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
Centre for Space, Atmospheric and Oceanic Sciences, University Of Bath, BA2 7AY, UK

Correspondence to: N. P. Hindley (n.hindley@bath.ac.uk)

Received: 9 December 2014 – Published in Atmos. Chem. Phys. Discuss.: 3 February 2015
Revised: 19 June 2015 – Accepted: 23 June 2015 – Published: 16 July 2015

Abstract. Nearly all general circulation models significantly fail to reproduce the observed behaviour of the southern winter-time 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 radio occultation (GPS-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.

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 longstanding problem and has been identified as a serious impediment to model progress, leading to discrepancies in properties including simulated Antarctic ozone trends and southern-hemispheric 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 waves from other latitudes propagate into this latitude belt in the real atmosphere or because the sources of gravity waves in the models underrepresent 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-hemispheric 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, 2011; M. J. Alexander et al., 2008). Several second-order hot spots include South Georgia (M. J. 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; He et al., 2008; M. J. Alexander et al., 2008; S. P. Alexander et al., 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 GCMs exhibit strong disagreement in the magnitude and distribution of the flux of horizontal pseudomomentum (hereafter referred to as momentum flux) due to gravity waves in the southern-hemispheric 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, many gravity waves are sub-grid-scale phenomena and their effects must be parametrized. Parametrizations vary greatly among 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 (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 Sect. 2, we present maps and cross sections of gravity wave energy in the Southern Hemisphere, with implications for oblique focusing and leeward propagation of gravity waves into the southern stratospheric jet. In Sect. 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 Sect. 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 Sect. 5, and in Sect. 6 the key results of the present study are summarised.

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 Liu 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 and 30 km, while Tsuda et al. (2011)
verified multiple profiles with nearby radiosonde flights, returning discrepancies typically less than 0.5 K between 5 and 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. 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.

### 1.2 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 \( \Delta Z \) is primarily limited by Fresnel diffraction as

\[
\Delta Z \approx 2(\lambda L_T)^{\frac{1}{2}} \approx 1.4 \text{ km},
\]

where \( \lambda = 19 \text{ cm} \) is the GPS L1 wavelength and \( L_T \approx 28500 \text{ km} \) 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 \( \Delta L \) 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 \( \Delta Z \). By a first-order geometric argument, \( \Delta L \) and \( \Delta Z \) are approximately related as

\[
\Delta L = 2(2R\Delta Z)^{\frac{1}{2}},
\]

where \( R \) is the radius of the atmosphere at the tangent point. For a vertical resolution \( \Delta Z \approx 1.4 \text{ km} \) in the stratosphere, the horizontal line-of-sight resolution \( \Delta L \) is around 270 km. Gravity waves with \( \lambda_H \lesssim 270 \text{ km} \) 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 \( \lambda_H \) in the line of sight may be longer. This means that some waves with \( \lambda_H < 270 \text{ km} \) may be resolved. As discussed by S. P. 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 \( \lambda_H \) 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 Sect. 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.

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.
calculate $E_p$, we first interpolate each dry temperature profile $T(z)$ to 100 m 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 second-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 $\sim 2$ km cannot be reliably disassociated with noise in GPS-RO temperature profiles (Marquardt and Healy, 2005); therefore we apply a second-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 $\sim 4$ km. The transmission functions for each step in our analysis is shown in Fig. 6 and discussed further in Sect. 3.1. For the calculation of $E_p$ in this section, only the blue and green dashed lines in Fig. 6 apply.

This analysis provides a dynamic cut-off for vertical features in $T'(z)$. Features with vertical scales $\sim 3$–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

\[
E_p(z) = \frac{1}{2} \left( \frac{g}{N} \right)^2 \left( \frac{T'(z)}{\bar{T}(z)} \right)^2 ,
\]

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 (S. P. 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 S. P. 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, and 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.

