Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 56 (2015) 619–633

ADVANCES IN SPACE RESEARCH (a COSPAR publication)

www.elsevier.com/locate/asr

Distribution functions and statistical parameters that may be used to characterize limb sounders gravity wave climatologies in the stratosphere

P. Alexander [a,*], D. Luna [b], A. de la Torre [c], T. Schmidt [d]

[a] Instituto de Física de Buenos Aires, CONICET, Ciudad Universitaria Pabellón 1, 1428 Buenos Aires, Argentina
[b] Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria Pabellón 1, 1428 Buenos Aires, Argentina
[c] Facultad de Ingeniería, Universidad Austral, Av. J. de Garay 125, 1063 Buenos Aires, Argentina
[d] Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Potsdam, Germany

Received 10 October 2014; received in revised form 6 May 2015; accepted 7 May 2015
Available online 14 May 2015

Abstract

The number of gravity wave (GW) activity climatologies in the stratosphere started to increase more than 10 years ago since the appearance of large amounts of limb and nadir satellite sounders data. There have been very few discussions regarding the adequate statistical description of GW activity in terms of a distribution function and its parameters. We put forward the question whether a general statistical functional representation adaptable to the characteristics of GW activity in diverse geographic regions and seasons exists. Here we approach this issue for two different types of limb sounders and in particular we try to find out which parameters may represent at best the climatological features. We study results for a region close to the Patagonian Andes and their prolongation in the Antarctic Peninsula, which is well-known for the generation by topography of intense stratospheric GW, specially during winter and spring. Global Positioning System (GPS) radio occultation (RO) records presently provide over 2000 profiles per day. We used 5 years of COSMIC (Constellation Observing System for Meteorology, Ionosphere and Climate) mission GPS RO data, which supplied almost 150,000 retrievals for our study. Three different distribution functions have been approached to describe the GW activity climatologies: gaussian, log-normal and gamma. The latter function has not been used in previous work. It has been shown here that it is a competitive option to the log-normal distribution. In addition, its use allows not only to quantify the GW activity level of each climatology in the stratosphere, but also to find out the number of significant modes that essentially determine it. Alternative parameters to the mean like the median may be used to characterize the climatologies. The use of the median may exhibit advantages in cases where the presence of spurious large GW activity measurements are suspected in GPS RO data. The mean is equally suitable to establish GW activity comparisons. As a priori we may not know if the above mentioned artifacts are present, in general it may be more appropriate to use the median. We perform a similar general study for data from the SABER (Sounding of the Atmosphere using Broadband Emission Radiometry) limb instrument, as it is presently also used to obtain global GW climatologies in the stratosphere. Although the observational window and data processing features are not identical for both instruments, the results exhibit many similarities.

Keywords: Gravity waves; Limb sounders; Statistical descriptions

* Corresponding author. Tel.: +54 11 4576 3353; fax: +54 11 4576 3357.
E-mail addresses: peter@df.uba.ar (P. Alexander), luna@df.uba.ar (D. Luna), adelaTorre@austral.edu.ar (A. de la Torre), tschmidt@gfz-potsdam.de (T. Schmidt).

http://dx.doi.org/10.1016/j.asr.2015.05.007
0273-1177/

1. Introduction

About 20 years ago it was possible to approach gravity wave (GW) activity climatologies for a limited amount of soundings in restricted areas and times (e.g., Allen and Vincent, 1995). Recent advances in satellite remote sensing technology have begun to provide valuable information on GW, so global and regional studies in the stratosphere have spread. Satellite data acquired using limb and nadir techniques have demonstrated great potential for studying GW sources and their propagation (e.g., Wu et al., 2006), but not all instruments can optimally contribute to the construction of GW climatologies. Some of them have no reliable records for several years or they provide a fair number of profiles per day or they exhibit a poor coverage over some geographical areas.

The Global Positioning System (GPS) radio occultation (RO) technique is able to provide huge amounts of retrievals (e.g., Baumgaertner and McDonald, 2007; Alexander et al., 2010b). A GPS RO occurs whenever a transmitting satellite from the global navigation network at an altitude about 20,000 km rises or sets from the standpoint of a low Earth orbit (LEO) receiving satellite at a height of about 800 km and the signal traverses the atmospheric limb. The Doppler frequency alteration produced through the refraction of the ray by the Earth's atmosphere in the trajectory between the transmitter and the receiver is detected and then may be converted through a sequence of established procedures into profiles of diverse variables in the neutral atmosphere and ionosphere (see e.g., Kursinski et al., 1997). The technique has a typical measurement time in the neutral atmosphere of 1 min (as compared to the much longer dynamical processes), it has global coverage, all-weather and all-time capability, no recurrent need of instrumental drift or bias corrections (observations are stable in the long-term due to permanent self-calibration) and it exhibits a vertical resolution of about 1 km in the stratosphere (see e.g., Kursinski et al., 1997; Kursinski et al., 2000; Marquardt and Healy, 2005).

In April 2006, the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) launched six LEO satellites. The aim of the mission was to produce up to 2500 GPS RO daily with global distribution. It is still obtaining real time and post-processed vertical profiles of temperature, pressure, refractivity, and water vapor in the neutral atmosphere and electron density in the ionosphere. It has generated an extensive amount of data to support operational global weather prediction and climate change monitoring, and also to improve the study of neutral atmosphere and ionospheric phenomena, space weather, and possible relations between meteorological or ionospheric processes with solar activity (e.g., Liou et al., 2007). The number of the daily retrievals is presently exhibiting a declining trend, as the expected lifetime has already been outranged. This study uses post-processed data (product version 2010.2640) from the COSMIC mission provided by CDAAC (COSMIC Data Analysis and Archive Center). We use the atmospheric vertical profiles of temperature to construct the GW activity climatologies as explained in the following section.

We perform a similar general study also for data from the SABER (Sounding of the Atmosphere using Broadband Emission Radiometry) instrument, as it is also used to obtain global GW climatologies in the stratosphere. The SABER instrument (e.g., Mlynczak, 1997) is an infrared emission limb sounder covering the upper troposphere, stratosphere, mesosphere and lower thermosphere and is on board of the Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) satellite. Here we use temperatures from version 2.0 retrievals. SABER and GPS RO share a similar observational sensitivity regarding the horizontal and vertical resolutions, but the latter technique has a slightly better performance (e.g., Sivakumar et al., 2011; Wright et al., 2011).

GW activity is usually quantified in each individual event by calculating the mean potential energy per unit mass E_p or nearly equivalently the mean temperature perturbation amplitudes or variances. The momentum flux may be also calculated in the presence of a sufficiently high density of retrievals (e.g., Wang and Alexander, 2010; Faber et al., 2013). In general, when characterizing climatologies through E_p, the most significant statistical parameter usually seems to be the arithmetic mean (e.g., von Storch and Zwiers, 2003). Its wide use is probably related to the fact that it is easy to calculate and due to the law of large numbers we know its functional distribution and dispersion. An implicit assumption which in some processes may not be applicable is that the observations exhibit distribution symmetry. In addition, the mean may not be an adequate representation of the typical state, as for example in a scenario described by a distribution with two far away Gaussian peaks. Moreover, in the analysis of phenomena that exhibit similarities to the normal distribution but clearly are not, alternative parameters to the mean and its range of variability may be considered.

Distributions based on symmetry may have more appeal than skewed ones. The normal shape is often assumed to describe the distribution of the occurrence of diverse processes. The curve can easily be described by two values: the mean and the standard deviation. However, many cases show skewed distributions, which sometimes emerge when the observed variable cannot be negative (e.g. energy). In some cases the variability around the mean may be clearly asymmetrical, for example subtracting one standard deviation from the mean produces a negative value (see below). Such skewed curves sometimes fit the log-normal shape. Log-normal distributions are usually characterized in terms of the log-transformed variable (Johnson et al., 1994). Specifically, a random variable X is said to be log-normally distributed if $\log(X)$ is normally distributed. The distribution is skewed to the left. Two parameters are needed to specify it: the mean m and the standard

deviation s of $\log(X)$. The median of X is e^m. A distinct aspect of the log-normal distribution as compared to the normal counterpart is that uncertainties could be expressed in an asymmetrical way and a remarkable feature of the median is that it is much less affected by outliers than the mean.

Preusse et al. (2000) compared the results from four different satellite instruments at a very early stage of global GW climatologies. Wu et al. (2006) made a very comprehensive review of limb and nadir sounders used to observe the GW. They recalled the instruments sensitivity limits in terms of vertical and horizontal wavelengths. McDonald and Hertzog (2008) mentioned previous work showing that the observational filter of each technique determines the portion of the full spectrum of GW that it may observe and they recalled in particular the effect in GPS RO observations. Preusse et al. (2008) showed in the horizontal and vertical wavelengths plane the sensitivity ranges covered by the diverse limb and nadir instruments, whereas Alexander et al. (2010a) also included the visibility limits in frequency. GPS RO has typically the best visibility when GW have wavelengths between a few and several km in the vertical direction, a few and several 100 km in the horizontal direction and periods above one hour. Hertzog et al. (2012) used various data sets to study the intermittency of GW in the lower stratosphere above Antarctica and the surrounding area. The obtained GW momentum flux distribution functions consistently showed long tails related with the occurrence of sporadic and large-amplitude cases. The tails were longest above mountains. It was found that these sporadic events represented the main contribution to the total momentum flux during winter, but its relevance significantly decreased in late spring and summer. It was also shown that, except above mountainous areas in winter, the momentum-flux distribution functions tended to look log-normal. Previously, Baumgaertner and McDonald (2007) used the log-normal distribution function to fit GW potential energy spectra above Antarctica. The log-normal distribution has appeared several other times when modeling observations of diverse quantities associated to GW (e.g., Nastrom and Gage, 1985; Plougonven et al., 2013).

