Manuscript prepared for Atmos. Meas. Tech.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 23 October 2015

Evaluation of methods for gravity wave extraction from middle atmospheric lidar temperature measurements

B. Ehard, B. Kaifler, N. Kaifler, and M. Rapp

Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

Correspondence to: B. Ehard (benedikt.ehard@dlr.de)

Abstract. This study evaluates commonly used methods of extracting gravity wave induced temperature perturbations from lidar measurements. The spectral response of these methods is characterized with the help of a synthetic dataset with known temperature perturbations added to a realistic background temperature profile. The simulations are carried out with the background temperature being either constant or varying in time to evaluate the sensitivity to temperature perturbations not caused by gravity waves. The different methods are applied to lidar measurements over new Zealand and the performance of the algorithms is evaluated. We find that the Butterworth filter performs best if gravity waves over a wide range of periods are to be extracted from lidar temperature measurements. The running mean method gives good results if only gravity waves with short periods are to be analyzed.

1 Introduction

Atmospheric gravity waves are well known to have a strong impact on the middle atmospheric circulation (e.g. Holton and Alexander, 2000; Fritts and Alexander, 2003). By transporting energy and momentum from the lower atmosphere into the middle atmosphere they are responsible for the formation of the cold polar summer mesopause (e.g. Lindzen, 1981). Although some processes related to gravity waves are believed to be well understood there are still open questions. For example, in how far gravity wave excitation, propagation and forcing is affected by a changing climate remains an open question (cf. Fritts and Alexander, 2003; Plougonven and Zhang, 2014).

Lidar technology has been used to study gravity waves in the middle atmosphere for the last three decades (e.g. Chanin and Hauchecorne, 1981; Gardner et al., 1989; Wilson et al., 1991; Whiteway and Carswell, 1995; Duck et al., 2001; Rauthe et al., 2008; Yamashita et al., 2009; Alexander et al., 2011; Kaifler et al., 2015b). Hence, lidar studies can potentially be used to infer long-term trends in gravity wave activ-

ity. Furthermore, lidars have the advantage of providing measurements throughout the entire middle atmosphere with high temporal and vertical resolution of typically $1\,\mathrm{h}$ and $1\,\mathrm{km}$. However, lidars generally provide one dimensional profiles and no information on the horizontal structure and the intrinsic properties of atmospheric waves can be retrieved. Exceptions are measurements from airborne lidars and multi-beam lidars.

Gravity waves are usually determined from lidar measurements by separating an estimated background temperature (density) profile from the measured profiles in order to derive temperature (density) perturbation profiles. Several methods have been developed and used over the last decades. For example Gardner et al. (1989), Rauthe et al. (2008) and Ehard et al. (2014) calculate a nightly mean profile and subtract it from the (time resolved) individual profiles. Yamashita et al. (2009) remove a background profile determined by a temporal running mean (in addition to vertical filtering). Perturbation profiles obtained through a fit of polynomial functions to the measured profiles are examined e.g. by Whiteway and Carswell (1995), Duck et al. (2001) or Alexander et al. (2011). Mzé et al. (2014) apply a variance method in order to determine perturbation profiles, while Chane-Ming et al. (2000) use spectral filtering.

All of these methods are most sensitive to different parts of the gravity wave spectrum. Thus, results from different lidar studies become hardly comparable because one cannot distinguish between variations that are caused by a different methodology and variations that are geophysically induced. Ehard et al. (2014) compared values of gravity wave potential energy density (GWPED) from different studies to their results. Due to potential methodological biases it remained unclear whether the differences were in fact of geophysical origin. Hence, they expressed the need for a standardized method to extract gravity wave amplitudes from lidar measurements.

To our knowledge, no literature is so far available which characterizes and evaluates the most commonly used methods to extract information on gravity waves from lidar profiles. Thus, we will evaluate and compare four methods in detail: subtraction of the nightly mean profile, subtraction of temporal running mean profiles, the sliding polynomial fit method proposed by Duck et al. (2001) and the application of a Butterworth filter. While the first two methods rely on filtering in time, the latter two methods apply a filter in space to determine wave induced temperature perturbations.

This paper is structured as follows: the four methods are described in detail in Sect. 2. The performance is studied in terms of their spectral response to synthetic data in Sect. 3. The results are then applied to measurement data in Sect. 4. Finally, the characteristics of the four methods as well as their suitability for extracting gravity wave induced temperature perturbations is discussed in Sect. 5 and conclusions are drawn in Sect. 6.

2 Methods

Lidar systems used for studies of the middle atmosphere measure the Rayleigh backscatter signal which is proportional to atmospheric density after range correction. The temperature is commonly retrieved by integration assuming hydrostatic equilibrium (Hauchecorne and Chanin, 1980). ~~Derived temperature~~ Recently Sica and Haefele (2015) proposed a temperature retrieval using optimal estimation methods. The derived temperature profiles typically range between 30 and 80–90 km altitude depending on signal-to-noise ratio. At the upper boundary, the temperature retrieval is commonly initalized with satellite data (e.g. Alexander et al., 2011) or resonance lidar measurements (e.g. Rauthe et al., 2008).

The combination with a resonance lidar system extends the altitude range of temperature measurements up to $\approx 105\,\mathrm{km}$. Temperatures below $30\,\mathrm{km}$ altitude can be retrieved by using a stratospheric Raman channel (e.g. Alpers et al., 2004). The large altitude range allows for studies of gravity wave propagation from the troposphere to the mesosphere. Hence, we discuss the extraction of gravity waves from temperature data rather than atmospheric density, although most of the results can be applied to density measurements as well. For different methods of extracting gravity waves from density measurements see e.g. Sica and Russell (1999) , Thurairajah et al. (2010) and Mzé et al. (2014) .

Lidar studies usually determine wave induced temperature perturbations $T'(z,t)$ (which are a function of altitude z and time t) from the measured temperature profile $T(z,t)$ by subtracting a background temperature profile $T_0(z,t)$:

$$T'(z,t) = T(z,t) - T_0(z,t)\,. \tag{1}$$

$T_0(z,t)$ ideally contains all contribution from radiative and chemical heating and other large scale effects such as planetary waves and tides. Hence, the temperature perturbations $T'(z,t)$ should be solely caused by gravity waves. Estimation of $T_0(z,t)$ is challenging due to the specific shape of the temperature profile with its changes in vertical temperature gradient, e.g. at the stratopause or mesopause.

The frequency range of gravity waves which may be present in $T'(z,t)$ can be inferred from the gravity wave dispersion relation which states that the relation

$$N > |\hat{\omega}| > f \tag{2}$$

between the intrinsic frequency $\hat{\omega}$, the Brunt–Väisälä frequency N and the Coriolis parameter f must be fulfilled at all times. Using a typical stratospheric value of $N = 0.02\,\mathrm{s}^{-1}$ and a Coriolis parameter for mid-latitudes of $f = 10^{-4}\ \mathrm{s}^{-1}$, the intrinsic period $\hat{\tau} = \frac{2\pi}{\hat{\omega}}$ ranges between $5\,\mathrm{min}$ and $17\,\mathrm{h}$. It is important to note that the lidar only detects the observed period τ which can be Doppler shifted to larger or smaller values, depending on local wind conditions. Typical vertical wavelengths of gravity waves measured by ground based instruments vary between 1 and $17\,\mathrm{km}$

(see Chane-Ming et al., 2000, their Table 2). The spatial scales combined with the temporal scales define the spectral requirements on the methods of extracting gravity wave induced temperature perturbations.

2.1 Time-averaged background profiles

A widely applied method is the use of the nightly mean temperature profile as background temperature profile (e.g Gardner et al., 1989; Rauthe et al., 2008; Ehard et al., 2014). Thereby it is assumed that the timescales of phenomena other than gravity waves affecting the temperature profile are considerably larger and the timescales of gravity waves are smaller than the measurement period, which is typically in the range of 3–12 h.