### 2.1 Geographic distribution of \( E_p \) in the Southern Hemisphere

Figure 2 shows \( E_p \) in the Southern Hemisphere for each month in 2010 over the height interval 26–36 km. Note that this 10 km averaging window generally undersamples 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; S. P. Alexander et al., 2009). Between June and November, we see in Fig. 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. 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 (~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. Sato et al. (2012, their Fig. 2) showed maximum \( E_p \) directly over the mountains of the southern Andes at 10 hPa. Likewise, Yan et al. (2010, their Fig. 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 (~22–5 hPa) in Fig. 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 \( \lambda_Z \) waves directly over the mountains to the \( E_p \) distributions in Fig. 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. 2 is very reminiscent of southern-hemispheric 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.
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 focussing of gravity waves into the southern stratospheric jet in the Kanto GCM. In a meridional cross section from 30° S to 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 ~ 1500–2000 km southward over the height region 100 hPa (~ 16 km) to 1 hPa (~ 48 km). The flow appeared to focus towards the centre of the jet, where mean zonal winds were strongest.

In our Fig. 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. 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 to 50 km is evident in Fig. 3. The lower section of the column traverses nearly 1500 km southward over the height region 22–35 km. This suggests a clear focussing 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. 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 ~ 35 km the horizontal gradient in zonal-mean wind speed is low and waves appear to generally propagate upward without further latitudinal drift.

Figure 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$^{-1}$. Note the $E_p$ has been normalised at each height level to highlight the vertical structure.
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 Fig. 5) showed that waves with horizontal ground-based phase velocity $c_{\text{H}} = 0$ and $\lambda_{\text{H}} = 300$ km launched from the southern Andes could propagate eastward and southward by up to around 2500 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 to 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. 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

$$\lambda_{\text{Z}} \approx \frac{2\pi U_{||}}{N}, \quad (4)$$

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 $4\lambda_{\text{Z}} \gtrsim 13$ (Fig. 6). From Eq. (4), mountain waves could have vertical wavelengths too long to be detected for $U_{||} \lesssim 40–50$ m s$^{-1}$. 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. 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.

2.3 Vertical distribution of $E_p$ over the Southern Ocean

We also investigate the vertical distribution of wave energies over the Southern Ocean. Figure 4 shows altitude-normalised $E_p$ in a zonal cross section from 40 to 60° S during August 2010. As in Fig. 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 is the projection in the zonal domain of the vertical column evident in Fig. 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 (M. J. 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 timescales to be revealed in our analysis.

The column at 70° W appears to persist over the full height range in Fig. 4. However, between 25 and 35 km the largest values are observed well eastward between 60° E and 60° W. These peaks are located in a deep region of increased $E_p$ between 20 and 40 km and 30° W and 90° E, which is the projection in the vertical domain of the long leeward region of increased $E_p$ seen in Fig. 2.

At first glance, Fig. 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 energy 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 in situ instabilities and spontaneous adjustment around the stratospheric may also play a part.

Hoskins and Hodges (2005) presented a detailed view of southern-hemispheric 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. 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. 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. 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 is likely dominated by (1) primary orographic waves with $\lambda_H > 350$ 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.

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.

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. 5.

We begin by extracting temperature perturbations $T'(z)$ from each profile (Fig. 5a) as described in Sect. 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. 5b). The purpose of this step is to focus on the height range of the

---

**Figure 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$^{-1}$. Note the $E_p$ has been normalised at each height level to highlight the vertical structure.
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 and 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. 5c). Note that the term “energy” is defined as the sum of the squares 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 0, 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

\[
C(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} T'(z) \psi^* \left( \frac{z - b}{a} \right) dz, \tag{5}
\]

where \( T'(z) \) is our normalised, windowed and zero-averaged perturbation profile and \( \psi^* \) is the complex conjugate of the analysing wavelet. We choose an eighth-order complex Gaussian wavelet for \( \psi \), 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 \( \lambda \) by

\[
\lambda \approx \frac{a \Delta z}{f_{\text{cent}}}, \tag{6}
\]

where \( \Delta z \) is sampling interval and \( f_{\text{cent}} \) is the dominant central frequency of the wavelet 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 \( \Delta z \).

We are thus able to describe the spectral coefficients \( C(a, b) \) in terms of vertical wavelength \( \lambda_z \) and altitude \( z \). The absolute magnitudes of the spectral coefficients \(|C(\lambda_z, z)|\) are plotted in Fig. 5d.

