMIC GPS radio occultation mission: Preliminary results. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 3813–3826.

Liu, H. L. and R. G. Roble, 2002: A study of a self-generated stratospheric sudden warming and its mesospheric-lower thermospheric impacts using the coupled TIME-GCM/CCM3. *J. Geophys. Res.*, **107**, doi:10.1029/2001JD001533.

Liu, W., F. P. Bretherton, Z. Liu, L. Smith, H. Lu, and C. J. Rutland, 2010: Breaking of Progressive Internal Gravity Waves: Convective Instability and Shear Instability. *Journal of Physical Oceanography*, **40**, 2243–2263, doi:10.1175/2010JPO4432.1.

Liu, Y.: 2009, Localization Operators for Two-Dimensional Stockwell Transforms. *New Developments in Pseudo-Differential Operators*, Rodino, L. and Wong, M. W., ed., Int Soc Anal Its Applicat & Computat, volume 189 of *Operator Theory Advances and Applications*, 287–296, 6th Congress of the International-Society-for-Analysis-Its-Applications-and-Computation, Middle East Tech Univ, Ankara, Turkey, AUG 13-18, 2007.

Liu, Y. and M. W. Wong: 2007, Inversion formulas for two-dimensional Stockwell transforms. *Pseudo-Differential Operators: Partial Differential Equations and time-frequency analysis*, Rodino, L and Schulze, BW and Wong, MW, ed., Int Soc Anal, Applicat & Computat; Acad Initiat Fund; York Univ , Dept Math & Statist; York Univ, Off Vice-President Acad, volume 52 of *Fields Institute Communications*, 323–330, Workshop on Pseudo-Differential Operators - Partial Differential Equations and Time-Frequency Analysis, Toronto, Canada, Dec 11-15, 2006.

London, J.: 1980, Radiative energy sources and sinks in the stratosphere and mesosphere. *Proceedings of the NATO Advanced Study Institute on Atmospheric Ozone: Its Variation and Human Influences*, 703–721.

Mansinha, L., R. G. Stockwell, and R. P. Lowe, 1997a: Pattern analysis with two-dimensional spectral localisation: Applications of two-dimensional S transforms. *Physica A.*, **239**, 286–295, doi:10.1016/S0378-4371(96)00487-6, Proceedings of the International Conference on Pattern Formation in Fluids and Materials CPiP 96 (Collective Phenomena in Physics 96), University of Western Ontario, London, Canada, June 13-15, 1996.

Mansinha, L., R. G. Stockwell, R. P. Lowe, M. Eramian, and R. A. Schincariol, 1997b: Local S-spectrum analysis of 1-D and 2-D data. *Physics of the Earth and Planetary Interiors*, **103**, 329–336, doi:10.1016/S0031-9201(97)00047-2, Conference on Geonomy in honor of Professor J A Jacobs, Edinburgh, Scotland, June 03-04, 1996.

Marquardt, C. and S. Healy, 2005: Measurement noise and stratospheric gravity wave characteristics obtained from GPS occultation data. *J. Meteor. Soc. Jap.*, **83**, 417–428, doi:10.2151/jmsj.83.417.

McDonald, A. J., 2012: Gravity wave occurrence statistics derived from paired COSMIC/FORMOSAT3 observations. *J. Geophys. Res.*, **117**, D15106, doi:10.1029/2011JD016715.

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

Mitchell, N. J., L. Thomas, and A. K. P. Marsh, 1991: LIDAR observations of long-period gravity waves in the stratosphere. **9**, 588–596.

Nappo, C., 2013: *An Introduction to Atmospheric Gravity Waves*. International geophysics series, Academic Press/Elsevier.

O'Sullivan, D. and T. J. Dunkerton, 1995: Generation of inertia-gravity waves in a simulated life-cycle of baroclinic instability. *J. Atmos. Sci.*, **52**, 3695–3716, doi:10.1175/1520-0469.

Pinnegar, C. R. and L. Mansinha, 2003: The S-transform with windows of arbitrary and varying shape. *Geophysics*, **68**, 381–385, doi:10.1190/1.1543223.

Plougonven, R., A. Arsac, A. Hertzog, L. Guez, and F. Vial, 2010: Sensitivity study for mesoscale simulations of gravity waves above Antarctica during Vorcore. *Quart. J. Roy. Meteor. Soc.*, 1371–1377, doi:10.1002/qj.639.

Plougonven, R., A. Hertzog, and L. Guez, 2013: Gravity waves over Antarctica and the Southern Ocean: consistent momentum fluxes in mesoscale simulations and stratospheric balloon observations. *Quart. J. Roy. Meteor. Soc.*, **139**, 101–118, doi:10.1002/qj.1965.

Plougonven, R. and F. Zhang, 2014: Internal gravity waves from atmospheric jets and fronts. *Reviews of Geophysics*, **52**, 33–76, doi:10.1002/2012RG000419.

