Review comments for "Smoothing data series by means of cubic splines: quality of approximation and introduction of an iterative spline approach" by Sabine Wüst et al.

The current paper introduces an alternative way of applying cubic splines for smoothing data series. Based on simple test data, the authors first showed the dependence of the approximation error on the sampling distance for the traditional way of applying cubic splines (Fig. 3). It was shown that when the sampling distance is equal or close to half of the wavelength, the approximation error does not decrease smoothly with decrease of the sampling distance, but oscillates strongly. In this case, the approximation error also depends very strongly on the phase of the wave/ exact positions of the sampling points (Fig. 4). In addition, artifical oscillations can appear in this case (Fig. 2b, c).

The authors suggested an alternative method (iterative approach) of applying cubic splines, which can reduce these phenomena. Indeed, using this iterative approach, the approximation error decreases quite smoothly with decrease of the sampling distance (Fig. 6a). Moreover, the dependence of the approximation error on exact positions of sampling points also reduces significantly (Fig. 6b). However, it should be noted that this improvement is only helpful for the sampling distances, which are equal or close to half of the wavelength. Another important note is that the amplitude of the smoothed series (and therefore of the residuals) changes significantly.

The iterative approach was also applied to another simple test dataset, which was created by superimposing 3 oscillations with different wavelengths on a realistic temperature background profile. Again, the approximation error decreases smoothly with decrease of the sampling distance. Finally, the iterative approach was applied to vertical temperature profiles measured by the SABER instrument. The residuals are then calculated and averaged over one year for the years 2010-2014. Results and their implication for gravity wave (GW) studying was discussed.

The paper is generally well written and has a clear structure. Introduction, Method description and Discussion are of appropriate length. My major and minor comments for this paper are given below:

Major comments

1. The iteractive approach has advantages of stabilizing the traditional approach: (1) the approximation error decreases smoothly with decrease of the sampling distance, and (2) the approximation error depends very slightly on the exact locations of sampling points. However, these advantages are only helpful if the sampling distance is equal or close to half of the wavelength.

   In practice, a constant sampling distance is chosen for all measurement profiles (for example, the authors chose 10 km for all SABER profiles in this study). It means the iteractive approach may be advanced for only some profiles, which contain waves with wavelength of $\sim$ 20 km. For other waves in the same profile or for other waves in other profiles, which have wavelengths different from $\sim$ 20 km, I would not expect advantages of the iteractive approach. It is well known that GWs have a very broad spectrum. Therefore, wavelengths, which are equal or close to a certain wavelength, can only be a very small part of that broad spectrum.

   It is therefore interesting to see how much this stabilizing contributes to the total approximation error:

(a) Starting with the test data, where the 3 oscillations of different wavelengths were superimposed on a realistic CIRA temperature background, can you please compare the approximation error of the traditional approach and the approximation error of the iteractive approach? (Using a constant sampling distance of 10 km, for example)

(b) Similarly for SABER profiles per year: For each profile, please calculate the approximation error δ_{1i} for traditional approach and δ_{2i} for iteractive approach, using a constant sampling distance of 10 km. For one year, the total approximation error is $\Delta_1 = \sum_i \delta_{1i}$ for traditional approach and $\Delta_2 = \sum_i \delta_{2i}$ for iteractive approach ($\sum_i$ means sum over all profiles of one year). Can you please compare Δ_1 and Δ_2 for the years 2010-2014?

2. The term called "approximation error" by the authors has the meaning of error in the case study, where the test data are described by an artificial sine. If an oscillation is superimposed on a realistic CIRA temperature background, the error of GW activity should be the difference between the residual (after the spline fit) and the original oscillation you used for superimposing. How does this error of GW activity vary with sampling distance for non-iteractive and iteractive approaches, for different superimposed wavelengths?