Another common method is to determine background temperature profiles by means of a running mean over a time window which is typically on the order of $3\,\mathrm{h}$ (e.g. Yamashita et al., 2009). Temperature variations with timescales larger than the window width are attributed to the background temperature profiles and are therefore not included in the extracted gravity wave spectrum.

2.2 Sliding polynomial fit

Duck et al. (2001) proposed a method of extracting temperature perturbations based on a sliding polynomial fit in the spatial domain. The method is sensitive to small vertical scales and ignores the temporal evolution of waves. The method is based on the assumption that temperature variations with large vertical scales can be attributed either to the climatological thermal structure of the atmosphere (i.e. the different vertical temperature gradients in the troposphere, stratosphere and mesosphere), the advection of colder or warmer air masses, or tides and planetary waves. Only variations with a spatial scale smaller than a certain threshold are identified as gravity waves.

The sliding polynomial fit method was designed to produce a background temperature profile which contains all perturbations with vertical scales larger than $15\,\mathrm{km}$. For each measured temperature profile Duck et al. (2001) applied a series of overlapping cubic polynomial fits to each range gate. Each fit was applied to an altitude window with a width of $L_\mathrm{f} = 25\,\mathrm{km}$. A weighted average was computed to reconstruct the background temperature profile from the individual polynomial fits using the weighting function

$$w(z)_i = \begin{cases} \exp\left(\frac{z-(z_{\mathrm{c},i}-\delta)}{\gamma}\right) & \text{if } z \leq z_{\mathrm{c},i} - \delta \\ 1 & \text{if } z_{\mathrm{c},i} - \delta < z < z_{\mathrm{c},i} + \delta \\ \exp\left(-\frac{z-(z_{\mathrm{c},i}+\delta)}{\gamma}\right) & \text{if } z \geq z_{\mathrm{c},i} + \delta\,. \end{cases} \tag{3}$$

Here $\delta = 0.5L_\mathrm{f} - L_\mathrm{w}$, L_w is the width of the weighting window, $z_{\mathrm{c},i}$ the center altitude of the individual fit and γ the e-folding width which defines how fast the weighting function decreases. Duck et al. (2001) used a weighting window length $L_\mathrm{w} = \frac{L_\mathrm{f}}{3}$ and $\gamma = 3\,\mathrm{km}$.

Duck et al. (2001) smoothed the resulting background temperature profiles with a $1.5\,\mathrm{km}$ boxcar mean. These profiles were then subtracted from the corresponding measured temperature profiles according to Eq. (1), yielding the temperature perturbation profiles.

In this study the following set of parameters is used: a fit length $L_\mathrm{f} = 20\,\mathrm{km}$, a weighting window length $L_\mathrm{w} = 3\,\mathrm{km}$ and an e-folding width $\gamma = 9\,\mathrm{km}$. These parameters are chosen because they yield the flattest spectral response for the altitude resolution used in this study (see Sect. 5 for further details). The boxcar smoothing showed to have a negligible effect. Hence, it is not applied in this study.

2.3 Spectral filter

Another method which can be applied to vertical profiles is spectral filtering (e.g. Chane-Ming et al., 2000). By applying a high-pass filter to individual temperature profiles, temperature perturbations can be retrieved. In order to yield perturbations caused by gravity waves, a filtering function has to be chosen which has an adequate spectral response.

In this study we use a 5th order Butterworth high-pass filter with a cutoff wavelength $\lambda_\mathrm{c} = 15\,\mathrm{km}$ and the transfer function

$$H(\lambda_z) = \left(1 + \left(\frac{\lambda_z}{\lambda_\mathrm{c}}\right)^{2n}\right)^{-\frac{1}{2}}, \tag{4}$$

where n is the order of the filter and λ_z is the vertical wavelength. The Butterworth filter is chosen due to its flat frequency response in the passband. The filter itself is applied in Fourier space. As the Fourier transformation assumes a cyclic dataset, the upper and lower end of the measured temperature profile are internally connected. This creates an artificial discontinuity which introduces a broad range of frequencies including frequency components that are in the passband of the filter. These frequency components contribute to temperature perturbations at the upper and lower end of the analyzed altitude window and thus artificially enhance gravity wave signatures. In order to mitigate this effect, the dataset is mirrored at the lowest altitude bin and attached to the original dataset before the filtering process. Thereby, the dataset can be cyclic extended without discontinuities at the lower end, where temperature perturbations are smallest and therefore artificial enhancements produce largest relative errors. After the filtering only the original half of the resulting perturbation profile is retained.

3 Application to synthetic data

In order to characterize the different methods regarding their ability to extract temperature perturbations from middle atmospheric temperature profiles, we apply them to a synthetic dataset with known temperature perturbations. These perturbations are added to a fixed, realistic background temperature profile $T_0(z)$. The latter is derived from the mean temperature profile above Lauder,

New Zealand, (45.0° S, 169.7° E) measured with the Temperature Lidar for Middle Atmosphere research (TELMA) from July until end of September (black line in Fig. 1a). The particular choice of the background temperature profile does not affect the results as long as the background temperature profile is realistic, smooth and does not contain contributions from gravity waves. For example, with a climatological or a model temperature profile, similar results can be derived.

Sinusoidal temperature perturbations with exponentially increasing amplitude were added to the background temperature profile according to

$$T_\mathrm{s}(z,t) = T_0(z) + T'_\mathrm{s}(z,t), \quad \text{with} \tag{5}$$

$$T'_\mathrm{s}(z,t) = A\cos\left(\frac{2\pi z}{\lambda_z} + \frac{2\pi t}{\tau}\right)\exp\left(\frac{z-z_0}{2H}\right), \tag{6}$$

with the amplitude A, the vertical wavelength λ_z, the observed period τ the scale height H and the lowest altitude of the analyzed altitude range z_0. An example of the perturbed background profile T_s can be seen in Fig. 1a (red line) and the corresponding temperature perturbations T'_s in Fig. 1b.

For each method, the spectral response $R_m(z)$ was calculated from the ratio between the time averaged absolute values of the determined temperature perturbations $\overline{|T'_m(z,t)|}$ and the synthetic temperature perturbations $\overline{|T'_\mathrm{s}(z,t)|}$ as

$$R_m(z) = \frac{\overline{|T'_m(z,t)|}}{\overline{|T'_\mathrm{s}(z,t)|}} \cdot 100\,\% \tag{7}$$

A spectral response larger than $100\,\%$ indicates an overestimation of gravity wave amplitude, while a value below $100\,\%$ indicates an underestimation of gravity wave amplitude.

All simulations conducted for this study use the realistic set of parameters $A = 1.2\,\mathrm{K}$, $H = 12\,\mathrm{km}$ and $z_0 = 25\,\mathrm{km}$. A height resolution of $\Delta z = 0.1\,\mathrm{km}$ was used, while the altitude interval ranged from 25 to $90\,\mathrm{km}$. A time interval of $8\,\mathrm{h}$, corresponding to the length of an average nighttime measurement period, with a resolution of $\Delta t = 0.5\,\mathrm{h}$ was used. For each simulation either λ_z or τ was kept constant, while the other was varied. The vertical wavelength λ_z was varied from 0.6 to $20\,\mathrm{km}$ in steps of $0.2\,\mathrm{km}$, while τ was varied from 0.15 to $14.95\,\mathrm{h}$ in steps of $0.1\,\mathrm{h}$.

3.1 Constant background temperature

As a first step, simulations were carried out with a constant background temperature profile $T_0(z)$. In order to reduce aliasing effects caused by even multiples of the analyzed time window ($8\,\mathrm{h}$), the period of simulated gravity waves was set to $\tau = 1.9\,\mathrm{h}$ while the vertical wavelength λ_z was varied. Figure 2 depicts the spectral response of the different methods as a function of vertical wavelength.