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 \( \lambda_z \) 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 both the profile and the analysing wavelet are real, have RSS “energy” equal to 1 and are 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. 5d, 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 \( \lambda_z \) at position \( z \) with coefficient greater than 0.6. If this condition is satisfied, the identification is positive and we record the vertical wavelength \( \lambda_{\text{peak}} \) and altitude...

W. N. P. Hindley et al.: The southern stratospheric gravity wave hot spot

Figure 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.

\( z_{\text{peak}} \) at \( C_{\text{max}} \). As a result of the Gaussian windowing, \( z_{\text{peak}} \) is almost always located within one wavelength \( \lambda_{\text{peak}} \) 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.

\( C_{\text{max}} \) can thus be regarded as a confidence metric for the existence of wave-like features in COSMIC perturbation profiles. In the example in Fig. 5d, \( C_{\text{max}} \approx 0.64 \) 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. 5b, \( T' \approx 2.3 \) K.

To summarise our requirements for a positive wave identification, we require that the wave (1) has an amplitude \( 1 \) K < \( T' \) < \( 10 \) K, (2) has a vertical wavelength \( 2 \) km < \( \lambda_{\text{peak}} \) < \( 20 \) km, (3) is located such that \( 20 \) km < \( z_{\text{peak}} \) < \( 40 \) km and (4) has a confidence metric \( C_{\text{max}} \geq 0.6 \) as described above.

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. Figure 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. 6, the combined analysis method is generally sensitive to gravity waves with \( 4 \) km < \( \lambda_z < 13 \) km, with a sharp cut-off below \( 4 \) km and a more gradual cut-off above \( 13 \) km.

The histogram in Fig. 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 \( 7 \) km < \( \lambda < 9 \) 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 traditional 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 \( \lambda_z \) 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 \( \lambda_z \), which may help to explain the slight mismatch between the histogram in Fig. 6 and the range of permitted wavelengths (solid black curve) for longer \( \lambda_z \) 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).

3.2 Wave identification results

In Fig. 7a 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. 7b 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. 7a is a time-averaged climatology of $E_p$ in the region, whereas Fig. 7b 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. 7b than in Fig. 7a. This is expected, since mean $E_p$ values in Fig. 7b 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. 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. 7, information about wave intermittency can be inferred. The peak of the distribution of $E_p$ in Fig. 7b 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. 7a. 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, 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 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. 7b.
Figure 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.

At first glance, the histograms of wave amplitudes in each sector in Fig. 9 appear broadly similar. Approximately 20,000 waves are identified in each sector and the modal amplitude is between 2 and 3 K. Upon closer inspection, however, some important differences become apparent.

Despite containing around 4.5 and 12 % fewer profiles than Sector B, 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 also 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. 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 ($3 < T' < 8$ K) waves and fewer small amplitude waves ($T' < 2.5$ 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 $3 < T' < 8$ K than the corresponding upwind region, though these large amplitude waves are still relatively rare. Since $E_p \propto (T')^2$, it is hence likely that the structured distribution of $E_p$ in Fig. 7b is 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 wave 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. (2012).

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 Sect. 5.

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 is 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; M. J. 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 both closely spaced and closely timed with near-parallel lines inter-profile spacing. The scarcity of multiple profiles that are not routinely been applied to COSMIC due to typically large $k$.

Faber et al. (2013), limitations in sampling density, aliasing and differing lines of sight restrict the use of their approach 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 M. J. 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.