3. When we average the temperature residuals over the whole year for enough number of profiles, due to the arbitrary distribution of phases and amplitudes of GWs, the non-squared mean residual should be approximately zero. However, as shown by Fig. 8d, the iteractive method still produces non-zero amplitude oscillations in the temperature residuals. In page 8, lines 20-22, the authors suggested that changing in temperature gradients could be one of the reasons for this problem. However, if it is the main reason, we should see non-zero amplitude oscillations only in limited altitude regions near the stratopause or mesopause, but not in the entire altitude range in Fig. 8d. There seem to be systematic errors that have not been removed. Can you please comment on this?

4. The authors recommended the iteractive approach for estimating squared residuals for studying GW activity. Perhaps, the most convincing way to demonstrate if the iteractive approach is suitable for GW studies, is comparing the GW squared temperature derived by this approach to the one derived by another method, using the same original data. For example, Ern et al. (2011) showed zonal averages of GW squared temperature for SABER measurements. For your method, taking the sampling distance of 10 km allows for vertical wavelengths up to 20 km. This can cover the main part of the GW spectrum in the stratosphere and mesosphere. It would be very interesting to see if the zonal averages provided by this iteractive approach is similar to the ones in Ern et al. (2011).

Minor comments

1. Page 2, Lines 5-7: "Conclusions about ... sampling points used". The Shannon's sampling theorem described in Appendix 3 of Gubbins, 2004 is about reconstruction of the original time series from its samples. Since you use a constant sampling distance through each entire data series, it is rather the sampling distance between two samples, which decides which shortest wavelength can be resolved. It is rather about the Nyquist theorem than the Shannon's sampling theorem. I would suggest to write straightforward that: "The shortest wavelength/period which can be resolved by the spline is twice of the sampling distance according to the Nyquist theorem."

2. Page 2, Line 15: "between 48°N ...and 15°O". Please shortly explain here why did you choose this region?

3. Page 2, Lines 19-22: "The distance of 10 km ... up to 20 km". At this point, it is not straightforward for all readers to understand why choosing 10 km distance will lead to maximum wavelength of 20 km. Later, in Sect. 2, paragraph 2, you explained this very nicely. I suggest to move the explanation here and shortly refer to it again in Sect. 2 later.

4. Page 2 Line 27 and Page 7, Line 33: For the motivation of this paper, the authors used Fig. 1, where squared temperature residuals were averaged for the whole year. An oscillation with a wavelength of about 10 km is found and the authors suggested that this oscillation is an artefact of the non-iteractive spline approach. Howerver, due to seasonal variations of GW sources and of the background wind, GW activity over the same location at a certain altitude can be different for different seasons. Therefore, the oscillation with a wavelength of about 10 km could also be an artefact of averaging. Can you please comment on this?

5. Page 3, Line 12: "Depending on ... approximated". This sentence is not clear to me. I guess the authors mean: "If we want to approximate the entire data series, the length of the data series must be an integer number of the distance between 2 spline sample points. That is why only certain distances between two consecutive spline sampling points can be chosen if the whole data series is approximated". If this is what the authors want to say, I suggest to rewrite this sentence. This also makes the next paragraph become more understandable.

6. Page 5, Line 7: "gravity waves ... ". Please cite the paper of Ern et al. (2011), which provides a comprehensive GW data set derived from SABER measurements.

7. Page 8, Line 20: Please clarify the "height regions"

Technical corrections

Page 2, Line 15 and Page 13, Lines 6, 7: °O and 15°O → 5°E and 15°E
Page 7, Line 15: might be surprising → is not straightforward
Page 8, Line 21: one reason → one of the reasons for
Page 19, Line 5: there are no dashed-dotted lines in Fig. 5e, f.

References

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

Answer to

Review comments for “Smoothing data series by means of cubic splines: quality of approximation and introduction of an iterative spline approach” by Sabine Wüst et al.

First of all, we would like to thank you for your valuable comments.
In order to avoid confusion, we would like to mention already here that we proposed to use “repeating” spline algorithm since the other reviewer had problems with the term “iterative”.
The attached manuscript with marked changes includes also the changes due to the comments of the other reviewer.