Preusse, P., A. Doernbrack, and S. Eckermann, 2002: Space-based measurements of stratospheric mountain waves by CRISTA 1. Sensitivity, analysis method, and a case study. *J. Geophys. Res.*, **107**, 8178, doi:10.1029/2001JD000699.

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.

Preusse, P., M. Ern, P. Bechtold, S. D. Eckermann, S. Kalisch, Q. T. Trinh, and M. Riese, 2014: Characteristics of gravity waves resolved by ECMWF. *Atmos. Chem. Phys.*, **14**, doi:10.5194/acp-14-10483-2014.

Preusse, P., S. Schroeder, L. Hoffmann, M. Ern, F. Friedl-Vallon, J. Ungermann, H. Oelhaf, H. Fischer, and M. Riese, 2009: New perspectives on gravity wave remote sensing by spaceborne infrared limb imaging. *Atmos. Meas. Tech.*, **2**, 299–311, doi:10.5194/amt-2-299-2009.

Rocken, C., R. Anthes, M. Exner, D. Hunt, S. Sokolovskiy, R. Ware, M. Gorbunov, W. Schreiner, D. Feng, B. Herman, Y. H. Kuo, and X. Zou, 1997: Analysis and validation of GPS/MET data in the neutral atmosphere. *J. Geophys. Res.*, **102**, 29849–29866, doi:10.1029/97JD02400.

Rocken, C., Y. H. Kuo, W. S. Schreiner, D. Hunt, S. Sokolovskiy, and C. McCormick, 2000: COSMIC system description. *Terrestrial Atmospheric and Oceanic Sciences*, **11**, 21–52.

Rosenhead, L., 1931: The formation of vortices from a surface of discontinuity. *Proceedings of the Royal Society of London*, **134**, 170–192.

Sato, K., S. Tateno, S. Watanabe, and Y. Kawatani, 2012: Gravity Wave Characteristics in the Southern Hemisphere Revealed by a High-Resolution Middle-Atmosphere General Circulation Model. *J. Atmos. Sci.*, **69**, 1378–1396, doi:10.1175/JAS-D-11-0101.1.

Sato, K., S. Watanabe, Y. Kawatani, Y. Tomikawa, K. Miyazaki, and M. Takahashi, 2009: On the origins of mesospheric gravity waves. *Geophy. Res. Lett.*, **36**, L19801, doi:10.1029/2009GL039908.

Savitzky, A. and M. J. E. Golay, 1964: Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, **36**, 1627–1639.

Scaife, A., N. Butchart, C. Warner, D. Stainforth, W. Norton, and J. Austin, 2000: Realistic quasi-biennial oscillations in a simulation of the global climate. *Geophy. Res. Lett.*, **27**, 3481–3484, doi:10.1029/2000GL011625.

Schreiner, W., C. Rocken, S. Sokolovskiy, S. Syndergaard, and D. Hunt, 2007: Estimates of the precision of GPS radio occultations from the COSMIC/FORMOSAT-3 mission. *Geophy. Res. Lett.*, **34**, doi:10.1029/2006GL027557.

Schreiner, W. S., S. V. Sokolovskiy, C. Rocken, and D. C. Hunt, 1999: Analysis and validation of GPS/MET radio occultation data in the ionosphere. *Radio Science*, **34**, 949–966, doi:10.1029/1999RS900034.

Shutts, G. J. and S. B. Vosper, 2011: Stratospheric gravity waves revealed in NWP model forecasts. *Quart. J. Roy. Meteor. Soc.*, **137**, 303–317, doi:10.1002/qj.763.

Smith, S. A., D. C. Fritts, and T. E. VanZandt, 1987: Evidence for a saturated spectrum of atmospheric graity waves. *J. Atmos. Sci.*, **44**, 1404–1410, doi:10.1175/1520-0469.

Steiner, A. K., G. Kirchengast, and H. P. Ladreiter, 1999: Inversion, error analysis, and validation of GPS/MET occultation data. **17**, 122–138, doi:10.1007/s005850050742.

Stockwell, R., L. Mansinha, and R. P. Lowe, 1996: Localization of the complex spectrum: the S-Transform. *IEEE Trans. Sig. Proc.*, **44**, 998–1001.

Stockwell, R. G., 1999: *"S-Transform Analysis of Gravity Wave Activity from a Small Scale Network of Airglow Imagers"*. Ph.D. thesis, "University of Western Ontario".

— 2007: A basis for efficient representation of the s-transform. *Digital Signal Processing*, **17**, 371 – 393, doi:http://dx.doi.org/10.1016/j.dsp.2006.04.006.

Tsuda, T., X. Lin, H. Hayashi, and Noersomadi, 2011: Analysis of vertical wave number spectrum of atmospheric gravity waves in the stratosphere using COSMIC GPS radio occultation data. *Atmos. Meas. Tech.*, **4**, 1627–1636, doi:10.5194/amt-4-1627-2011.

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**, 7257–7273, doi:10.1029/1999JD901005.