The nightly mean method (Fig. 2a) and the $3\,\mathrm{h}$ running mean method (Fig. 2b) both exhibit an almost uniform spectral response at all altitudes and wavelengths. However, the running mean slightly overestimates the extracted temperature perturbations ~~.~~ which is due to the choice of a specific period of $\tau = 1.9\,\mathrm{h}$ (cf. Fig. 3e). The sliding polynomial fit method (Fig. 2c) shows a reduced spectral response for vertical wavelengths larger than $\approx 13\,\mathrm{km}$. For shorter vertical wavelengths the spectral

response is close to 100 % at most altitudes. Vertical wavelengths of $\approx 9\,\mathrm{km}$ show a slight reduction in spectral response over the entire altitude range. At the upper and lower $5\,\mathrm{km}$ of the analyzed altitude window vertical wavelengths larger than $5\,\mathrm{km}$ are strongly damped. The spectral response of the Butterworth filter (Fig. 2d) is very similar to the sliding polynomial fit. The main difference is that the Butterworth filter exhibits no underestimation of temperature perturbations at $9\,\mathrm{km}$ vertical wavelength.

Figure 2e and f shows mean extracted temperature perturbations. The blue line (here underneath the green line) depicts the original temperature perturbations added to the background temperature profile. As evident from Fig. 2e, the sliding polynomial fit method underestimates temperature perturbations at vertical wavelengths around $9\,\mathrm{km}$. In agreement with the filter design both vertical filtering methods, the sliding polynomial fit and the Butterworth filter, show a decrease in extracted temperature perturbations for vertical wavelengths larger than $13\,\mathrm{km}$. This decrease is almost linear with increasing vertical wavelength. As a consequence, amplitudes are effectively reduced by a factor of 3 at $\lambda_z = 20\,\mathrm{km}$.

In the first simulation setup the vertical wavelength λ_z was varied, while the period τ was kept constant. We now proceed by varying the period τ with a fixed $\lambda_z = 6\,\mathrm{km}$ (Fig. 3). The spectral response of the nightly mean method (Fig. 3a) is close to 100 % at all altitudes. Temperature perturbations with periods larger than $10\,\mathrm{h}$ are damped and periods around $6\,\mathrm{h}$ are slightly underestimated. For $\tau = 15\,\mathrm{h}$ the reduction in amplitude is $\approx 20\%$ (green line in Fig. 3e and f). Like the nightly mean method, the $3\,\mathrm{h}$ running mean (Fig. 3b) exhibits a uniform spectral response at all altitudes. However, waves with periods longer than $3.5\,\mathrm{h}$ are strongly damped. At a period of $6\,\mathrm{h}$ temperature perturbations are underestimated by a factor of 2 and for $\tau = 2.5\,\mathrm{h}$ amplitudes are overestimated by $\approx 20\%$ (orange line in Fig. 3e and f). The spectral response of the filter for waves with shorter periods oscillates between over- and underestimation as τ approaches zero. In contrast, the sliding polynomial fit method (Fig. 3c) and the Butterworth filter (Fig. 3d) both exhibit an almost uniform spectral response for most periods. Only for very long periods the spectral response oscillates between over- and underestimation with increasing altitude, indicating a slight phase delay between simulated and extracted temperature perturbations. This oscillation is not seen in Figure 3e and f due to the vertical averaging over 10 km.

3.2 Varying background temperature

While in the previous section the simulated background temperature was kept constant, we now examine the influence of a time dependent variation of the background temperature on the different methods. Slow variations of the form

$$T_0'(z,t) = \alpha\, t \sin\left(\frac{2\pi\,(z - z_0)}{60\,\mathrm{km}}\right) \exp\left(\frac{z - z_0}{H_0}\right) \tag{8}$$

were added to Eq. (5), where $\alpha = 0.5\,\mathrm{K\,h^{-1}}$ is the heating/cooling rate and $H_0 = 65\,\mathrm{km}$ is the scale height of the background temperature variation. This results in a warming of the stratosphere and a cooling of the mesosphere over time, representing a very simplified effect of a propagating planetary wave with a vertical wavelength of $60\,\mathrm{km}$. All other parameters are the same as before.

Filter characteristics are shown for a varying vertical wavelength in Fig. 4. Compared to the steady background simulations (e.g. Fig. 2), the nightly mean method exhibits an enhanced spectral response around 35 and $65\,\mathrm{km}$ altitude (Fig. 4a). From Fig. 4e it can be determined that the nightly mean method overestimates temperature perturbations by roughly $25\,\%$ between 30 and $40\,\mathrm{km}$ altitude. No change in spectral response is detected for the $3\,\mathrm{h}$ running mean method (Fig. 4b), the sliding polynomial fit method (Fig. 4c) and the Butterworth filter (Fig. 4d).

The filters exhibit similar characteristics if the gravity wave period is varied instead of the vertical wavelength. The nightly mean method (Fig. 5a) overestimates temperature perturbations in the same altitude bands as shown for the simulations with varying vertical wavelength (cf. Fig. 4a). The filter characteristics of the $3\,\mathrm{h}$ running mean method (Fig. 5b), the sliding polynomial fit method (Fig. 5c) and the Butterworth filter (Fig. 5d) are not affected by the varying background temperature.

4 Application to measurement data

Rayleigh lidar measurements at Lauder, New Zealand, (45.0° S, 169.7° E) were obtained with the TELMA instrument from mid June to mid November 2014 (Kaifler et al., 2015a). We use temperature data with a temporal resolution of $10\,\mathrm{min}$ and a vertical resolution of $100\,\mathrm{m}$. The effective vertical resolution of the temperature data is $900\,\mathrm{m}$ due to smoothing of the raw data before processing. Measurement uncertainties are typically on the order of 2–$3\,\mathrm{K}$ at $70\,\mathrm{km}$ altitude and generally lower than $1\,\mathrm{K}$ below $60\,\mathrm{km}$ altitude.

4.1 Case study: 23 July 2014

A detailed analysis with the four different methods of extracting temperature perturbations is shown for the dataset obtained on 23 July 2014 in Fig. 6. This case was chosen because the gravity wave analysis depicts many previously noted characteristics of the four methods.

The main features of the mean temperature profile (Fig. 6b) are the stratopause between 45 and $55\,\mathrm{km}$ altitude with $T \approx 245\,\mathrm{K}$ and the temperature minimum of approximately $200\,\mathrm{K}$ at $73\,\mathrm{km}$ altitude below a mesospheric inversion layer. The time evolution of the temperature measurements (Fig. 6a) shows an increase of the temperature at the stratopause and a jump in stratopause height around 08:00 UTC. Afterwards, the stratopause descends slowly. The structure of the mesospheric inversion layer varies also over time with the minimum temperature below the inversion layer reaching $\approx 175\,\mathrm{K}$ around 14:00 UTC.

The temperature perturbations as determined by the nightly mean method (Fig. 6c) exhibit a vertically broad maximum descending from about 80 km altitude down to 50 km altitude over the 12 h measurement period. Temperature perturbations within this descending maximum reach values of up to ± 20 K. Below 50 km altitude temperature perturbations are generally on the order of ± 5 K.

The 3 h running mean method on the other hand (Fig. 6d) shows strongly tilted patterns. Below 50 km altitude the phase lines tend to be steeper than above. The magnitude of the temperature perturbations generally increases with altitude from approximately ± 5 K below 60 km altitude to approximately ± 15 K above 60 km altitude.

The sliding polynomial fit method (Fig. 6e) and the Butterworth filter (Fig. 6f) extract almost identical patterns of temperature perturbations, with the Butterworth filter inferring slightly larger amplitudes. The phase lines in the Fig. 6e and f decrease more slowly in altitude compared to the 3 h running mean method. Below 60 km altitude temperature perturbations are below ± 10 K for both filters and increase to ± 15 K above 60 km altitude.