Major comments:

1. (a) As proposed, we compare the approximation error of both approaches based on the test data, where the three oscillations of different wavelengths were superimposed on a realistic CIRA temperature background. In all subfigures of figure 7, additional curves for the traditional approach are added. In order to make the pictures easier to read, we changed the wave amplitudes from 0.5 to 2.

1. (b) We already provided the sum of squared residuals which we denote as approximation error in figure 8) and b) as well as in figure 1 for a constant sampling point distance of 10 km.

2. Actually, the dashed line in figure 7a describes a spline approximation (repeating approach, difference between spline sampling points 10 km) of the background and not the original background. We are sorry for this non-optimal figure description in the original manuscript. However, one can see that the background is already very well adapted for this sampling point distance. Main reason for the variation of the "approximation error” is the increasing adaption of the superimposed oscillations (see also figure 5 and 6). When calculating the difference between the residuals and the original oscillation for different sampling point distances, the corresponding does not look not very instructive any more from our point of view. When the spline is able to approximate one of the original oscillations, the residuals do not contain this oscillation any more. Subtracting the original oscillation from the residuals in this case, leads to a larger difference than before. That means “your” approximation error grows again.
   Since the term “approximation error” is ambiguous (also the other reviewer seems having problems with it, see point 1 of his major comments), we leave it out now in the entire manuscript and call it sum of squared residuals.

3. We made some further tests in order to answer your question. We de-trended the CIRA temperature values between ca. 14 and 110 km height (as we did it for the SABER data) with three different sampling point distances (2.5 km, 5 km and 10 km). CIRA data do not show small-scale fluctuations and so we can test how well the spline adaption works for a smooth background. For the sampling point distances of 2.5 and 5 km, it becomes clear that the amplitude of the oscillation shows local maxima at the upper and lower border (see figure 1). For a sampling pont distance of 10 km, the oscillation covers the whole height range. We add this information to the manuscript (page 8 last paragraph).

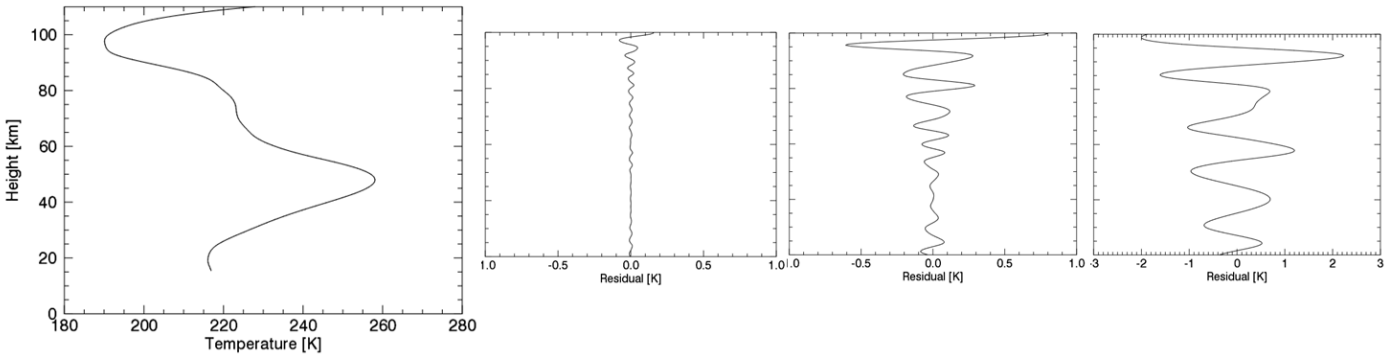


Figure 1 *First panel: CIRA-temperature profile which was detrended between 14 and 110 km height (as the SABER data). Second to fourth panel: Residuals for a sampling point distance of 2.5 k, 5 km and 10 km between 20 and 100 km height. Please be aware of the changing x-scale.*