Some problems appear in the analysis of skewed processes. On one hand, skewed distributions produce large values that may appear to be outliers. Such observations could be eliminated and the analysis would then be performed without them, thereby making the observations to look more symmetrical, but at the expense of introducing unknown bias. On the other hand, skewed data are often grouped together and averaged, whereby the mean is normally distributed (law of large numbers) and used for further analyzes. By following this way, important features of the data may remain hidden. This is one aspect we will explore below regarding GW climatologies. Alexander et al. (2009) used medians instead of the mean to avoid skewing COSMIC GW temperature variances by a few very large values. We are not aware of any specific study assessing the pertinence of the mean to describe E_p according to its distribution. We would like to take this challenge and analyze the observed distributions of E_p and thereafter consider if alternative parameters should be used to describe it. We should stress that there are for the moment no specific explanations of the observed skewed E_p distribution shape in terms of basic physical concepts or of the observational window or errors of the sounding technique. We therefore search below from basic physical and statistical principles for a distribution model that suits these expectations. In addition, the new distribution function will lead to two statistical parameters, that may be interpreted in terms of elementary GW activity characteristics.

In general, the median as an estimator has the significant advantages that (i) it is more robust in the presence of outliers and (ii) it better characterizes skewed distributions than the mean. Regarding (i) it makes this estimator less sensible to large E_p fictitious cases due to GPS RO profile initialization problems (see e.g., Ao et al., 2006), which usually show up as several Kelvin temperature oscillations above about 25 km height and typically correspond in our work to events about $E_p > 4$ J/kg. In reference to (ii) the median may be understood as percentile 50, regardless of the distribution function (normal, log-normal or any other possible choice). On the contrary, the interpretation of the mean is dependent on the distribution function, for example in bimodal distributions its meaning is not straightforward.

Section 2 describes the GPS RO and SABER data and their transformation into GW activity climatologies. In Section 3 we evaluate the adequateness of the normal and log-normal distribution functions and the mean and the median to describe GW activity with GPS RO data and the possible impact in these results of potential outliers due to sporadic retrieval artifacts. We also argue in favor of a new model to represent the E_p distribution function. In Section 4 we repeat calculations for SABER data. Conclusions are drawn in Section 5.

2. Data and analysis method

A vast range of satellite, balloon and numerical modeling studies have shown that there is a zone in the Southern Hemisphere, close to the Patagonian Andes and their prolongation in the Antarctic Peninsula, that exhibits a strong GW activity in the lower stratosphere in winter and spring (see e.g., Baumgaertner and McDonald, 2007; Alexander et al., 2010b; Sato et al., 2012; Plougonven et al., 2013; Jiang et al., 2013), mainly due to topography (see e.g., Alexander and Teitelbaum, 2011). We take advantage of this characteristic to perform our study in the area 90–30W and 30–90S. The region was divided in 12 zones to study any geographical variations: 10° latitudinal bands to the West and to the East of the rough longitude of the highest topography. In Fig. 1 we show the studied region and the bands.

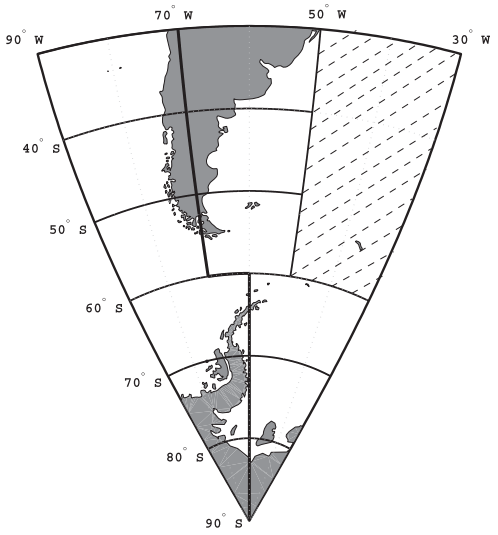

Fig. 1. The studied region divided in 10° latitude bands to the West and East of 70W at 60–30S (50–30W excluded) and of 60W at 90–60S.

We used COSMIC GPS RO temperature profiles between day 200 of year 2006 and day 200 of year 2011. The 5 years supplied 149,715 retrievals. In the region studied here Alexander et al. (2009) found that about 41% of COSMIC RO have a LOS within 30° of the North–South direction, which is a value somewhat lower to the 52% for the CHAMP (Challenging Minisatellite Payload) LEO satellite. This preferential meridional alignment of the LOS is advantageous regarding the detection of GW originated by the topography in this region, as it is expected that this fact will favor the observation of the searched signatures (see e.g., Baumgaertner and McDonald, 2007).

The mean potential energy per unit mass E_p was here calculated through the average relative temperature T variance between altitudes z_1 and z_2 in each profile (Wilson et al., 1991)

$$E_p = \frac{1}{z_2 - z_1} \int_{z_1}^{z_2} \frac{1}{2} \left(\frac{g}{N}\right)^2 \left(\frac{T'}{T_b}\right)^2 dz \quad (1)$$

where g is the gravitational acceleration and N represents the Brunt-Väisälä frequency, which may be derived from each background temperature profile (T_b). Notice that E_p is a positive quantity. The ratio of perturbation and background temperatures T'/T_b was obtained as follows in each case. The T profiles were low pass filtered, with a cutoff at 15 km, obtaining T_b. The filter applied is nonrecursive and a Kaiser window was used (e.g., Hamming, 1998). The filter was applied again to the difference $T - T_b$, now with a cutoff at 2 km, giving T' profiles, which isolate wavelengths between 2 and 15 km. GW in this region typically belong to this range (e.g., de la Torre and Alexander, 2005; Alexander et al., 2008b), whereby smaller wavelengths cannot be detected due to noise and processing of the original raw GPS RO data (e.g., Marquardt and Healy, 2005). The z_1 to z_2 vertical column for the integral was 15–33 km (i.e. lower stratosphere in this region). Although this procedure does not actively remove planetary waves, it has been used in some works as the effects due to GW are generally much larger than those due to planetary waves (McDonald et al., 2010; Alexander et al., 2011). We applied the lapse rate definition of the World Meteorological Organization to identify the tropopause height in each retrieval. Profiles with the tropopause above 15 km were discarded. Although tropospheric data are available, it is standard use to study only the stratosphere because a problem arises around the tropopause, as the sharp change in temperature gradient sign leads to an artificial enhancement in wave activity when using digital filters to isolate the components (e.g., de la Torre et al., 2010). In addition, the much lower amounts of water vapor in the stratosphere imply a much better quality for the temperature retrievals than in the troposphere. The vertical resolution of the temperature profiles typically ranges between 0.5 and 1.4 km from the troposphere to the stratosphere (see e.g., Kursinski et al., 1997).

We then repeated similar calculations with SABER temperature profiles for the same time period. We obtained for each profile the E_p value between 20 km and 42 km height using the same band-pass filter but with cutoffs at 2 and 20 km. GPS RO profiles were limited to a maximum of 33 km height due to the their low reliability at higher altitudes (Luna et al., 2013). The larger height interval allows for SABER profiles the possible inclusion in the analysis of longer vertical wavelengths. This fact and the higher top altitude of the analyzed SABER retrievals (it allows a further growth of wave amplitude with height) lead to an expectation of larger E_p for data obtained from this instrument, and therefore also the corresponding medians and means.

3. Results with GPS RO data

In Fig. 2 we show the number of GPS RO soundings per band and month. Variations against band are more remarkable than per month, where the former are related to the combined effect of growing band surfaces towards lower latitudes and the preferential occurrence of GPS RO at middle latitudes due to the orbital characteristics of the COSMIC satellites. There are no order of magnitude variations over time per zones. Then, we may infer that there were no significant drop-offs in the number of GPS RO due to malfunction over the period studied, so no biases should be expected in certain months or seasons due to this potential problem.

In Fig. 3 we show the corresponding E_p means and medians. The maximum levels of activity are found during austral winter and spring. The effect is most visible at the highest latitudes for both quantities. In the latitudes 40–60S there are local minima about June due to the appearance of an April surge. Notice that the highest latitudes exhibit the lowest values during summer and autumn. The median values are always lower than the mean values

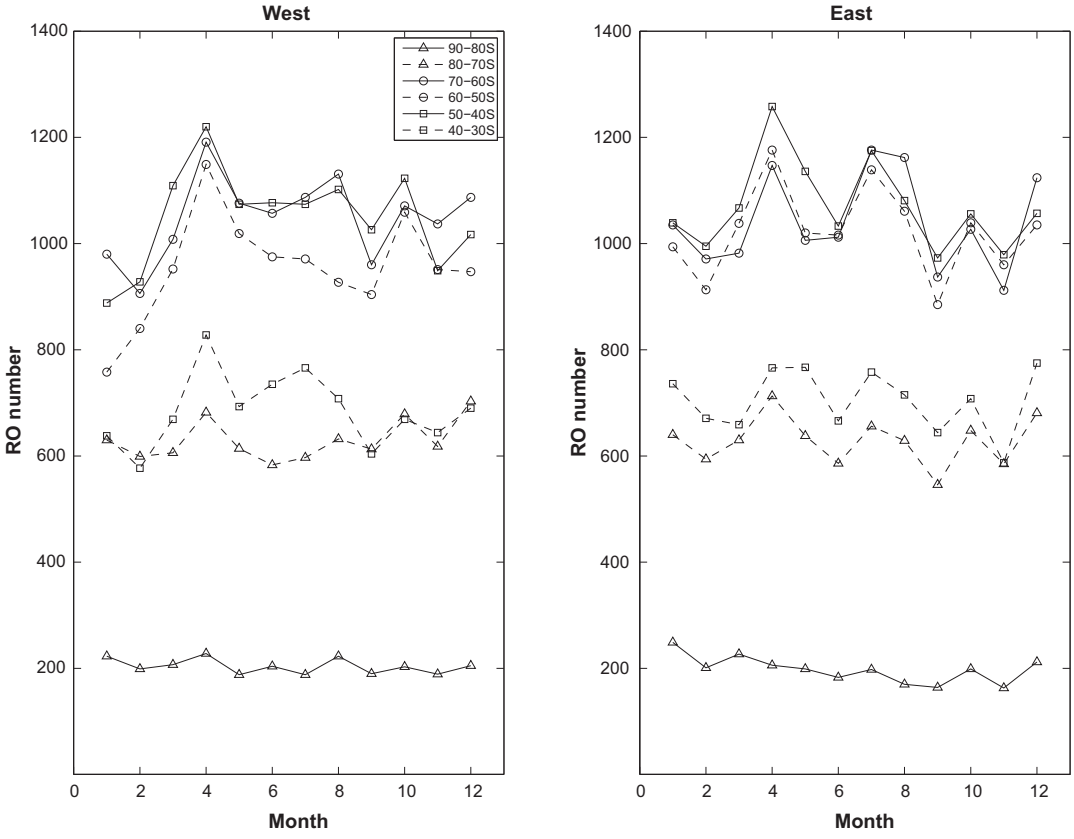

Fig. 2. Amount of GPS RO per band against month number (1 is January and so on) between day 200 of year 2006 and day 200 of year 2011.