Vallis, G. K., 2006: *Atmospheric and Oceanic Fluid Dynamics*. Cambridge University Press, Cambridge, U.K., 745 pp.

VanZandt, T. E., 1985: A model for gravity wave spectra observed by Doppler sounding systems. *Radio Science*, **20**, 1323–1330, doi:10.1029/RS020i006p01323.

Wang, L. and M. J. Alexander, 2010: Global estimates of gravity wave parameters from GPS radio occultation temperature data. *J. Geophys. Res.*, **115**, D21122, doi:10.1029/2010JD013860.

Wang, L., M. J. Alexander, T. P. Bui, and M. J. Mahoney, 2006: Small-scale gravity waves in ER-2 MMS/MTP wind and temperature measurements during CRYSTAL-FACE. *Atmos. Chem. Phys.*, **6**, 1091–1104.

Watanabe, S., Y. Kawatani, Y. Tomikawa, K. Miyazaki, M. Takahashi, and K. Sato, 2008: General aspects of a T213L256 middle atmosphere general circulation model. *J. Geophys. Res.*, **113**, D12110, doi:10.1029/2008JD010026.

Wickert, J., C. Reigber, G. Beyerle, R. Knig, C. Marquardt, T. Schmidt, L. Grunwaldt, R. Galas, T. K. Meehan, W. G. Melbourne, and K. Hocke, 2001: Atmosphere sounding by gps radio occultation: First results from champ. *Geophy. Res. Lett.*, **28**, 3263–3266, doi:10.1029/2001GL013117.

Wright, C. J., 2010: *Detection Of Stratospheric Gravity Waves Using HIRDLS Data*. Ph.D. thesis, Trinity College, University of Oxford.

— 2012: A one-year seasonal analysis of martian gravity waves using MCS data. *Icarus*, **219**, 274–282, doi:10.1016/j.icarus.2012.03.004.

Wright, C. J. and J. C. Gille, 2011: Hirdls observations of gravity wave momentum fluxes over the monsoon regions. *J. Geophys. Res.*, **116**, n/a–n/a, doi:10.1029/2011JD015725, d12103.

— 2013: Detecting overlapping gravity waves using the S-Transform. *Geophy. Res. Lett.*, **40**, 1850–1855, doi:10.1002/grl.50378.

Wright, C. J., N. P. Hindley, and N. J. Mitchell, 2016: Combining airs and mls observations for three-dimensional gravity wave measurement. *Geophy. Res. Lett.*, **43**, 884–893, doi:10.1002/2015GL067233, 2015GL067233.

Wright, C. J., N. P. Hindley, A. C. Moss, and N. J. Mitchell, 2015a: Multi-instrument gravity-wave measurements over tierra del fuego and the drake passage - part 1: Potential energies and vertical wavelengths from airs, cosmic, hirdls, mls-aura, saamer, saber and radiosondes. *Atmos. Chem. Phys. Disc.*, **8**, 6797–6876, doi:10.5194/amtd-8-6797-2015.

Wright, C. J., S. M. Osprey, J. J. Barnett, L. J. Gray, and J. C. Gille, 2010: High Resolution Dynamics Limb Sounder measurements of gravity wave activity in the 2006 Arctic stratosphere. *J. Geophys. Res.*, **115**, doi:10.1029/2009JD011858.

Wright, C. J., S. M. Osprey, and J. C. Gille, 2013: Global observations of gravity wave intermittency and its impact on the observed momentum flux morphology. *J. Geophys. Res.*, **118**, 10,980–10,993, doi:10.1002/jgrd.50869.

— 2015b: Global distributions of overlapping gravity waves in HIRDLS data. *Atmos. Chem. Phys. Disc.*, 4333–4382, doi:10.5194/acpd-15-4333-2015.

Wu, D. L., 2004: Mesoscale gravity wave variances from amsu-a radiances. *Geophy. Res. Lett.*, **31**, 1944–8007, doi:10.1029/2004GL019562, l12114.

Wu, D. L. and J. H. Jiang, 2002: MLS observations of atmospheric gravity waves over Antarctica. *J. Geophys. Res.*, **107**, 4773, doi:10.1029/2002JD002390.

Wu, D. L. and J. W. Waters, 1996: Satellite observations of atmospheric variances: A possible indication of gravity waves. *Geophy. Res. Lett.*, **23**, 3631–3634, doi:10.1029/96GL02907.

Yan, A., W. Zhou, Q. Yuan, S. Yuan, Q. Wu, X. Zhao, and J. Wang, 2015: Automatic seizure detection using Stockwell transform and boosting algorithm for long-term EEG. *Epilepsy & Behavior*, **45**, 8–14, doi:10.1016/j.yebeh.2015.02.012.

Yan, X., N. Arnold, and J. Remedios, 2010: Global observations of gravity waves from High Resolution Dynamics Limb Sounder temperature measurements: A year-long record of temperature amplitude. *J. Geophys. Res.*, **115**, D10113, doi:10.1029/2008JD011511.