4. We agree that the most convincing argument for this approach would be a comparison to other ones. We choose Shuai et al. (2014). Different to Ern et al. (2011), which you proposed, those authors provide the same parameter as we do—squared temperature fluctuations. Ern et al. (2011) use gravity wave amplitudes which stand in relation to the squared temperature fluctuations but this makes the comparison unnecessarily more complicated.
Shuai et al. (2014) use an earlier version of TIMED-SABER data (1.07) and a different de-trending procedure as we do. In their figure 2, they show monthly averages of the squared temperature fluctuations for the years 2002–2010. Different to us, they provide this parameter in dB ($10 \cdot log_{10}\,(T'^{2}_{GW})$) with the squared temperature fluctuation T'^{2}_{GW} which makes a short calculation necessary.

For 100 km height, we extract a yearly mean of ca. 21 dB for 50°N from their figure 2 that means

$$10 \cdot log_{10}\,(T'^{2}_{GW}) = 21 \;\Leftrightarrow log_{10}\,(T'^{2}_{GW}) = 2.1 \Leftrightarrow T'^{2}_{GW} = 10^{2.1} \approx 126$$

For 25 km height, we read am mean value of ca. 4 dB:

$$10 \cdot log_{10}\,(T'^{2}_{GW}) = 4 \;\Leftrightarrow log_{10}\,(T'^{2}_{GW}) = 0.4 \Leftrightarrow T'^{2}_{GW} = 10^{0.4} \approx 2.5$$

These values agree very well with the ones provided by us, however the colour-code used in Shuai et al. (2014) makes it difficult to give a profound answer to the question which spline approach agrees better. Nevertheless, we can argue that the overall structure which is characterized by a nearly constant or slow increase of gravity wave activity in the upper stratosphere can be observed in their figure 2 and also our figure 8a. We include this comparison also in the manuscript.

Minor comments:

1. Corrected in the entire manuscript.

2. Done.

3. Done.

4. You are right, after figure 1 we cannot exclude that the oscillation is an artefact of averaging over the whole year. However, we made exactly the same analysis with the repeating approach. We did not change the data basis or anything else, we just used the other spline approach and the oscillation became less pronounced. This contradicts the assumption that the oscillation is an

artefact of averaging. Additionally, we had a look into monthly data but the result is the same (see figure 2)

5. This is exactly what we meant. We re-formulated this paragraph.

6. Done.

7. We mean the tropo-, strato-, and mesopause corrected in the text.

Technical corrections:

1. Corrected.

2. You probably meant line 5, which we re-formulated.

3. Corrected.

4. Corrected

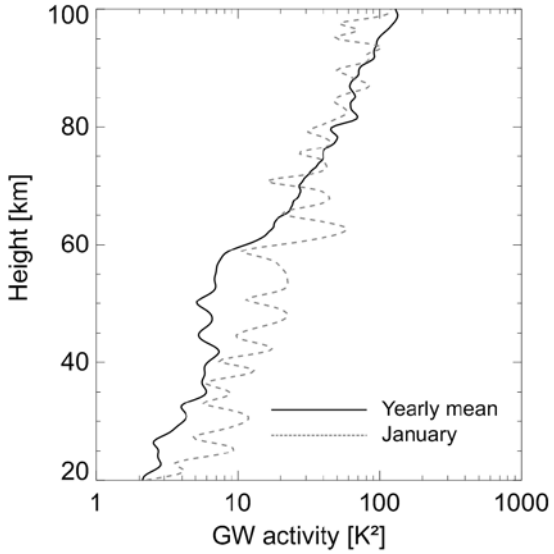


Figure 2 *Both, the yearly mean of the year 2014 and also the mean over January 2014, show oscillations (repeating spline approach).*

Smoothing data series by means of cubic splines: quality of approximation and introduction of a~~n iterative~~repeating spline approach