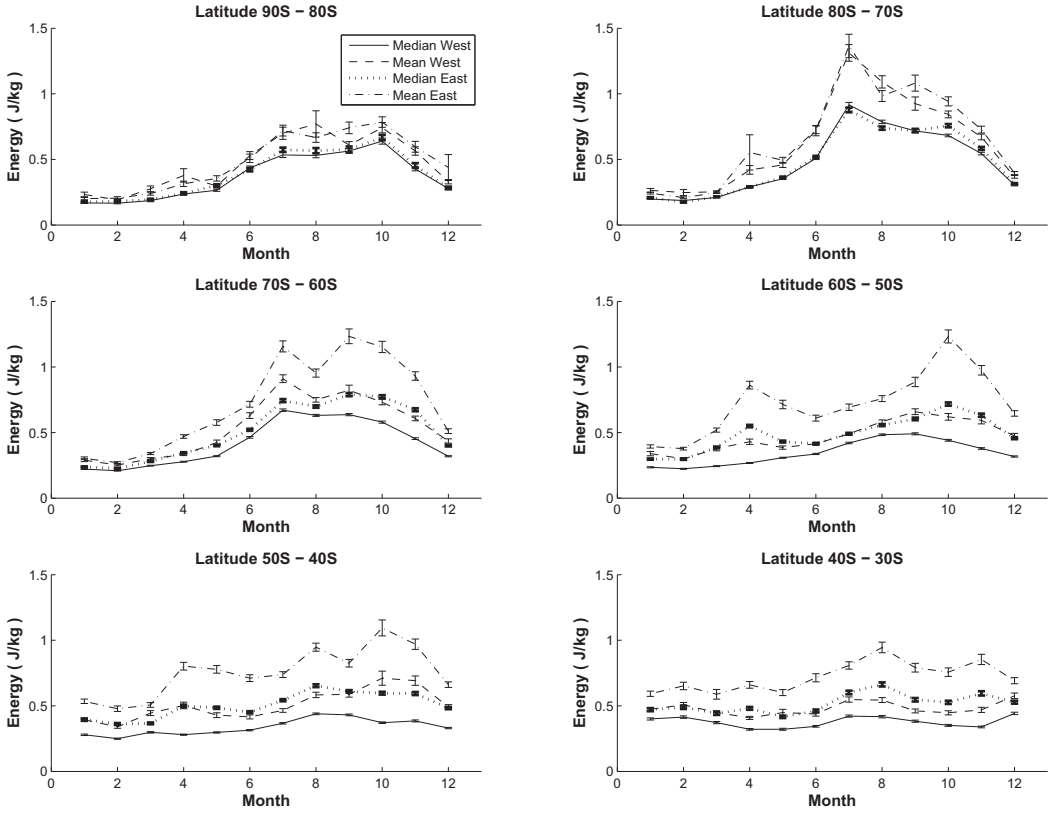

Fig. 3. Mean and median E_p for GPS RO data and their error bars per zones against month number.

(the distribution functions are always skewed to the left), but they both follow the same trends. In addition for both parameters, values to the East are nearly systematically larger than to the West, which confirms the significant role of topography for GW generation in the studied region.

As mentioned above, there is a significant increase of activity around April between latitudes 40–60S in Fig. 3. A consistent April enhancement in E_p was also seen in this region using radiosondes by Moffat-Griffin et al. (2013). A possible explanation for this surge may be the presence of an additional source of GW besides topography, which is flow imbalance adjustment (Uccellini and Koch, 1987; Fritts and Alexander, 2003; Sato et al., 2012). Possible generation of GW by jet streams was extensively studied experimentally and theoretically by Gavrilov and Fukao (1999a,b, 2001). Jet streams may generate downward and upward propagating modes, whereby inertia-GW are the typical byproduct of flow imbalance adjustments and they belong to the optimal observational range of RO (Alexander et al., 2008a). With GPS RO data de la Torre et al. (2006) detected inertia-GW at jet levels above the Southern Hemisphere mid-latitude Andes mountains, which were generated by flow imbalance adjustment while the atmosphere was trying to restore equilibrium. The effect waned with increasing horizontal distance to the mountains. In Fig. 4 we show NCEP (National Center for Environmental Prediction) reanalysis zonal wind climatological (1981–2010) data U at 250 mb. In the studied region about early autumn the subtropical jet stream changes its nearly stable (for some months) latitude from about 50S to 35S and it is therefore a possible flow imbalance source.

One of the possible alternatives to present GW activity climatologies is through an average of E_p taken over a given region in a determined time interval or season (in some cases along various years), as shown e.g. by Baumgaertner and McDonald (2007) and John and Kumar (2012). A normal distribution representation of E_p would be expected to be centered at some positive value (as defined above it must be positive). In several cases in our work the standard deviation of E_p in a given zone and month exceeded the mean value and the distribution function exhibited clear skewness. These aspects may be visualized in the histogram example in Fig. 5 for October over the 5 years. A non-symmetrical function as the log-normal distribution seems then to be a more adequate representation (see e.g., Baumgaertner and McDonald, 2007) regardless of the cause of the shape, which could be the true wave characteristics or the limitations of the measurement technique. The log-normal distribution has been used because heuristically it provides a good statistical representation of E_p, but a physical substantiation of this aptitude has not been given. Therefore, below we use χ^2 goodness-of-fit tests to assess both hypotheses for the

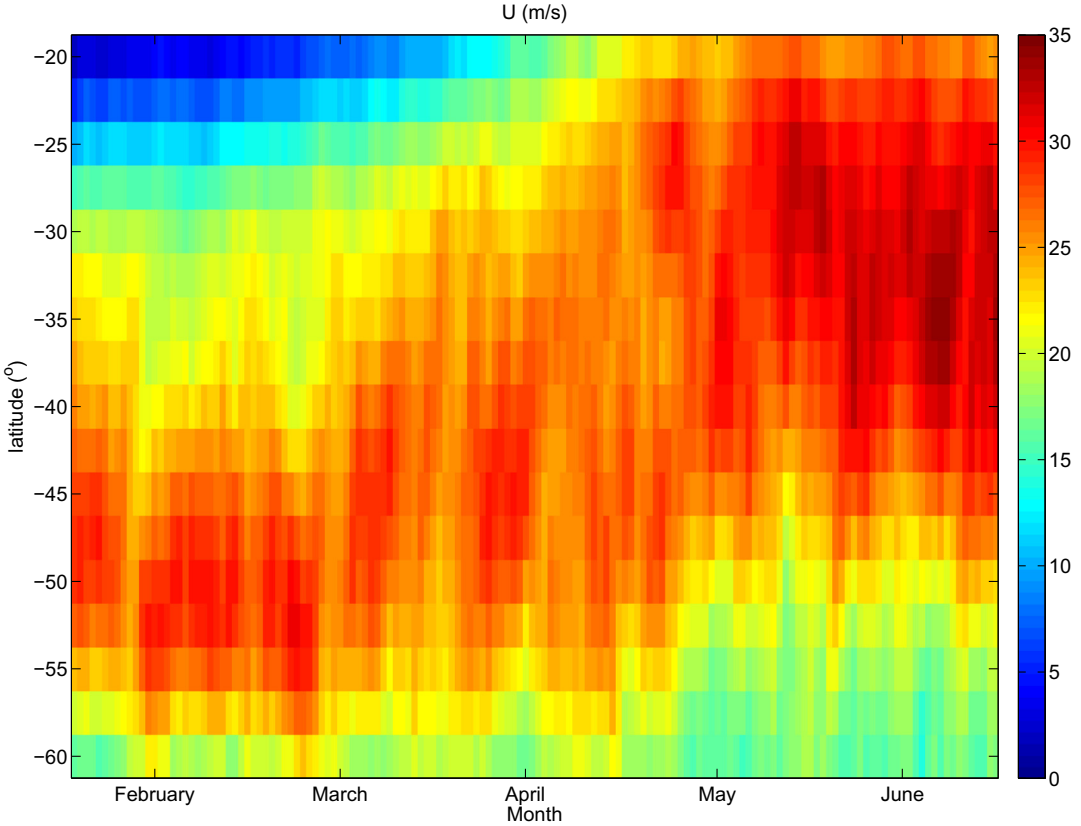

Fig. 4. Zonal wind in the studied region at 250 mb from climatological reanalysis NCEP data around April, where the subtropical jet stream latitude against time may be identified.