4.2 Statistical performance

A quantity often used as a proxy for gravity wave activity is the gravity wave potential energy density (GWPED) per mass

$$E_p = \frac{1}{2}\frac{g^2}{N^2}\overline{\left(\frac{T'}{T_0}\right)^2}, \text{ with} \tag{9}$$

$$N^2 = \frac{g}{\overline{T_0}}\left(\frac{\mathrm{d}\overline{T_0}}{\mathrm{d}z} + \frac{g}{c_p}\right), \tag{10}$$

where g denotes the acceleration due to gravity and c_p the heat capacity of dry air under constant pressure, in addition to the previously defined variables. The mean GWPED is determined as the average over one measurement period – typically 5–12 h in our case – which is denoted by the overline in Eq. (9). Due to the decrease in density with altitude, GWPED per mass increases exponentially with altitude in the case of conservative wave propagation. For a more detailed description and physical interpretation of the GWPED see e.g. Rauthe et al. (2008) and Ehard et al. (2014).

From TELMA observations above New Zealand over the period 1 July 2014 to 30 September 2014 we determined the mean GWPED per mass using the four methods of gravity wave extraction discussed in this study (Fig. 7). Relative uncertainties of the GWPED for all methods are on the order of 0.5% in the stratosphere and increase to approximately 5% at 80 km altitude which is considerably smaller than the variations of the GWPED due to the geophysical variability. The absolute value of the GWPED varies by as much as one order of magnitude depending on which method is used. The largest relative deviations appear in the lower stratosphere between the 3 h running mean method and the Butterworth filter. Above 65 km altitude all methods produce similar results. A distinct feature of Fig. 7 is the larger growth of GWPED with altitude if the running mean method is used instead of the vertical filtering methods. Additionally, the 3 h running mean method yields the lowest GWPED

values. If a 4 h running mean is used instead, the GWPED profile is shifted towards slightly larger values. Below 45 km altitude the nightly mean method produces values comparable to the sliding polynomial fit and the Butterworth filter. Above 45 km altitude the nightly mean method shows the largest values of all methods. The sliding polynomial fit and the Butterworth filter produce generally similar results, with the Butterworth filter yielding a slightly larger GWPED. Another striking feature in Fig. 7 is the ~~increase in~~ enhanced GWPED below 35 km altitude which is detected by both vertical filtering methods. This ~~increase~~ enhancement is not detected by the running mean method.

5 Discussion

5.1 Temporal filters

The nightly mean method has been applied in many studies (e.g. Gardner et al., 1989; Blum et al., 2004; Rauthe et al., 2008; Ehard et al., 2014). The major disadvantage is that a varying length of measurement periods results in a variation of the sensitivity to different timescales. This effect is clearly demonstrated in Fig. 3e showing that gravity waves with periods larger than 10 h are significantly underestimated if an 8 h long timeseries is used. If the timeseries is shortened, the cutoff period is smaller as well (not shown) and the spectral response for long period waves is reduced even further. Strictly speaking, this implies that gravity wave analyses of timeseries of different length cannot be compared.

In practice measurement periods vary typically in length between a few hours up to a whole night as weather conditions can change rapidly during an observational period. Moreover, there is a seasonal dependency because most middle atmospheric lidars are capable of measuring in darkness only. This results in shorter measurement periods in summer and longer measurement periods in winter. Hence, the nightly mean method is sensitive to different parts of the gravity wave spectrum depending on weather conditions as well as season. For example, Rauthe et al. (2006) compared winter and summer measurements of gravity wave activity determined by the nightly mean method. ~~Their winter measurements were restricted to observational~~ They resolved gravity waves with periods of 1.5–12 h ~~and the summer measurements to observational periods of~~ during winter and 1.5–3.5 h during summer. Hence, Rauthe et al. (2006) limited their analysis to 3–5 h long measurement periods in order to reduce the variation of the spectral response.

The use of the nightly mean method in gravity wave analysis is further complicated by the fact that there are processes besides gravity waves which occur on similar timescales. For example tides with periods of 8, 12 and 24 h are within the sensitivity range of this method. In the analysis of radar data, the removal of tidal signals is a standard procedure (e.g. Hoffmann et al., 2010). With lidar data, however, this is problematic due to generally shorter and often intermitted measurement periods. Figure 6c shows an example of a tidal signal extracted with the nightly mean method. The broad descending maximum in temperature perturbations is caused by the semidiurnal tide, which

was confirmed by a composite analysis over several days (not shown). Note, that the nightly mean method is not a suitable method for tidal analysis. Tidal signals are generally extracted from lidar measurements by means of ~~a~~ the previously mentioned composite analysis (e.g. Lübken et al., 2011).

The running mean method (e.g. Yamashita et al., 2009) tries to compensate for some of the shortcomings of the nightly mean method. The spectral response is limited to timescales on the order of the window width of the running mean – which is typically 3 h – resulting in the suppression of tides and planetary waves. However, due to this limitation, only a very small part of the gravity wave spectrum is retained in the analysis (e.g. Fig. 3e). As stated previously, gravity wave periods can range from about 5 min to 17 h. Thus the limitation to short timescales excludes a major part of the gravity wave spectrum. Figure 7 shows that as the length of the running mean window increases, the GWPED increases as well. Still, gravity waves with long periods are suppressed. Additionally, the running mean method overestimates periods slightly shorter than the chosen window width ~~and shows aliasing effects for even shorter~~ (Fig. 3e). The strongly oscillating spectral response of the running mean method for short periods (Fig. 3e) ~~.~~ arises due to the coarse temporal resolution of 0.5 h used in the simulations, which is a typical temporal resolution of lidar measurements. If the temporal resolution of the simulations is increased these sharp peaks for periods shorter than 1 h vanish (not shown).

The beginning and the end of the measurement period poses an additional problem for the application of the running mean method. At the beginning of the measurement period, a centered running mean of 3 h lacks the first 1.5 h of observations necessary for determining the background temperature. Thus, if in the beginning of the measurement only 1.5 h of data are available for averaging, the spectral response differs at the beginning of the measurement period compared to later times when 3 h of measurements are available. The same is true at the end of the measurement period as well as in the presence of measurement gaps. Thus, when requiring the same spectral response at all times, the "spin-up" time of the running mean method would have to be discarded. However, this would result in a significantly reduced dataset because one window width of data would have to be discarded from each measurement period, in addition to another window width for each measurement gap.

Note that the resolved high frequency range of the gravity wave spectrum is limited by the sampling frequency of the lidar system which ranges typically between 10 min and 1 h, depending on lidar performance. This is a fundamental limitation to the extractable part of the gravity wave spectrum which affects all methods of extracting gravity wave induced temperature perturbations in the same way. The same holds true for the effective vertical resolution of the temperature profiles.

5.2 Spatial filters

Filtering in the spatial domain, either by using the sliding polynomial fit or the Butterworth filter, has the advantage that the spectral response in the time domain is independent of the length of the measurement period and the presence of measurement gaps. This makes it possible to derive

temperature perturbations associated with gravity waves from observational periods which are too short to yield meaningful results if temporal filtering methods are applied. In addition, both spatial filtering methods are capable of detecting waves with periods larger than $12\,\mathrm{h}$ (Fig. 3c and d). One disadvantage of both spatial filtering methods is the dampening of vertical wavelengths larger than $5\,\mathrm{km}$ at the upper and lower edge of the analyzed altitude window due to edge effects.