Sabine Wüst [1], Verena Wendt [1,2], Ricarda Linz [1], Michael Bittner [1,3]

[1]Deutsches Fernerkundungsdatenzentrum, Deutsches Zentrum für Luft- und Raumfahrt, 82234 Oberpfaffenhofen, Germany

[2]Umweltforschungsstation Schneefernerhaus, Zugspitze, Germany, now at: Institut für industrielle Informationstechnik, Hochschule Ostwestfalen-Lippe, Ostwestfalen-Lippe, Germany

[3]Institut für Physik, Universität Augsburg, 86159 Augsburg, Germany

Correspondence to: Sabine Wüst (sabine.wuest@dlr.de)

Abstract. Cubic splines with equidistant spline sampling points are a common method in atmospheric science for the approximation of ~~undisturbed~~ background conditions by means of filtering superimposed fluctuations from a data series. What is defined as background or superimposed fluctuation depends on the specific research question. The latter also determines whether the spline or ~~Often, not only the background conditions are of scientific interest but also~~ the residuals—the subtraction of the spline from the original time series—are further analysed.

Based on test data sets, we show that the quality of approximation is not increasing continuously with increasing number of spline sampling points / decreasing distance between two spline sampling points. Splines can generate considerable artificial oscillations in the background and the residuals~~data~~.

We introduce a~~n iterative~~repeating spline approach which is able to significantly reduce this phenomenon. We apply it not only to the test data but also to TIMED-SABER temperature data and choose the distance between two spline sampling points in a way that we are sensitive for a large spectrum of gravity waves.

1 Introduction

It is essential for the analysis of atmospheric wave signatures like gravity waves that these fluctuations are properly separated from the background. Therefore, particular attention must be attributed to this step during data analysis. Splines are a common method in atmospheric science for the approximation of atmospheric background conditions. The shortest wavelength or period which can be resolved by the spline is twice the sampling point distance according to the Nyquist theorem.~~The choice of a sufficiently low number of spline sampling points ensures a smoothing of the original time series. Conclusions about the shortest wavelength / period which is still uniquely resolvable by the spline can be drawn from the application of Shannon's sampling theorem (see e.g. Gubbins, 2004) based on the number of spline sampling points used.~~ Depending on the field of interest, either the smoothed data series or the residuals—the subtraction of a spline from the original time series—are further analysed (see for example the work of Kramer et al., 2016; Baumgarten et al., 2015; Zhang et al., 2012; Wüst and Bittner, 2011; Wüst and Bittner, 2008; Young et al., 1997; Eckermann et al., 1995).

Algorithms for the calculation of splines are implemented in many programming languages and in various code packages making them easy to use. Nevertheless, spline approximations need sometimes to be handled with care when it comes to physical interpretation.

Figure 1 explains our motivation for the work presented below. It shows the squared temperature residuals averaged over one year for the years 2010–2014 versus height between 44° N and 48° N and 5° E~~O~~ and 15° E~~O~~ (approximately 500 profiles per year). This region includes the Alps where gravity waves are supposed to be generated. The vertical temperature profiles are derived from the SABER (Sounding of the Atmosphere using Broadband Emission Radiometry) instrument on board the satellite TIMED (Thermosphere Ionosphere Mesosphere Energetics Dynamics), data version 2.0 (details about this data version can be found in Wüst et al. (2016) and references therein, for example). For the calculation of the residuals, we applied a cubic spline routine with equidistant sampling points. As mentioned above, the shortest wavelength which can be resolved by the spline is two times the distance between two consecutive spline sampling points. At the same time, this wavelength is the largest resolvable one in the residuals. The number of spline sampling points (and the length of the data series) therefore determines the sensitivity of the spline to specific wavelengths. The distance of 10 km between two spline sampling points makes sure that we are sensitive for a large spectrum of gravity waves. Therefore, we take the squared temperature residuals as a simple proxy for gravity wave activity with vertical wavelengths up to 20 km. Obviously, the mean squared residuals do not only reveal a strong and continuous increase with height (note the logarithmic x-axis) as it is expected since gravity wave amplitudes should increase due to the exponentially decreasing atmospheric background pressure with altitude. Superimposed on this general increase of gravity wave activity are well-pronounced oscillations with a wavelength of ca. 10 km, which is nearly equal to the distance between two spline sampling points.