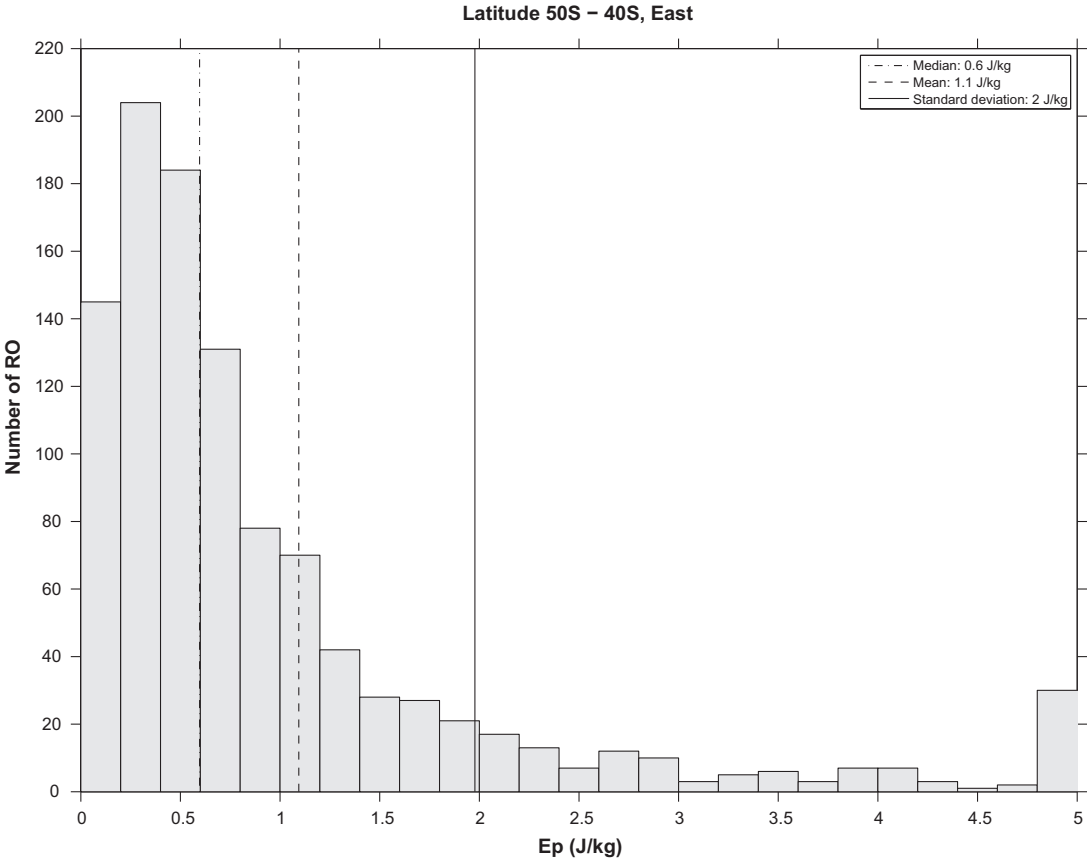

Fig. 5. A E_p histogram for the soundings in latitude band 30–40S to the East of 70W during October of the 5 years studied.

distribution function. This kind of test is used to compare a given sample against a reference probability density distribution function, whereby the significance level (we chose 5%) will lead in each case to rejection or insufficient evidence for rejection (which is not the same as acceptance) of the null hypothesis (the sample resembles the reference distribution). We performed the tests for the 12 zones and the 12 months (144 samples).

Before doing the tests we explore if another distribution function might be appropriate, trying to give a physical interpretation to the observed skewed statistical distribution of E_p. The relative temperature perturbation amplitudes of the modes are all positive, many of them probably having small values, whereas it is physically expected that they tend to disappear for very large values as energy cannot tend to infinite. We might model this general behavior by a half-normal distribution. We notice that we could observe positive or negative perturbations at different measurement points, but we are now focusing on the amplitude of every mode, which is the governing factor in the calculation of the potential energy in a given zone. The measured fluctuations are the sum of the positive and negative perturbation values of the modes at every observational point. As defined in Eq. (1) and when using any technique to decompose the perturbations into orthogonal modes, E_p should be considered proportional to the sum of the square of the corresponding amplitudes. We may perform our calculation over an integer number of sinusoidal cycles if we assume a statistical compensation between excess and defect contributions of modes in a real ensemble, so

$$E_p = \frac{1}{2}\left(\frac{g}{N}\right)^2 \frac{\sum_i X_i^2}{2} \tag{2}$$

where X_i are the relative temperature perturbation amplitudes, which we conjecture to follow a half-normal distribution. It is then expected that E_p should resemble a Γ distribution (see Appendix A). By optimizing the fit for any data-set we may obtain two parameters: (i) the typical number of significant modes that determine the E_p calculation and (ii) a mean value of the half-normally distributed relative temperature perturbation amplitudes. Regarding the first parameter, we recall that the application of standard spectral techniques to detect the GW present in a measured atmospheric profile often lead to the identification of a few modes that have significantly larger amplitudes than the rest (see e.g., de la Torre et al., 1994).

From the 144 samples analyzed, the normal hypothesis became rejected in 67 cases (47%). Theses results rule out the Gaussian description as a faith model for the E_p distribution. The results do not exhibit a specific dependence on the latitudinal band or month. The log-normal hypothesis was rejected in no one of the samples, whereas the gamma distribution was rejected in just 6 samples (4%). In Fig. 6

we show the results derived from the physical interpretation of the Γ distribution for all months and zones. Regarding the second Γ model parameter, the relative value has been converted into an absolute measure after multiplying by 220 K, which is a typical stratospheric temperature. We may notice that the number n of prominent modes range between 2 and 6 and the perturbation mean amplitudes μ between 0.2 and 0.7 K. The only latitude band with clearly more modes on the Eastern side is 90–80S, whereas the opposite happens in 70–60S. However, it must be recalled that the 90–80S band has less data than the other regions because it is much smaller and due to the GPS RO geographical coverage characteristics. The other regions have a mixed behavior throughout the year. In the zone 70–30S there is a surplus of the Eastern amplitudes compared to the Western ones throughout the year, as expected from topographic GW generation in the Antarctic Peninsula and continental Andes, mainly between late autumn and spring (see e.g., Eckermann and Preusse, 1999; Alexander et al., 2008b; Hei et al., 2008; Alexander et al., 2009; Yan et al., 2010; Sato et al., 2012).

We further spatially zoom in the GW activity and divide the studied area in cells $3° \times 6°$ and zoom out the time interval over 3 months, whereby the mean and median energy are calculated for two seasons which exhibit different behavior regarding E_p: autumn (March, April and May) and spring (September, October and November). The size of the cells is determined by the achievement of a sufficient number of cases to perform statistical calculations in each of them and the period is determined by the similarity of E_p values over time. In Fig. 7 we show the mean, the median and the number of cases per cell for autumn. GW activity is high to the East of the Patagonian Andes. In Fig. 8 we show the results for spring. This is the season with largest activity, specially to the East of the Andes and their prolongation in the Antarctic Peninsula. There is a somewhat smaller activity above the Drake Passage. This reinforces the idea of the topographic origin of the waves: although the absence of mountains would imply null activity in the 55–65S latitude range, some of the neighboring North and South intense activity may be leaking into the zone.

An intense GW activity may be noticed in Figs. 7 and 8 to the East of the Patagonian Andes in both seasons, but only during spring to the East of the Antarctic Peninsula. We would like to establish comparisons by using the mean $\bar{E}_p$ and the median $\hat{E}_p$ and their corresponding uncertainties. We defined the former zone by 35–55S and 70–60W and the latter by 60–80S and 60–30W to perform our calculations. The mean uncertainty was evaluated by the standard error of the mean. The median uncertainty was assessed by two different methods (Sheather and Maritz, 1983; Ratel, 2006), which led to the same result for the most significant digit, which is exhibited in each case.

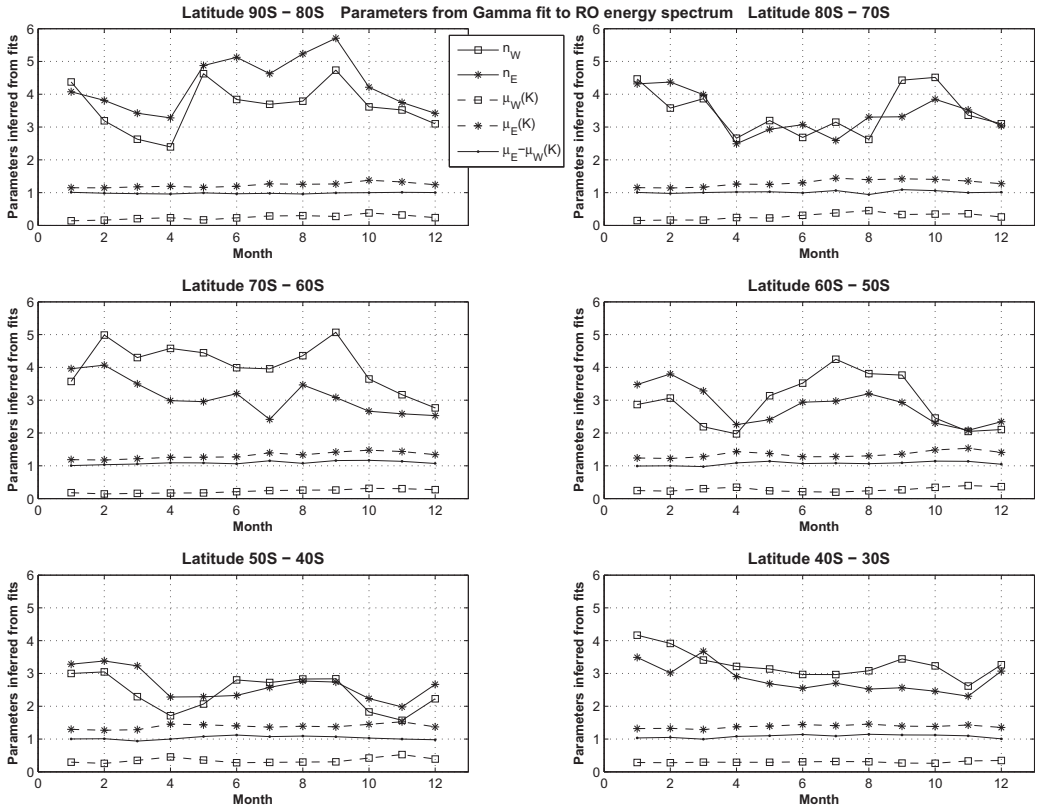

Fig. 6. Quantities derived from fits of E_p to a Γ distribution. Notice that μ_E has been vertically shifted by 1 unit to avoid the superposition with μ_W. The difference $\mu_E - \mu_W$ is shown centered at 1. Subscripts W and E refer to roughly the West and East of Andes.