4.1 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 min 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 $\lambda_H \geq 270$ km in the line of sight as discussed in Sect. 1.2. 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 ($\pm 1.5$ km) is identified in both profiles using the Wave-ID method described in Sect. 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. A2a), time separations of less than 1 min 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 $\sim 75,000$ to $\sim 14,000$ 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 M. J. Alexander et al. (2008). We first apply a Gaussian window of FWHM = 22 km centred at 30 km altitude as described in Sect. 3.1. We next compute the CWT of each profile. The resulting transform $\hat{T}(z, \lambda_Z)$ is a complex valued function of altitude $z$ and vertical wavelength $\lambda_Z$. For the two profiles $a$ and $b$, the cospectrum $C_{a,b}$ is computed as

$$C_{a,b} = \hat{T}_a^{\ast} \hat{T}_b = \hat{T}_a \hat{T}_b \text{exp}(i \Delta \phi_{a,b}),$$

(7)

where $\hat{T}$ is the magnitude and $\Delta \phi_{a,b}$ is the phase difference between the two profiles for each $\lambda_Z$ 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 $\lambda_{DOM}$ common to both profiles at altitude $z_{DOM}$. We then compute the phase difference between the two profiles $\Delta \phi_{a,b}$ as

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

(8)

where $\text{Re}(C_{a,b})$ and $\text{Im}(C_{a,b})$ 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

\[
k_H = \frac{\Delta \phi_{a,b}}{\Delta \lambda_{a,b}},
\]

where \(\Delta \lambda_{a,b}\) 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 \(\lambda_H\) 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 \(100 \leq \lambda_H \leq 5000\) km to exclude unphysically short or extremely long horizontal wavelengths. This exclusion is discussed in more detail in Appendix A.

Generally, the horizontal separation \(\Delta r\) 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 \(\Delta r\) 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 \(T'_{a,b}(z)\) and take the mean.

Finally we compute an estimate of the absolute value of momentum flux \(M_{\text{flux}}\) as

\[
M_{\text{flux}} = \frac{\hat{\rho} \lambda_Z}{2 \lambda_H} \left( \frac{g}{N} \right)^2 \left( \frac{T'}{T} \right)^2,
\]

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

### 4.2 COSMIC momentum flux results

Figure 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 M. J. 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. 10a, 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. This region of longer vertical wavelength also extends further south over the Antarctic Peninsula.

The corresponding HIRDLS analysis in Fig. 10d shows typically longer \(\lambda_Z\) values overall due likely to the increased sensitivity of HIRDLS to waves with long \(\lambda_Z\) as a result of the larger usable height range in HIRDLS profiles. Like our COSMIC analysis, Fig. 10d 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 and 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 \(\lambda_H\) analysis from COSMIC profile pairs is presented in Fig. 10b. 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 \(\lambda_H\) along the axis between the two profiles, it is a upper-bound estimate heavily dependent on 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 \(\sim NE–SW\) or \(\sim NW–SE\) across this region, and hence estimates of \(\lambda_H\) 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 \(\lambda_H\) 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 \(\lambda_H\) 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 (\(~10\) km) between profile pairs than HIRDLS (\(~80\) km for an ascending–descending pair, \(~120\) km for a descending–ascending pair; e.g. Wright et al., 2015). This means that any absolute error in phase difference \(\Delta \phi_{a,b}\) between COSMIC pairs will bias our results towards shorter \(\lambda_H\) 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 $\lambda_H$ estimation with regards to horizontal separation is discussed more fully in Appendix A.

Figure 10c 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 (M. J. Alexander et al., 2008) and the maximum over the Antarctic Peninsula is in good agreement with results from the Vorcor super-pressure 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. 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 in Fig. 10f 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 Vorcor. 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 Sect. 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.

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 Sect. 2.3, Sato et al. (2012) suggested that waves from the mountains of the southern tip of South America and northern tip of the Antarctic Peninsula can propagate 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 waves rarely propagate east of the prime meridian, regardless of launch angle. Hence, the distribution of increased $E_p$ shown here eastwards of around 20$^\circ$ E is not likely to be explained by the downwind propagation of waves with $\lambda_H$ less than approximately 350 km. This suggests that the distribution of increased $E_p$ eastwards of around 20$^\circ$ E may be the result of (1) downwind propagating mountain waves with $\lambda_H > 350$ 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. 2 is morphologically reminiscent of southern-hemispheric 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 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 $\lambda_Z$ too short ($\lesssim3$ km) to be resolved by COSMIC below $\sim20$ 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 Figs. 2 and 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 Sect. 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$^\circ$ S could be greatly reduced in the Canadian Middle Atmosphere Model (CMAM) through the inclusion of non-specific orographic gravity wave drag in the stratosphere. One hypothesis for the missing drag is unparametrized mountain waves from small islands and around the Southern Ocean that are sub-grid-scale 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 focussing of waves in Fig. 3 into the jet around 60$^\circ$ 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 focussing is significant since many parametrization schemes used operationally in GCMs do not include such focussing phenomena (Preusse et al., 2014).