The sliding polynomial fit has been applied in several studies (e.g. Duck et al., 2001; Alexander et al., 2011; Kaifler et al., 2015b). Different authors use temperature data with different altitude resolutions and slightly different parameter setups for L_f, L_w and γ. The fit length L_f determines the cutoff wavelength of the spectral response. The weighting window length L_w and the e-folding width γ must be adapted to the altitude resolution of the data used. For example, the parameter setup $\gamma = 3\,\mathrm{km}$ and $L_\mathrm{w} = L_\mathrm{f}/3$ used by Duck et al. (2001) results in a flat spectral response for their altitude resolution of $\Delta z = 2\,\mathrm{km}$ and fit length $L_\mathrm{f} = 25\,\mathrm{km}$. If a different altitude resolution is chosen, a different set of parameters is needed in order to achieve a flat spectral response in the passband. For the altitude resolution of $\Delta z = 0.1\,\mathrm{km}$ used in this study, a flat spectral response was found for $\gamma = 9\,\mathrm{km}$ and $L_\mathrm{w} = 3\,\mathrm{km}$. However, vertical wavelengths of $\approx 9\,\mathrm{km}$ are still slightly underestimated with this parameter set. The fit length of $L_\mathrm{f} = 20\,\mathrm{km}$ was chosen following Kaifler et al. (2015b). Additional high-pass filtering, as applied by Alexander et al. (2011) or Kaifler et al. (2015b), was found to be unnecessary because the long vertical wavelengths are already strongly suppressed by the sliding polynomial fit itself.

The sliding polynomial fit method is sensitive to large changes of the temperature gradient and may falsely overestimate temperature perturbations for example in the presence of mesospheric inversion layers (not shown). The Butterworth filter tends to overestimate sudden changes in the temperature gradient of the measured temperature profile as well. However, the magnitude of the overestimation is generally lower than for the sliding polynomial fit method. Furthermore, the Butterworth filter has the advantage that it can be easily adjusted if a different cutoff wavelength is desired.

5.3 Application to measurement data

All the previously discussed characteristics influence the gravity wave spectrum which is extracted from lidar temperature measurements. This becomes visible if the mean GWPED of a set of measurements is computed using different methods as shown in Fig. 7. The running mean method extracts only a small part of the gravity wave spectrum and thus shows the lowest GWPED values. The GWPED increases if the window width of the running mean is increased. The nightly mean method yields the largest GWPED values at higher altitudes. This can be attributed to the insufficient suppression of tides and other processes unrelated to gravity waves which happen on longer timescales. In the lower stratosphere the sliding polynomial fit method and the Butterworth filter yield the largest GWPED values. This is most likely caused by the inclusion of long period waves such as quasi-stationary mountain waves. These waves have the largest impact on GWPED in the

lower stratosphere above Lauder during winter (Kaifler et al., 2015a). Above 30 km altitude GWPED values are reduced. A possible mechanism is that mountain waves with very large amplitudes become unstable at these altitudes and break. This has for example been observed by Ehard et al. (2015) who detected a self-induced critical layer around 30 km altitude caused by a strong mountain wave event above northern Scandinavia.

The fact that the Butterworth filter exhibits a lower growth rate of GWPED compared to the running mean method (Fig. 7) may ~~by~~ be evidence that short period gravity waves can propagate more easily to higher altitudes than gravity waves with long periods. This complicates the comparison and interpretation of GWPED growth rates (generally expressed in terms of scale heights) of different studies. For example Rauthe et al. (2006) deduced a GWPED scale height of 9–11 km with the nightly mean method for a mid-latitude site. On the other hand, Kaifler et al. (2015b) reported a GWPED scale height of approximately 7 km determined with the sliding polynomial fit method for measurements conducted at Antarctica. A large part of the difference in retrieved scale height can be attributed to different wave propagation conditions at the two sites. However, it remains an open question in how far the results are affected by the use of different methods to extract gravity waves.

6 Conclusions

We evaluated four commonly used methods of extracting gravity wave induced temperature perturbations from lidar measurements. A widely used method – the nightly mean method – relies on filtering in time by subtraction of the nightly mean temperature. Thereby, it is sensitive to all temperature changes occurring on the timescale of the measurement period including temperature changes induced by planetary waves and tides. Because measurement periods can vary substantially in length and the spectral response of the nightly mean method depends on the length of the measurement period, the extracted gravity wave spectrum can vary from observation to observation. This makes the nightly mean method an improper choice for compiling gravity wave statistics if a dataset with a varying length of observational periods is analyzed.

The second method which relies on filtering in time, the running mean method, provides a more stable spectral response with regard to a varying length of the measurement period. However, it extracts only a small fraction of the gravity wave spectrum, with long period waves being strongly suppressed. Moreover, the running mean method exhibits a variation in the spectral response at the beginning and end of a measurement period as well as in the presence of measurement gaps.

The sliding polynomial fit method is not only capable of extracting waves over a broad range of temporal scales but also suppresses tides and planetary waves due to their large vertical wavelengths. In addition, it is unaffected by measurement gaps. However, the parameters used for the sliding

polynomial fit need to be adjusted to the altitude resolution of the measured temperature profiles in order to provide a flat spectral response in the passband.

The Butterworth filter provides an alternative to the sliding polynomial fit method which is not only easy to implement but also easily adjustable to a desired cutoff wavelength. Also, the filter is largely independent of the altitude resolution while providing all the advantages of the sliding polynomial fit method. Furthermore, sudden changes in the background temperature gradient affect the Butterworth filter less than the sliding polynomial fit method.

Based on the results presented here, two methods are recommended for gravity wave extraction from lidar temperature measurements covering a large altitude range: the running mean method is the most suitable method if the analysis is focused on short period gravity waves with large vertical wavelengths. On the other hand, if a broad passband is desired which covers a large part of the gravity wave spectrum, the Butterworth filter is the method of choice. Additional advantages are the insensitivity to measurement gaps, a varying length of observational periods and the altitude resolution of the measured temperature profile.

Acknowledgements. We thank the NIWA personnel at Lauder station for their support during the DEEPWAVE measurement campaign. This work was supported by the project “Processes and Climatology of Gravity Waves” (PACOG) in the framework of the research unit “Mesoscale Dynamics of Gravity Waves” (MS-GWaves) funded by the German Research Foundation.

The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

References

Alexander, S., Klekociuk, A., and Murphy, D.: Rayleigh lidar observations of gravity wave activity in the winter upper stratosphere and lower mesosphere above Davis, Antarctica (69°S, 78°E), J. Geophys. Res., 116, D13 109, doi:10.1029/2010JD015164, 2011.

Alpers, M., Eixmann, R., Fricke-Begemann, C., Gerding, M., and Höffner, J.: Temperature lidar measurements from 1 to 105 km altitude using resonance, Rayleigh, and Rotational Raman scattering, Atmos. Chem. Phys., 4, 793–800, doi:10.5194/acp-4-793-2004, 2004.

Blum, U., Fricke, K. H., Baumgarten, G., and Schöch, A.: Simultaneous lidar observations of temperatures and waves in the polar middle atmosphere on the east and west side of the Scandinavian mountains: A case study on 19/20 January 2003, Atmos. Chem. Phys., 4, 809–816, doi:10.5194/acpd-4-969-2004, 2004.

Chane-Ming, F., Molinaro, F., Leveau, J., Keckhut, P., and Hauchecorne, A.: Analysis of gravity waves in the tropical middle atmosphere over La Reunion Island (21°S, 55°E) with lidar using wavelet techniques, Ann. Geophys., 18, 485–498, doi:10.1007/s00585-000-0485-0, 2000.

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

Duck, T., Whiteway, J., and Carswell, A.: The Gravity Wave-Arctic Stratospheric Vortex Interaction, J. Atmos. Sci., 58, 3581–3596, 2001.

Ehard, B., Achtert, P., and Gumbel, J.: Long-term lidar observations of wintertime gravity wave activity over northern Sweden, Ann. Geophys., 32, 1395–1405, doi:10.5194/angeo-32-1395-2014, 2014.

Ehard, B., Achtert, P., Dörnbrack, A., Gisingar, S., Gumbel, J., Khaplanov, M., Rapp, M., and Wagner, J.: Combination of lidar and model data for studying deep gravity wave propagation, Mon. Wea. Rev., doi:10.1175/MWR-D-14-00405.1, submitted, 2015.