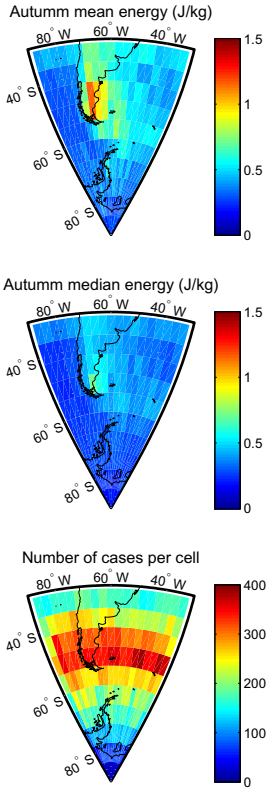

Fig. 7. The mean and the median E_p and the number of GPS RO per cell in the analyzed zone for autumn over the 5 years studied.

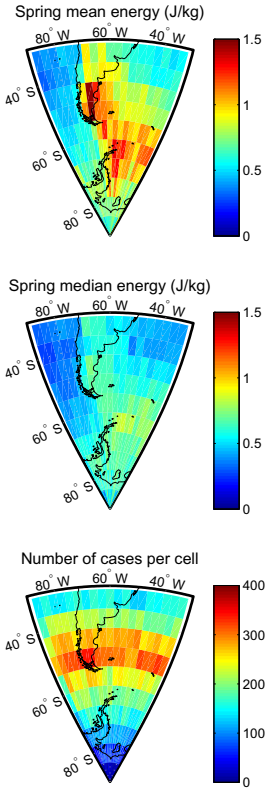

Fig. 8. The same as Fig. 7 for spring.

The former work estimates the uncertainty by $1/(2k^{1/2}f(m))$ where k is the amount of sample values and $f(m)$ the density estimator evaluated at the sample median, whereas the latter publication uses for the calculation $1.858M/\sqrt{k-1}$, where M is the median of the absolute deviations from the sample median. We also evaluated the effect of the cases with $E_p > 4$ J/kg, the potentially fictitious retrievals mentioned above, by repeating the calculations without them. All the results are shown in Table 1. It may be seen that the extreme tail events produce some effect on the mean and the median values (they represent here about 3% of the soundings, whereas in the histogram of Fig. 5 they were about 4% of the cases). However, E_p changes between autumn and spring remain statistically significant in both regions, regardless wether the artifacts are present or not or the mean or the median are used (68 % confidence intervals from Table 1 do not overlap). However, the median is remarkably more robust in the presence of the outliers. But it should be also observed that the mean highlights the absolute variations more than the median, because the latter is always smaller due to the skewed distribution of E_p, independently from the season or the region. Even if the mean is used to present results, it does not imply that credibility on the normal distribution becomes restored.

Monte Carlo simulations were performed in order to quantify the possible effects of outliers on the mean and the median and the possible effect of under-sampling in their calculations. We constructed GW fields according to the following procedure. We considered the East box 40–50S during the month of October, which has more than 1000 RO available (Fig. 2) to construct a power spectral density (PSD) against vertical wavenumber k_z, which resembled the typical k_z^{-3} relation (see e.g., Smith et al., 1987) from a least squares fit. The E_p mean is 1.1 J/kg, the median 0.6 J/kg and about 4% of the values are above

Table 1
Mean $\bar{E}_p$ and median $\hat{E}_p$ energy values describing GW activity to the East of the Patagonian Andes and the Antarctic Peninsula, their associated uncertainties and the number N_s of GPS RO soundings used in the calculations (superscript 4 refers to the removal of cases with $E_p > 4$ J/kg). Uncertainties were calculated as explained in the main text.

	Autumn	Spring
Patagonia		
$\bar{E}_p$ (J/kg)	0.81 ± 0.02	1.07 ± 0.03
$\bar{E}_p^4$ (J/kg)	0.70 ± 0.01	0.86 ± 0.01
$\hat{E}_p$ (J/kg)	0.49 ± 0.01	0.64 ± 0.02
$\hat{E}_p^4$ (J/kg)	0.48 ± 0.01	0.61 ± 0.01
N_s	3361	2914
N_s^4	3297	2822
Antarctica		
$\bar{E}_p$ (J/kg)	0.45 ± 0.02	1.03 ± 0.02
$\bar{E}_p^4$ (J/kg)	0.41 ± 0.01	0.89 ± 0.01
$\hat{E}_p$ (J/kg)	0.31 ± 0.01	0.72 ± 0.01
$\hat{E}_p^4$ (J/kg)	0.31 ± 0.01	0.71 ± 0.01
N_s	5116	4655
N_s^4	5101	4541

4 J/kg (see Fig. 5). For the same geographical area and month we created synthetic vertical temperature profiles between 15 and 33 km height, by adding background and perturbation components. The former parts consisted of 1100 vertical profiles of temperature between 18 km and 33 km from daily NCEP reanalysis randomly selected between the latitude and longitude limits of the box within days of October between years 2006 and 2011. Five sinusoidal waves were added to each background profile, where each phase and vertical wavelength (between 2 and 18 km) were randomly assigned and the amplitude was determined according to the constructed PSD(k_z). We randomly selected 4 % of the profiles to introduce a downward dying 4 K amplitude oscillation between 33 km and 25 km to simulate the GPS RO initialization issue as a possible source of the outliers. We computed E_p for all the profiles with the above explained standard procedure, including the use of the digital band-pass filter. We finally considered 50 random sets of each subgroup size of 100, 200, 300,...profiles and computed their dispersion with and without the 4% of outliers. The impact of the outliers is quite clear on the mean and the median in Fig. 9, but the effect on the latter is less significant in magnitude. In addition, the lower dispersion of the median is also quite visible (for samples of diverse size). Note the similarity of all these effects with those of Table 1.

It should be stressed that Luna et al. (2013) have shown that the effect of the temperature measurement uncertainties on the individual energy calculation of each sounding is much less significant than the data processing procedures (mainly the choice of the detailed separation of background and waves and the height interval analyzed), so the same applies to the propagation into the statistical representation of the GW energy. It may be then also inferred that the distribution shape may be very similar among different GW studies for the same region and season, but the climatologies and in particular the values of the statistical parameters are only comparable if the same measurement technique and data processing have been employed.

4. Results with SABER data

In Fig. 10 we show the number of soundings per band and month. Polar regions are not covered during certain time intervals (John and Kumar, 2012), so some empty groups will be found in our calculations. In Fig. 11 we exhibit the E_p means and medians. Again the maximum levels of activity are found during austral winter and spring. The effect is most visible at the highest latitudes for both quantities. Also the median values are always lower than the mean values, but they both follow the same trends. In addition for both parameters, values to the East are nearly systematically larger than to the West, which as with GPS RO data confirms the significant role of topography for GW generation in the studied region. SABER lacks data between 60 and 90S for some months, so comparisons with GPS RO are hard to establish in that region. Fig. 12 compares to Fig. 5 and it may be seen that the distribution

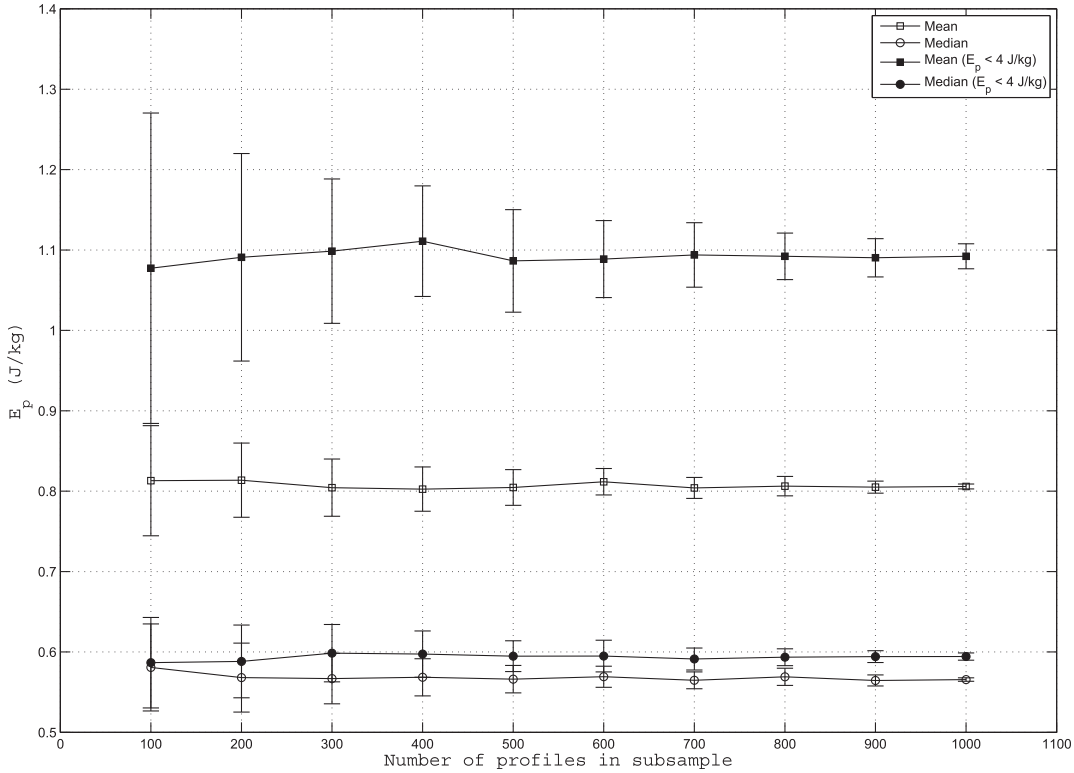

Fig. 9. Mean and median E_p values and dispersion for different sampling sizes according to Monte Carlo simulations for 30–40S, East box during typical October background conditions and the possible effect of outliers on GPS RO calculations.