Since we are not aware of any physical reason for this oscillation, we formulate the hypothesis that this is an artefact of the analysis. In order to avoid or at least reduce such problems we here propose an ~~iterative~~repeating variation of the cubic

spline approach, which we explain in section 2. In section 3, we apply the original and the ~~iterative~~repeating approach to test data sets. The results are discussed in section 4. A brief summary is given in section 5.

2 Methods and algorithms

The approach we investigate here relies on cubic splines with equidistant sampling points. Since spline theory is well-elaborated, we will therefore not go into much detail here. The algorithm we use is based on Lawson and Hanson (1974).

The first step for the adaption of a spline function to a data series on an interval [a, b] is the choice of the number of spline sampling points (also called knots). These points divide the interval for which the spline is calculated into sub-intervals of equal length. For each sub-interval a third-order polynomial needs to be defined, that means the coefficients have to be determined. At the spline sampling points, not only the function value, but also the first and second derivatives of the two adjacent polynomials need to be equal. The optimal set of coefficients is calculated according to a least square approach where the sum of the squared differences between the data series and the spline is minimized.

As mentioned above, the number of spline sampling points (and the length of the data series) determines the sensitivity of the spline to specific wavelengths. Since the length of the data series must be an integer number of the distance between two spline sample points only certain distances between two consecutive spline sampling points can be chosen if the whole data series is approximated. ~~According to the sampling theorem, the shortest wavelength / period which can be resolved by the spline is two times the distance between two consecutive spline sampling points. At the same time, this wavelength / period is the largest resolvable one in the residuals. Depending on the length of the data series only specific distances between two consecutive spline sampling points can be realised if the whole data series is approximated. The number of spline sampling points (and the length of the data series) therefore determines the sensitivity of the spline to specific wavelengths.~~

~~With SABER data we analysed temperature profiles which extend over a great height distance.~~ We would like to operate the spline algorithm in the way that we provide ~~independently of the ratio between length of the data series and number of spline sampling points~~ the shortest wavelength which shall be resolved by the spline. That means that we have to cut the upper part of the profile from case to case. This is only possible for data sets of sufficient length as the SABER temperature profiles we used for our purpose. In detail, our spline algorithm works as follows. The scheme includes the ~~iterative~~repeating as well as the non-~~iterative~~repeating algorithm.

- *Step 1: Provision of shortest wavelength*
  We provide the algorithm the shortest wavelength which shall be resolved by the spline (in the following denoted as lim). It is equal to the doubled distance between two spline sampling points; therefore the distance between two spline sampling points is equal to lim/2.

- *Step 2: Determination of x-values of the spline sampling points*

The minimal x-value of the data series is subtracted from the maximal x-value, the difference is divided by lim/2. If the result is a whole-number, 1 is added. If this is not the case, the closest integer less than the result is calculated and 1 is added. This is the number of spline sampling points used for the next step. It is denoted as n:

$$n = \left(\frac{x_{max}-x_{min}}{lim/2} - \frac{x_{max}-x_{min}}{lim/2} \bmod 1\right) + 1 \quad (1)$$

Knowing lim/2 and the minimal x-value, the x-values of the further spline sampling points can be calculated.

- *Step 3: Calculation of spline approximation*
  The spline approximation is calculated based on Lawson and Hanson (1974). If the length of the data series is not equal to an integer multiple of lim/2, the surplus part at the end of the data series is not subject of this step. For the non-~~iterative~~repeating approach, the spline algorithm stops here.