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

Gardner, C. S., Miller, M. S., and Liu, C. H.: Rayleigh Observations of Gravity Wave Activity in the Upper Stratosphere at Urbana, Ilinois, J. Atmos. Sci., 46, 1838–1854, 1989.

Hauchecorne, A. and Chanin, M.: Density and temperature profiles obtained by lidar between 35 and 70 km, Geophys. Res. Lett., 7, 565–568, doi:10.1029/GL007i008p00565, 1980.

Hoffmann, P., Becker, E., Singer, W., and Placke, M.: Seasonal variation of mesospheric waves at northern middle and high latitudes, Journal of Atmospheric and Solar-Terrestrial Physics, 72, 1068–1079, doi:10.1016/j.jastp.2010.07.002, 2010.

Holton, J. and Alexander, M.: The role of waves in the transport circulation of the middle atmosphere, Geophys. Monogr., 123, 21–35, 2000.

Kaifler, B., Kaifler, N., Ehard, B., Dörnbrack, A., Rapp, M., and Fritts, D.: Influences of source conditions on mountain wave penetration into the stratosphere and mesosphere, Geophys. Res. Lett., doi:10.1002/2015GL066465, 2015a.

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

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

Lübken, F.-J., Höffner, J., Viehl, T., Kaifler, B., and Morris, R.: First measurements of thermal tides in the summer mesopause region at Antarctic latitudes, Geophys. Res. Lett, 38, L24 806, doi:10.1029/2011GL050045, 2011.

Mzé, N., Hauchecorne, A., Keckhut, P., and Thetis, M.: Vertical distribution of gravity wave potential energy from long-term Rayleigh lidar data ata northern middle-latitude site, J. Geoph. Res. Atmos., 119, doi:10.1002/2014JD022035, 2014.

Plougonven, R. and Zhang, F.: Internal gravity waves from atmospheric jets and fronts, Rev. Geophys., 52, 33–76, doi:10.1002/2012RG000419, 2014.

Rauthe, M., Gerding, M., Höffner, J., and Lübken, F.-J.: Lidar temperature measurements of gravity waves over Kühlungsborn (54°N) from 1 to 105 km: A winter-summer comparison, J. Geophys. Res., 111, D24 108, doi:10.1029/2006JD007354, 2006.

Rauthe, M., Gerding, M., and Lübken, F.-J.: Seasonal changes in gravity wave activity measured by lidars at mid-latitudes, Atmos. Chem. Phys., 8, 6775–6787, doi:10.5194/acp-8-6775-2008, 2008.

Sica, R. and Haefele, A.: Retrieval of temperature from a multiple-channel Rayleigh-scatter lidar using optimal estimation method, Appl. Opt., 54, 1872–1889, doi:10.1364/AO.54.001872, 2015.

Sica, R. and Russell, A.: Measurements of the effects of gravity waves in the middle atmosphere using parametric models of density fluctuations. Part I: vertical wavenumber and temporal spectra, J. Atmos. Sci., 56, 1308–1329, 1999.

Thurairajah, B., Collins, R. L., Harvey, V. L., Liebermann, R. S., and Mizutani, K.: Rayleigh lidar observations of reduced gravity wave activity during the formation of an elevated stratopause in 2004 at Chatnika, Alaska (65°N, 147°W), J. Geophys. Res., 115, D13 109, doi:10.1029/2009JD013036, 2010.

Whiteway, J. A. and Carswell, A. I.: Lidar observations of gravity wave activity in the upper stratosphere over Toronto, J. Geophys. Res., 100, 14 113–14 124, doi:10.1029/95JD00511, 1995.

Wilson, R., Chanin, M., and Hauchecorne, A.: Gravity waves in the middle atmosphere observed by Rayleigh lidar: 2. Climatology, J. Geophys. Res., 96, 5169–5183, doi:10.1029/90JD02231, 1991.

Yamashita, C., Chu, X., Liu, H., Espy, P. J., Nott, G. J., and Huang, W.: Stratospheric gravity wave characteristics and seasonal variations observed by lidar at the South Pole and Rothera, Antarctica, J. Geophys. Res., 114, D12 101, doi:10.1029/2008JD011472, 2009.

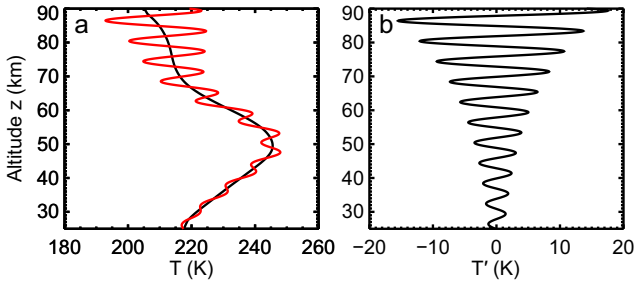

Figure 1. (a) Background temperature profile T_0 used for the simulations (black) and perturbed temperature profile T (red). (b) The temperature perturbations T' added to T_0. Temperature perturbations in both panels were constructed using Equation 6 with the following set of parameters: $t = 4$ h, $A = 1.2$ K, $\lambda_z = 6$ km, $\tau = 1.9$ h, $H = 12$ km.

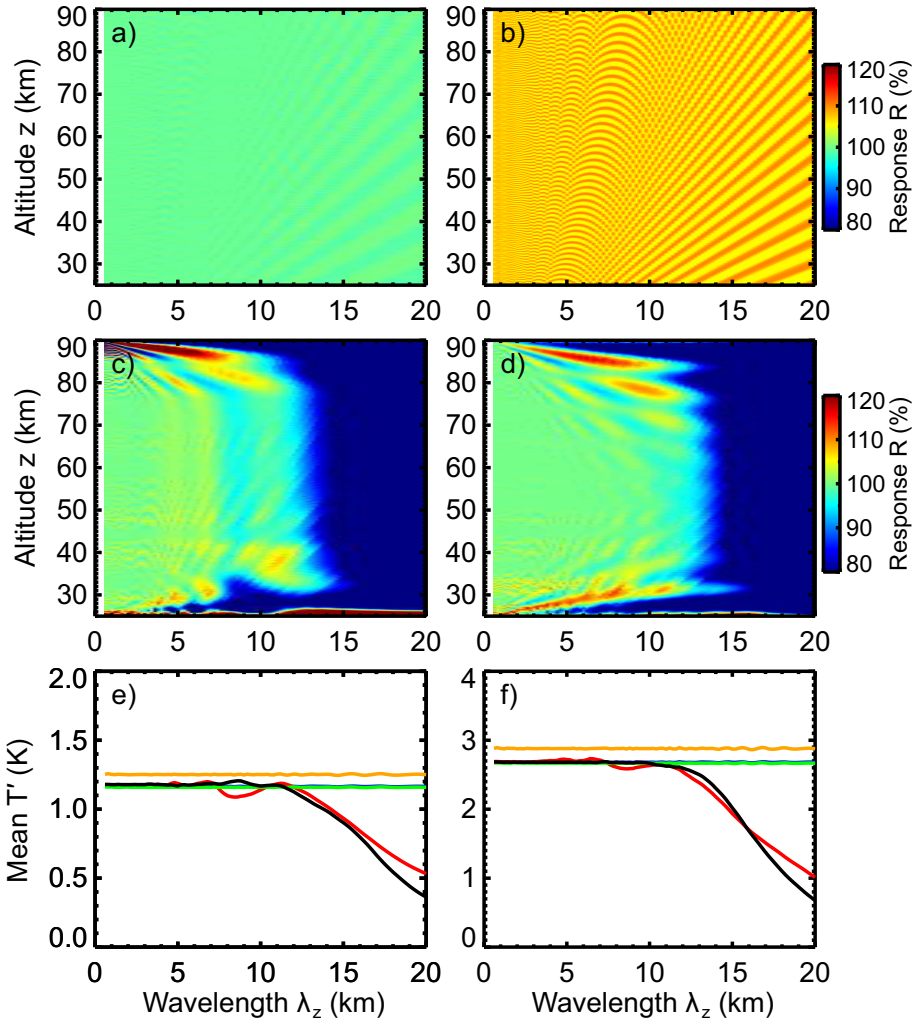

Figure 2. Spectral response of different methods of determining temperature perturbations as a function of vertical wavelength λ_z: Nightly mean (a), 3 h running mean (b), sliding polynomial fit (c) and Butterworth filter (d). Panels (e) and (f) depict mean extracted temperature perturbations between 30–40 km (e) and 50–60 km (f) as well as the simulated temperature perturbations (blue line). The different methods are color coded as follows: Nightly mean – green, 3 h running mean – orange, sliding polynomial fit – red, Butterworth filter – black. Please note that the blue line in this case lies exactly underneath the green line. All simulations were carried out with $\tau = 1.9$ h and a background temperature profile constant in time.