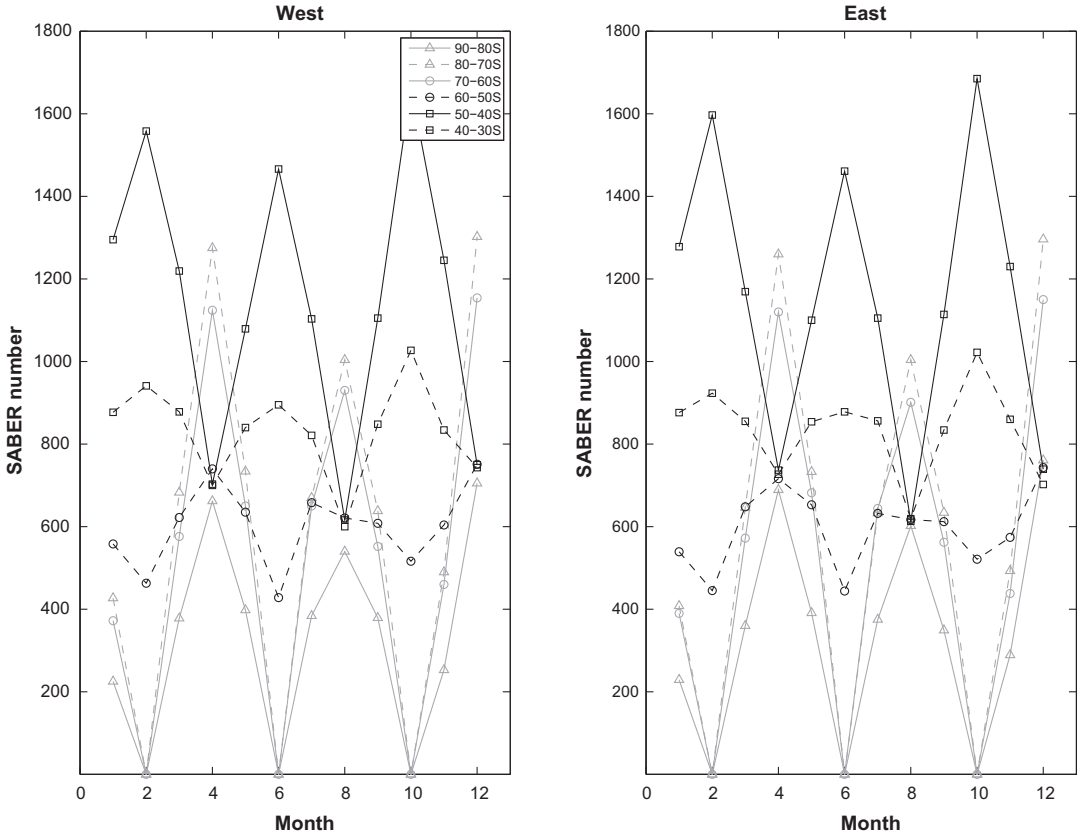

Fig. 10. Amount of SABER profiles per band against month number. Lines corresponding to cases with some null values have been drawn with light gray to avoid a messy look due to excessive superposition.

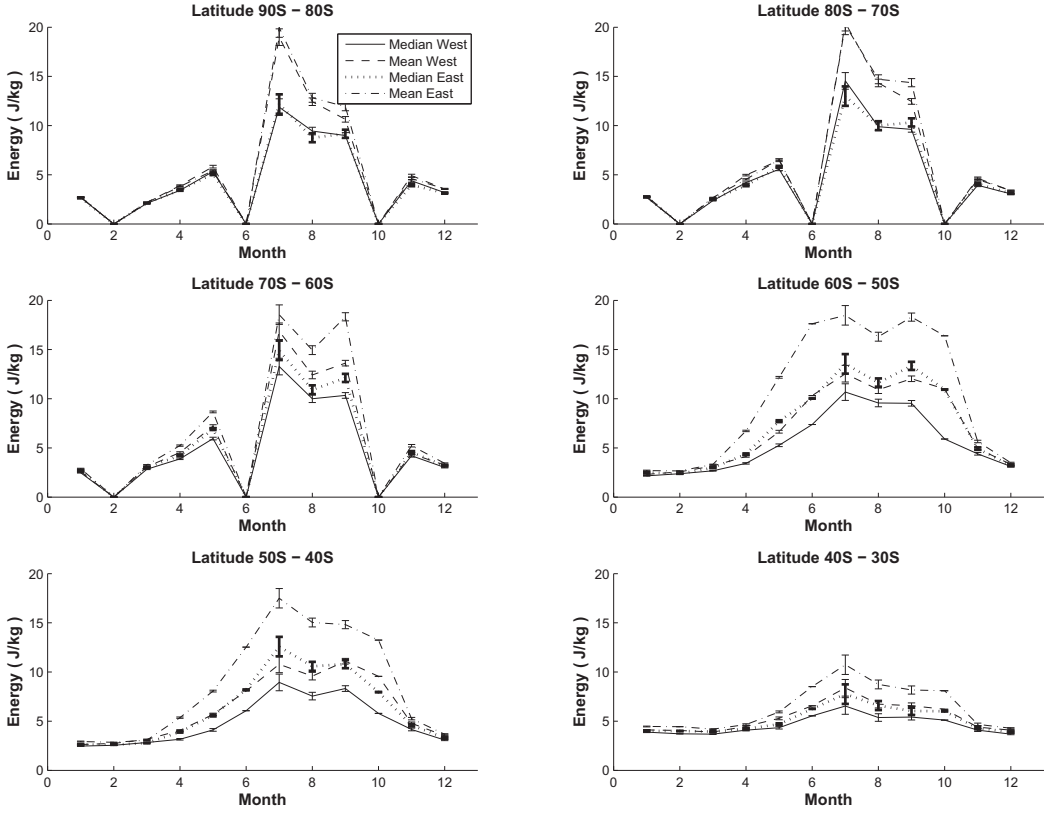

Fig. 11. Mean and median E_p and their error bars for SABER profiles per zones against month number.

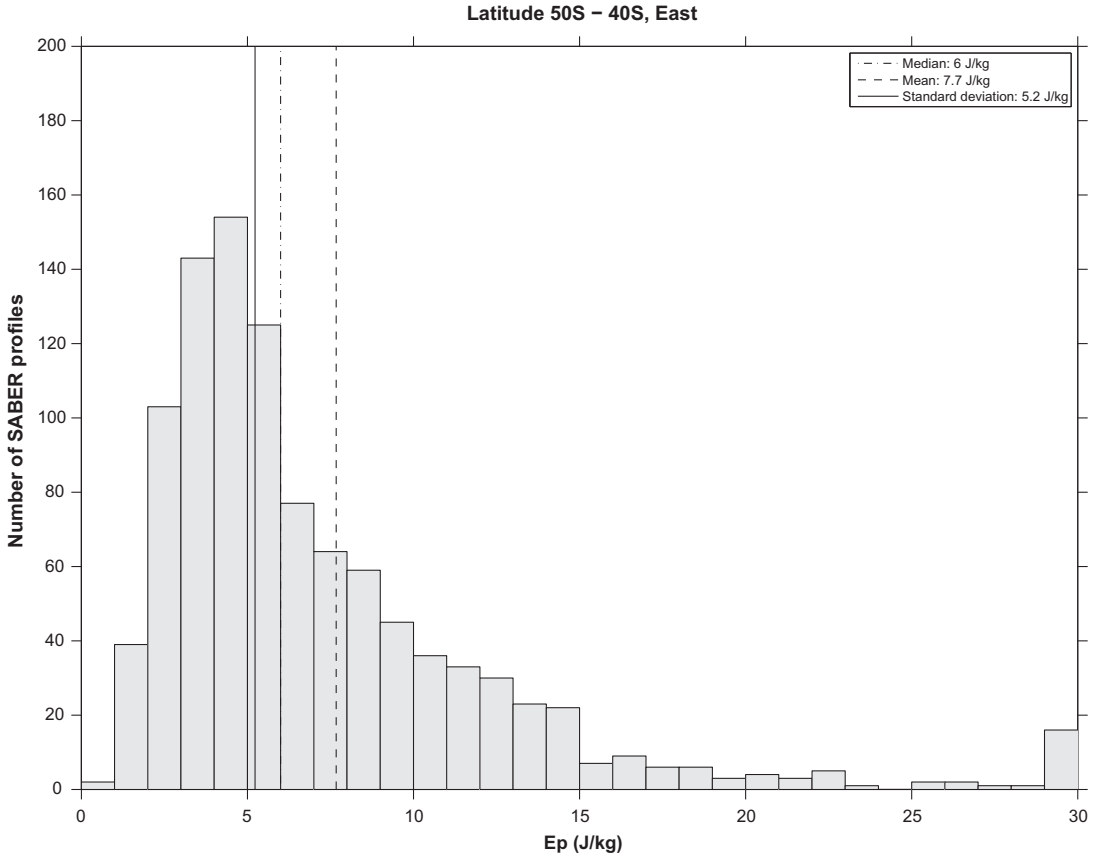

Fig. 12. A E_p histogram for the SABER soundings in latitude band 30–40S to the East of 70W during October of the 5 years studied.