In Sect. 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 rare, 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. 7b 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 Sect. 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. 7b 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$^\circ$ 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. 7b 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$^\circ$ W–40$^\circ$ 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; M. J. 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 Appendix A.

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 Sect. 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 six low-inclination (24°) satellites.

6 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 to 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 $k_l$ 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 spaceborne 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.
Appendix A: On the determination of $\lambda_H$ 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 $\Delta r \sim 10$ km. The method for the estimation of $\lambda_H$ described in Sect. 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 $\lambda_H$ estimates to estimates from other studies as in Sect. 4.2, it is difficult to independently quantify the error and reliability of these estimates. This Appendix discusses the effect of short horizontal separations of COSMIC profile pairs on the estimation of $\lambda_H$ in comparison to profile pairs from the HIRDLS mission.

A1 Horizontal profile pair separations

An estimate of horizontal wavelength $\lambda_H$ can be calculated from the horizontal separation and phase difference between two adjacent profile pairs via the relation in Eq. (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. (2015, 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 $\sim 80$ and $\sim 120$ km respectively (see blue bars in our Fig. A2a).

Figure A1 shows histograms of gravity wave phase differences between COSMIC, HIRDLS 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 $s \leq 6$. HIRDLS profile pairs are processed using the method described by Wright and Gille (2013).

All three horizontal separations in Fig. A1 indicate a general preference towards small ($\Delta \phi < \frac{\pi}{2}$) 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. A1 indicates that COSMIC pairs with $\Delta r \sim 10$ km generally have more small ($\Delta \phi < \frac{\pi}{2}$) and fewer large ($\Delta \phi > \frac{\pi}{2}$) 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 $\lambda_H$ estimation methodology, particularly its application to COSMIC profile pairs.

A2 Biases from small phase differences

Even if the methodology is valid for horizontal separations as short as $\sim 10$ km, errors 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 $\pm 0.1 (\sim \frac{\pi}{20})$ radians, then absolute phase differences of $0 \leq \Delta \phi \leq \frac{\pi}{20}$ 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, $2 \times \Delta r \approx 20$ km. However, the requirement of $\lambda_H \gtrsim 270$ 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 $\lambda_H$ estimates are unlikely to be physical. A cut-off of $\lambda_H \gtrsim 100$ 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 $\pm \frac{\pi}{30}$, $\lambda_H$ estimates from COSMIC profile pairs with $\Delta r \sim 10$ km may be accurate for $100 \lesssim \lambda_H \lesssim 600$ km. For larger horizontal separations and/or more accurate phase difference determinations, the upper limit is larger. Figure A2d 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 Sect. 3. Dashed black lines show lines of constant $\lambda_H$ estimated via the relation in Eq. (9). For the majority of detected waves $250 \lesssim \lambda_H \lesssim 5000$ km. A low-bias effect on the estimation of $\lambda_H$ for short horizontal separations $\Delta r \sim 10$ 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 Sect. 4 are likely to introduce a low bias in the estimation of $\lambda_H$ due to error in the determination of $\Delta \phi$. We suspect that this is the reason for the differences in $\lambda_H$ estimates between COSMIC and HIRDLS in Fig. 10e. 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.
Acknowledgements. This study is supported by the UK Natural Environment Research Council (NERC): C. J. Wright and N. J. Mitchell are supported by NERC grant NE/K015117/1 and N. P. Hindley by a PhD studentship awarded to the University of Bath. We are grateful to Andrew Moss for helpful suggestions and discussions. We would also like to acknowledge Harrower and Brewer (2003) for the ColorBrewer colour maps featured in this study.

Edited by: F. Fierli

References


Yan, X., Arnold, N., and Remedios, J.: Global observations of gravity waves from High Resolution Dynamics Limb