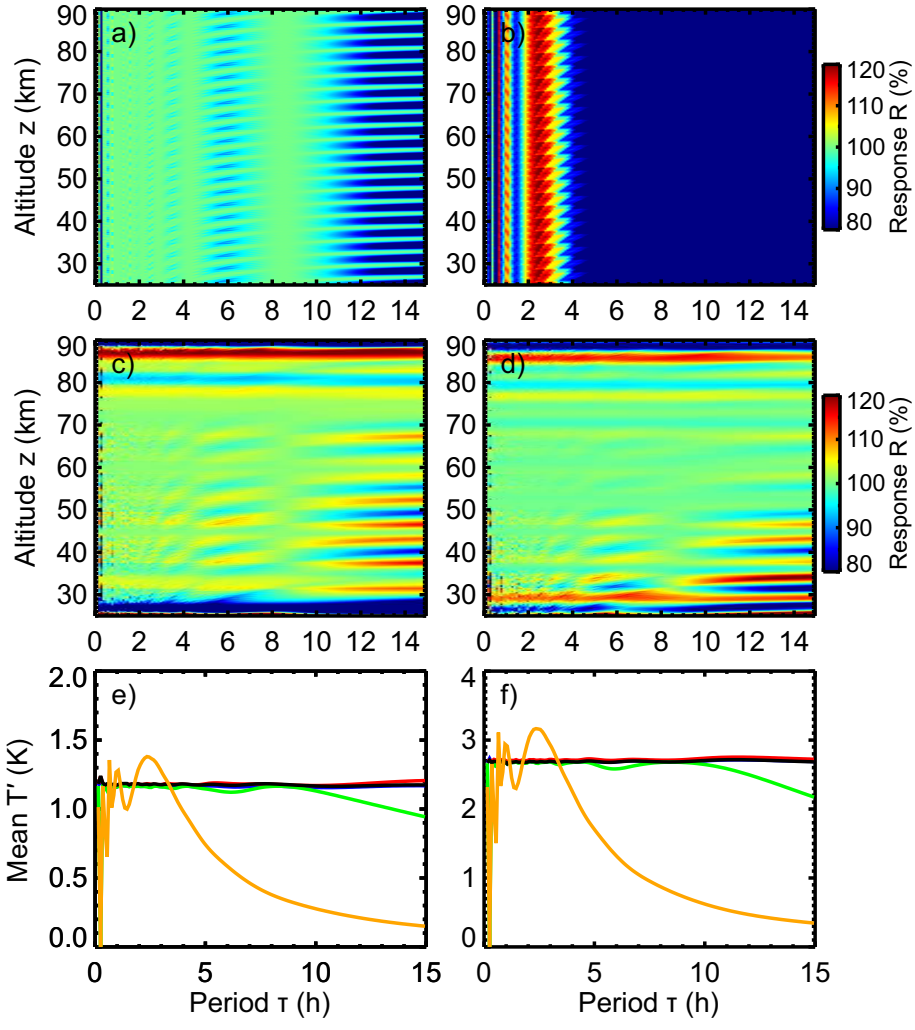

Figure 3. Same as Figure 2 but as a function of period τ. All simulations were carried out with a fixed vertical wavelength of 6 km and a background temperature profile constant in time. Note that the blue and black lines in panels (e) and (f) are lying on top of each other.

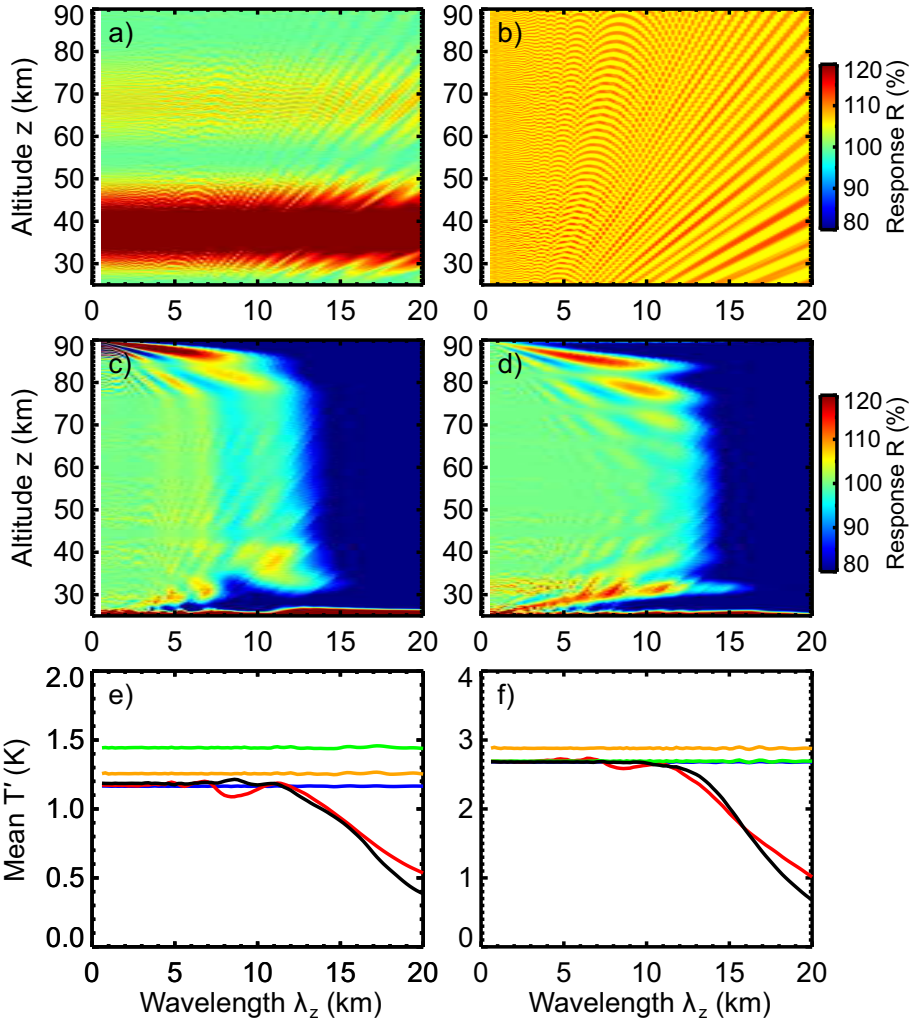

Figure 4. Same as Figure 2 but with a varying background temperature (see Section 3.2 for details).