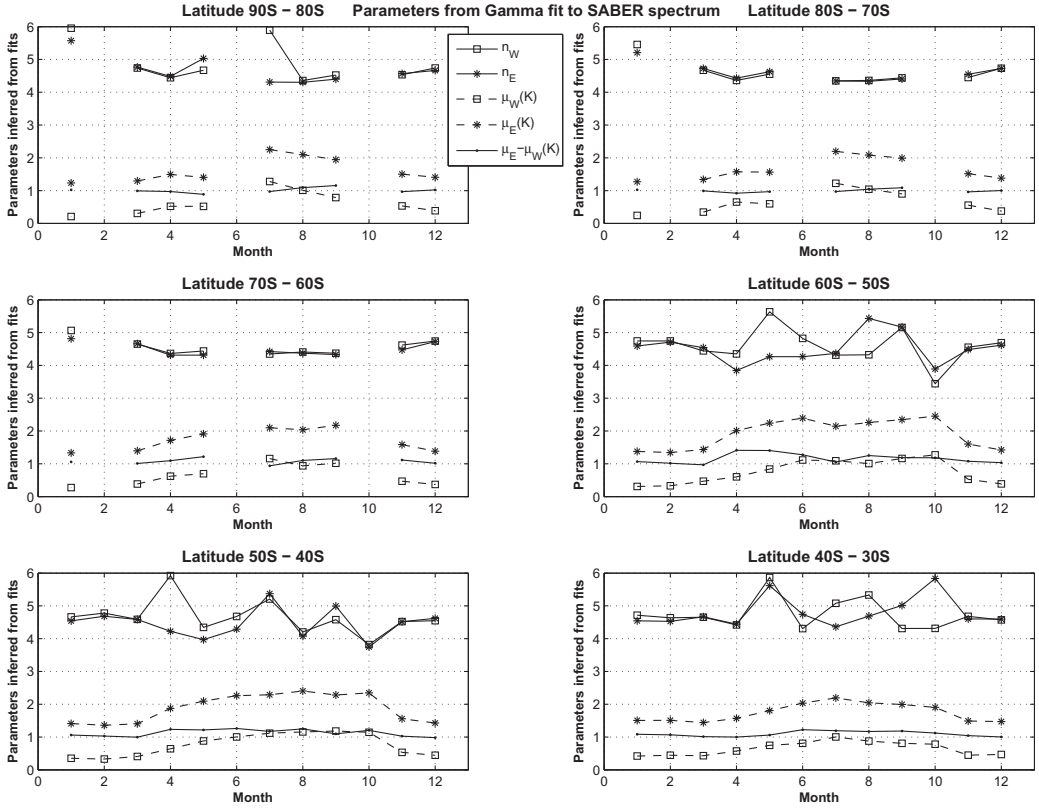

Fig. 13. Quantities derived from fits of E_p from SABER soundings to a Γ distribution. Notice that μ_E has been vertically shifted by 1 unit to avoid the superposition with μ_W. The difference $\mu_E - \mu_W$ is shown centered at 1. Subscripts W and E refer to the West and East of 70W.

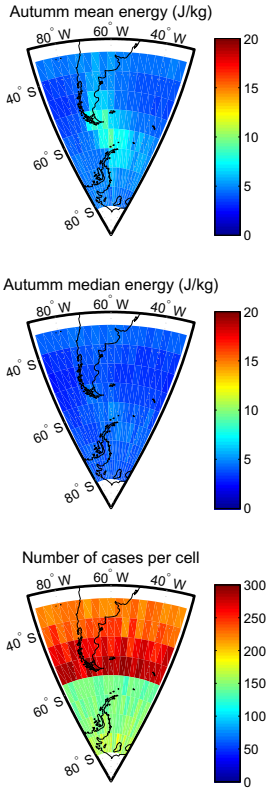

Fig. 14. The mean and the median E_p and the number of SABER profiles per cell in the analyzed zone for autumn over the 5 years studied.

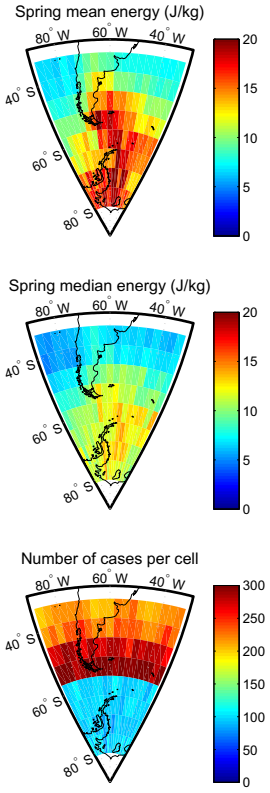

Fig. 15. The same as Fig. 14 for spring.

shape is similar but values are typically different. It is not only that both techniques have different observational windows, but also as stated above the different processing and assumptions at the retrieval stage lead to differences in the outcomes. Fig. 13 is the SABER equivalent to Fig. 6 for GPS RO. It should be said that differences in the typical number of significant modes may be expected from both observational techniques because although their filtering window is similar, it is partially different in respect to the observed horizontal and vertical wavelengths. The same applies to the calculated mean amplitudes. Finally, energies in Figs. 14 and 15 resemble Figs. 7 and 8, mainly the hot spot to the East of Southern Patagonia and the Antarctic Peninsula and most importantly during Spring.

5. Conclusions

We used 5 years of GPS RO data from the COSMIC mission in a zone close to the Patagonian Andes and their prolongation in the Antarctic Peninsula that exhibits a strong GW activity during several months per year. We verified diverse distribution hypotheses for GW activity with χ^2 goodness-of-fit tests. The normal and log-normal functions were analyzed, but we also investigated the Γ function, which was not used in previous work. It was shown to be a reasonable alternative to the log-normal curve as a good representation of the E_p distribution. For GPS RO data the normal hypothesis became rejected in 47% of all the tests, so it may be discarded as an adequate representation of the E_p distribution. The log-normal hypothesis was rejected in no sample and the gamma distribution in just 4% of the tests. For SABER profiles in the same geographical area and time span the rejection rates were larger, respectively 78%, 2% and 7%. This implies that all the representations performed worse for SABER. The results do not exhibit a specific dependence on the latitudinal band or month. The Γ function is slightly not as good as the log-normal representation to describe the E_p histograms but it has the advantage that it offers an explanation of the observed distribution in terms of GW properties: the obtained parameter values may lead to extra information in relation to the modes contributing to the observed climatology. The Γ interpretation allows not only to quantify the GW activity level of each climatology, but also to find out the number of significant modes that essentially determine it. Number differences between West and East boxes are typically less than one so the higher E_p values on the East side should be attributed to strength rather than to the number of modes.

Although the success of the log-normal distribution function may suggest a priori that the median might be a more adequate parameter than the mean to represent the level of GW activity in a given region and time, it was found that the results of both parameters are nearly equivalent regarding the statistical significance of regional or seasonal variations. It was shown that the calculated mean

values with GPS RO data could be significantly affected in the presence of spurious large activity cases located in the distribution tail (possible artifacts generated by initialization problems in some profiles that were not ruled out in the quality control process). If those spurious cases are absent or minimal in the data, then the mean should be used, as it tends to better highlight than the median the absolute variations. However, if those artifacts are significantly present, then the median should be used. The mean has the advantage that it has a clearer meaning. It is an estimator of the average state of the atmosphere in a given area, height interval and season given possible variabilities in GW potential energy within the three corresponding ranges. The meaning of the median is not as straightforward, but its advantage relies on the fact that its definition is independent from the distribution, so we may even ignore the latter. All these calculations and tests may be extended in the future to other geographical regions in order to be then eventually able to reach a higher degree of generalization of the results.

Acknowledgements

Manuscript prepared under grant CONICET PIP 11220120100034, ANPCYT PICT 2013-1097 and German BMBF grant 01DN14001. P. Alexander and A. de la Torre are members of CONICET. GPS RO data downloaded from cdaac-ftp.cosmic.ucar.edu and SABER data from saber.gats-Inc.com. Reanalysis data for Fig. 4 provided by the NOAA-ESRL Physical Sciences Division, Boulder Colorado from www.esrl.noaa.gov/psd.

Appendix A. The derivation of the distribution for E_p

If a quantity $X_i \epsilon [0, \infty)$ and it exhibits a so-called half-normal distribution, then it depends on a single scale parameter σ. A property relevant for our work is that

$$(X_i/\sigma)^2 \sim \chi_1^2$$

i.e. chi-squared distribution with one degree of freedom (e.g., Johnson et al., 1994). Then, if successive X_i are independent and have the same half-normal probability distribution,

$$\sum_{i=1}^{n} (X_i/\sigma)^2 \sim \chi_n^2$$

i.e chi-squared distribution with n degrees of freedom. The last equation is a consequence of a well-known property of the addition of independent chi-squared variables.

If we recall that X_i are the relative temperature perturbation amplitudes of the modes, then from Eq. (2) we obtain that

$$E_p/(1/4(g/N)^2\sigma^2) \sim \chi_n^2 \tag{A.1}$$

Let us consider a random variable

$$Z \sim \chi_n^2$$

and

$$a = 1/4(g/N)^2\sigma^2 \tag{A.2}$$

Then for $a > 0$

$$Y = aZ \sim \Gamma(n/2, 2a)$$

as shown e.g. by Johnson et al. (1994).

From the fit of each E_p data-set to the Γ distribution we obtain two optimal values: one for $n/2$ and another one for $2a$. From the first Γ parameter value we may infer the typical number of relevant modes n that determine relation (A.1). Non-integer values are obtained as it reflects an estimate over all the measurements that were made. To obtain σ we use the second Γ parameter value with the aid of Eq. (A.2). The σ result should be considered a measure of the typical relative amplitudes of the modes. This follows from the fact that they exhibit a half-normal distribution, so

$$\mu = \sigma\sqrt{2/\pi}$$

where μ is the mean. We used $g = 9.8$ m s^{-2} and $N = 2.2 \times 10^{-2}$ s^{-1} in the stratosphere when calculating the E_p values in Eq. (1).

References

Alexander, M.J., Teitelbaum, H., 2011. Three-dimensional properties of Andes mountain waves observed by satellite: a case study. J. Geophys. Res. 116, D23110.

Alexander, P., de la Torre, A., Llamedo, P., 2008a. Interpretation of gravity wave signatures in GPS radio occultations. J. Geophys. Res. 113, D16117.

Alexander, M.J., Gille, J., Cavanaugh, C., et al., 2008b. Global estimates of gravity wave momentum flux from high resolution dynamics limb sounder observations. J. Geophys. Res. 113, D15S18.

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

Alexander, M.J., Geller, M., McLandress, C., et al., 2010a. Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models. Q.J.R. Meteorol. Soc. (1103–1124), 1103–1124

Alexander, P., Luna, D., Llamedo, P., de la Torre, A., 2010b. A gravity waves study close to the Andes mountains in Patagonia and Antarctica with GPS radio occultation observations. Ann. Geophys. 28, 587–595.

Alexander, S.P., Klekociuk, A.R., Pitts, M.C., McDonald, A.J., Arevalo-Torres, A., 2011. The effect of orographic gravity waves on Antarctic polar stratospheric cloud occurrence and composition. J. Geophys. Res. 116, D06109.

Allen, S.J., Vincent, R.A., 1995. Gravity wave activity in the lower atmosphere: seasonal and latitudinal variations. J. Geophys. Res. 100, 1327–1350.

Ao, C.O., Hajj, G.A., Ijima, B.A., 2006. Sensitivity of stratospheric retrievals from radio occultations on upper boundary conditions. In: Atmosphere and Climate. Springer, Berlin, pp. 17–26.

Baumgaertner, A., McDonald, A., 2007. A gravity wave climatology for Antarctica compiled from challenging minisatellite payload/global positioning system (CHAMP/GPS) radio occultations. J. Geophys. Res. 112, D05103.

de la Torre, A., Alexander, P., 2005. Gravity waves above Andesdetected from GPS radio occultation temperature profiles: mountain forcing?. Geophys. Res. Lett. 32 L17815.
de la Torre, A., Giraldez, A., Alexander, P., 1994. Saturated gravity wave spectra measured with balloons in Mendoza (Argentina). Geophys. Res. Lett. 21, 2039–2042.
de la Torre, A., Alexander, P., Llamedo, P., et al., 2006. Gravity waves above the Andes detected from GPS radio occultation temperature profiles: jet mechanism? Geophys. Res. Lett. 33, L24810.
de la Torre, A., Llamedo, P., Alexander, P., et al., 2010. Estimated errors in a global gravity wave climatology from GPS radio occultation temperature profiles. Adv. Space Res. 46, 174–179.
Eckermann, S.D., Preusse, P., 1999. Global measurements of stratospheric mountain waves from space. Science 286, 1534–1537.
Faber, A., Llamedo, P., Schmidt, T., de la Torre, A., Wickert, J., 2013. On the determination of gravity wave momentum flux from GPS radio occultation data. Atmos. Meas. Tech. 6, 31693180.
Fritts, D.C., Alexander, M.J., 2003. Gravity wave dynamics and effects in the middle atmosphere. Rev. Geophys. 41, 1003.
Gavrilov, N.M., Fukao, S., 1999a. The MU radar measurements and modeling of the gravity wave climatology and sources in the atmosphere. Adv. Space Res. 24, 565–574.
Gavrilov, N.M., Fukao, S., 1999b. A comparison of seasonal variations of gravity wave intensity observed by the MU radar with a theoretical model. J. Atmos. Sci. 56, 3485–3494.
Gavrilov, N.M., Fukao, S., 2001. Hydrodynamic tropospheric wave sources and their role in gravity wave climatology of the upper atmosphere from the MU radar observations. J. Atmos. Sol. Terr. Phys. 63, 931–943.
Hamming, R.W., 1998. Digital Filters, third ed. Dover Publications, Mineola, New York.
Hei, H., Tsuda, T., Hirooka, T., 2008. Characteristics of atmospheric gravity wave activity in the polar regions revealed by GPS radio occultation data with CHAMP. J. Geophys. Res. 113, D04107.
Hertzog, A., Alexander, M.J., Plougonven, R., 2012. On the intermittency of gravity wave momentum flux in the stratosphere. J. Atmos. Sci. 69, 34333448.
Jiang, Q., Doyle, J.D., Reinecke, A., et al., 2013. A modeling study of stratospheric waves over the southern andes and drake passage. J. Atmos. Sci. 70, 1668–1689.
John, S.R., Kumar, K.K., 2012. TIMED/SABER observations of global gravity wave climatology and their interannual variability from stratosphere to mesosphere lower thermosphere. Clim. Dyn. 39, 14891505.
Johnson, N.L., Kotz, S., Balakrishnan, N., 1994, . Continuous Univariate Distributions, second ed., vol. 1. Wiley, New York.
Kursinski, E.R., Hajj, G.A., Schofield, J.T., et al., 1997. Observing Earths atmosphere with radio occultation measurement using the global positioning system. J. Geophys. Res. 102, 23429–23465.
Kursinski, E.R., Hajj, G.A., Leroy, S.S., Herman, B., 2000. The GPS radio occultation technique. Terr. Atmos. Ocean. Sci. 11, 53–114.
Liou, Y.-A., Pavelyev, A.G., Liu, S.-F., et al., 2007. FORMOSAT-3/COSMIC GPS radio occultation mission: preliminary results. IEEE Trans. Geosci. Remote Sens. 45, 3813–3826.
Luna, D., Alexander, P., de la Torre, A., 2013. Evaluation of uncertainty in gravity wave potential energy calculations through GPS radio occultation measurements. Adv. Space Res. 52, 879–882.
Marquardt, C., Healy, S., 2005. Measurement noise and stratospheric gravity wave characteristics obtained from GPS occultation data. J. Meteorol. Soc. Jpn. 83, 417–428.
McDonald, A.J., Hertzog, A., 2008. Comparison of stratospheric measurements made by CHAMP radio occultation and Strateole/Vorcore in situ data. Geophys. Res. Lett. 35, L11805.
McDonald, A.J., Tan, B., Chu, X., 2010. Role of gravity waves in the spatial and temporal variability of stratospheric temperature measured by COSMIC/FORMOSAT-3 and Rayleigh lidar observations. J. Geophys. Res. 115, D19128.
Mlynczak, M.G., 1997. Energetics of the mesosphere and lower thermosphere and the SABER experiment. Adv. Space Res. 20, 11771183.
Moffat-Griffin, T., Jarvis, M.J., Colwell, S.R., Kavanagh, A.J., Manney, G.L., Daffer, W.H., 2013. Seasonal variations in lower stratospheric gravity wave energy above the Falkland Islands. J. Geophys. Res. 118, 10861–10869.
Nastrom, G.D., Gage, K.S., 1985. A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. J. Atmos. Sci. 42, 950960.
Plougonven, R., Hertzog, A., Guez, L., 2013. Gravity waves over Antarctica and the Southern Ocean: consistent momentum fluxes in mesoscale simulations and stratospheric balloon observations. Q.J.R. Meteorol. Soc., 101–118
Preusse, P., Eckermann, S.D., Offermann, D., 2000. Comparison of global distributions of zonal-mean gravity wave variance inferred from different satellite instruments. Geophys. Res. Lett. 27, 38773880.
Preusse, P., Eckermann, S.D., Ern, M., 2008. Transparency of the atmosphere to short horizontal wavelength gravity waves. J. Geophys. Res. 113, D24104.
Ratel, G., 2006. Median and weighted median as estimators for the key comparison reference value (KCRV). Metrologia 43, S244–S248.
Sato, K., Satoshi, T., Watanabe, S., et al., 2012. Gravity wave characteristics in the southern hemisphere revealed by a high-resolution middle-atmosphere general circulation model. J. Atmos. Sci. 69, 1378–1396.
Sheather, S.J., Maritz, J.S., 1983. An estimate of the asymptotic standard error of the sample median. Aust. J. Stat. 25, 109–122.
Sivakumar, V., Vishnu Prasanth, P., Kishore, P., Bencherif, H., Keckhut, P., 2011. Rayleigh LIDAR and satellite (HALOE, SABER, CHAMP and COSMIC) measurements of stratosphere-mesosphere temperature over a southern sub-tropical site, Reunion (20.8°S; 55.5°E): climatology and comparison study. Ann. Geophys. 29, 649–662.
Smith, S.E., Fritts, D.C., VanZandt, T.E., 1987. Evidence of saturated spectrum of atmospheric gravity waves. J. Atmos. Sci. 44, 14041410.
Uccellini, L.W., Koch, S.E., 1987. The synoptic setting and possible energy sources for mesoscale wave disturbances. Mon. Weather Rev. 115, 721–729.
von Storch, H., Zwiers, F.W., 2003. Statistical Analysis in Climate Research. Cambridge University Press, Cambridge, UK.
Wang, L., Alexander, M.J., 2010. Global estimates of gravity wave parameters from GPS radio occultation temperature data. J. Geophys. Res. 115, D21122.
Wilson, R., Chanin, M.L., Hauchecarne, A., 1991. Gravity waves in the middle atmosphere observed by Rayleigh lidar 1. Case studies. J. Geophys. Res. 96, 5153–5167.
Wright, C.J., Rivas, M.B., Gille, J.C., 2011. Intercomparisons of HIRDLS, COSMIC and SABER for the detection of stratospheric gravity waves. Atmos. Meas. Tech. 4, 1581–1591.
Wu, D.L., Preusse, P., Eckermann, S.D., et al., 2006. Remote sounding of atmospheric gravity waves with satellite limb and nadir techniques. Adv. Space Res. 37, 22692277.
Yan, X., Arnold, N., Remedios, J., 2010. Global observations of gravity waves from high resolution dynamics limb sounder temperature measurements: a yearlong record of temperature amplitude and vertical wavelength. J. Geophys. Res. 115, D10113.