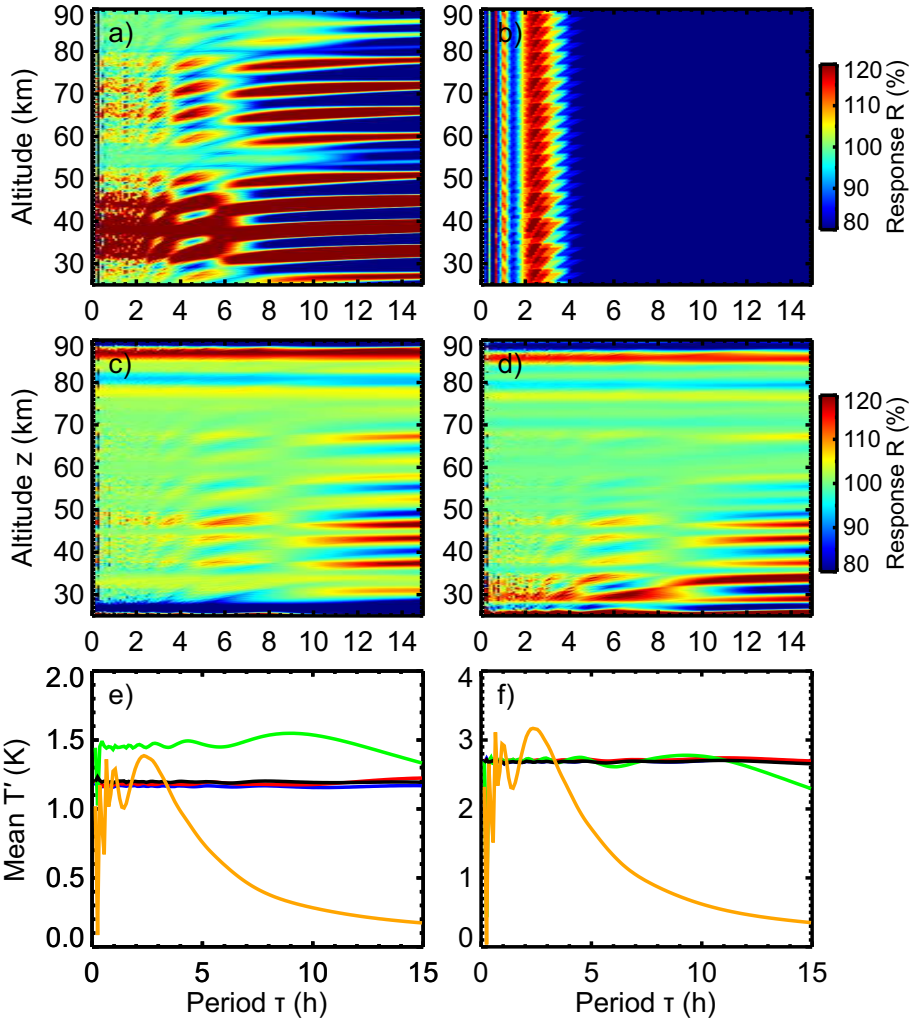

Figure 5. Same as Figure 3 but with a varying background temperature (see Section 3.2 for details).

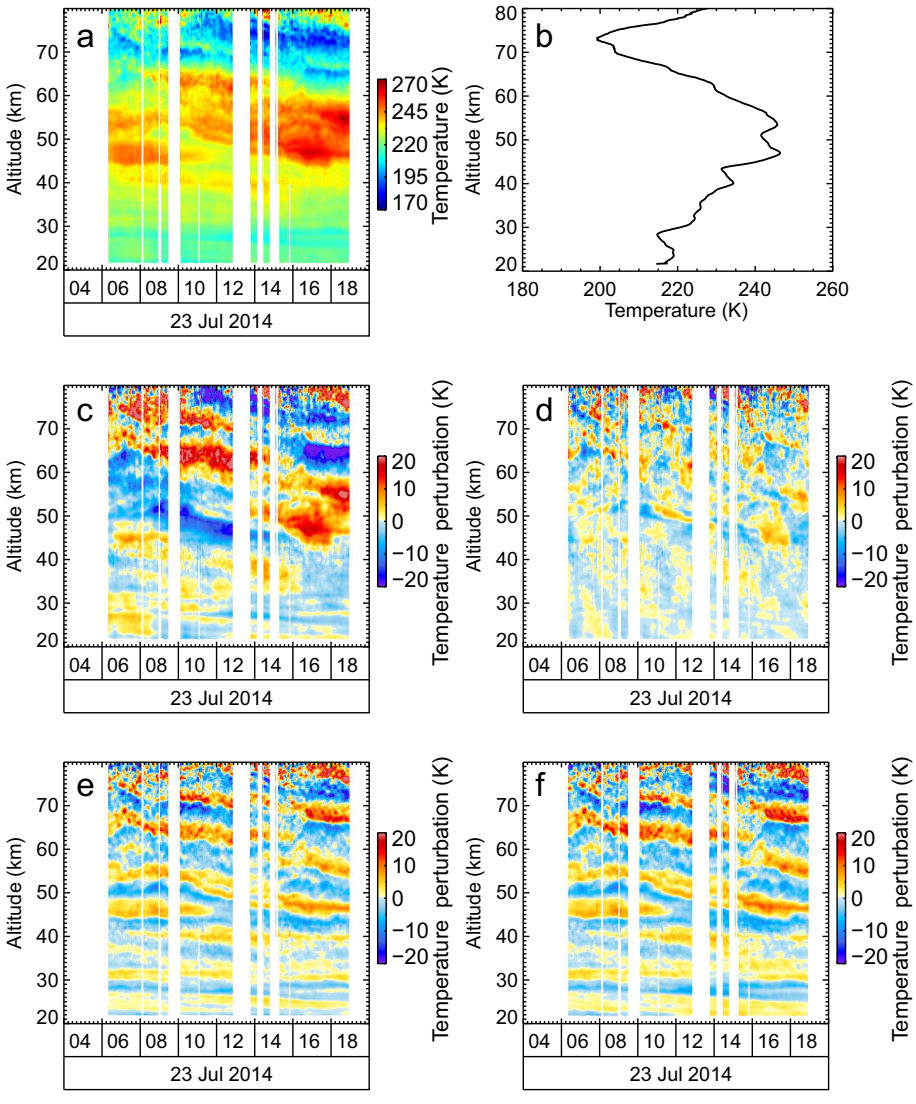

Figure 6. Temperature (a), mean temperature profile (b) and derived temperature perturbations obtained by different methods (c)–(f) over Lauder, New Zealand, (45.0° S, 169.7° E) on 23 July 2014. The following methods were used for the different panels: Nightly mean (c), 3 h running mean (d), sliding polynomial fit (e), Butterworth filter (f). Time is given in UTC.

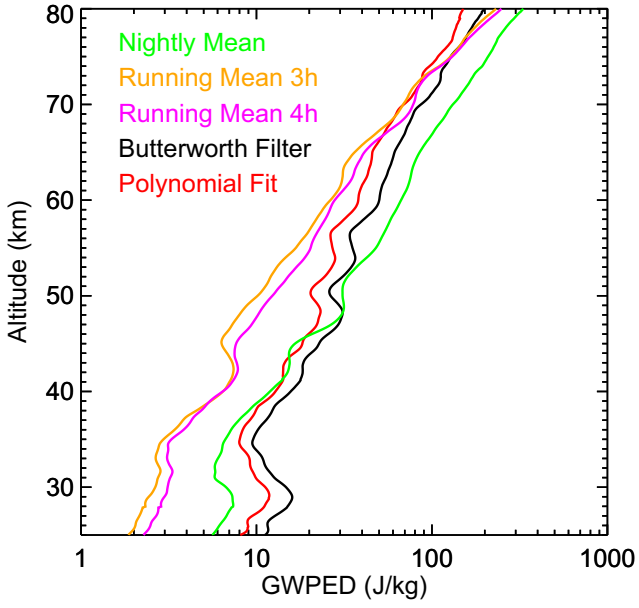

Figure 7. Mean gravity wave potential energy density (GWPED) per mass over Lauder, New Zealand, (45.0° S, 169.7° E) between 1 July and 30 September 2014. The methods used to determine the GWPED are color coded. The profiles were smoothed by a vertical running mean with a window width of 3 km.