- *Step 4 (only in the case of the ~~iterative~~repeating approach): Iteration of starting point*
  The first point of the data series is removed and step 2 and 3 are repeated. If the starting point is equal to the original minimal x-value plus lim/2, the algorithm proceeds with step 5.

- *Step 5 (only in the case of the ~~iterative~~repeating approach): Calculation of the final spline*
  The mean of all splines derived before is calculated. That is the final (~~iterative~~repeating) spline.

For the ~~iterative~~repeating approach, the length of the data series is not the same in each iteration since data at the beginning and the end of the data series are not necessarily part of each iteration: at the beginning of the data series, this holds for all x-values between the minimal x-value and the minimal x-value plus lim/2 (see step 4), at the end of the data series, this is the case for all values between the maximal x-value and the maximal x-value minus lim/2 (see step 3).

For the non-~~iterative~~repeating approach, data are cut only at the end of the data series if the length of the data series is not equal to an integer multiple of lim/2.

3 Case studies

We generate a basic example using an artificial sine with a vertical wavelength of 3 km, a phase of zero and an amplitude of one. The function is sampled every 375 m (that means at its zero-crossings, at its extrema and once in between the zero-crossing and the next extremum / the extremum and the next zero-crossing).

The values for the sampling rate and the vertical wavelength are set arbitrarily. However, the spatial resolution of 375 m is motivated through the spatial resolution of TIMED-SABER, an instrument which is commonly used for the investigation of gravity waves (e.g. Zhang et al., 2012; Ern et al., 2011; Wright et al., 2011; Krebsbach and Preusse, 2007) and which delivered also the temperature profiles we used in fig. 1.

Fig. 2 (a) shows the test data series (dotted line) between 15 km and 100 km height. This great height range is chosen since it facilitates the demonstration of our results. A (non-~~iterative~~repeating) spline with a distance of 1.5 km between two spline sampling points is fitted (solid line). According to the ~~sampling~~ Nyquist theorem, the chosen distance between two spline sampling points is small enough for resolving the oscillation in our test data. In part (b) and (c) of fig. 2, a spline with a distance of 1.6 km and 1.4 km between two spline sampling points is calculated. Part (d) to (f) of fig. 2 focus on the height range of 15 km to 50 km of fig. 2 (a) to (c): here, the height-coordinates of the spline sampling points are plotted additionally (dashed-dotted lines). The asterisks mark the sampling points of the original sine. The spline adaption in fig. 2 (a) / (d) differs significantly from the spline adaption in fig. 2 (b) / (e) and 2 (c) / (f): apart from a slight oscillation at the beginning / end of the height interval, the spline is equal to zero in fig. 2 (a) / (d). The spline approximation plotted in fig. 2 (b) and (c) shows a beat-like structure in the whole height range.

In order to give an overview concerning the quality of adaption not only for some chosen examples as they were shown in fig. 2, the test data set is approximated by a cubic spline with varying numbers of spline sampling points. The squared differences between the spline and the test data are summed up between 20 km and 40 km (this height interval is chosen in order to be consistent with fig. 7 later). We call this value the sum of squared residuals which is equal to the approximation error in this case. It is not decreasing continuously with increasing number of spline sampling points / decreasing distance between two spline sampling points but it is characterized through a superimposed oscillation which reaches its maximum for a distance of ca. 1.5 km between two spline sampling points (fig. 3, solid line). When changing the phase of the test data set to $\pi/2$ (instead of zero), the sum of squared residuals~~approximation error~~ for a distance of ca. 1.5 km between two spline sampling points is much lower (fig. 3, dashed line), that means the sinusoidal oscillation is better adapted in this case. This makes clear that the sum of squared residuals~~approximation error~~ depends on the phase of the oscillation (one can also say on the exact position of the spline sampling points).

The analysis described above is repeated, but the phase of the oscillation is varied between 0 and 2π. The sum of squared residuals~~approximation error~~ (between 20 km and 40 km) is calculated for three different distances between two spline sampling points: 1.5 km (fig. 4, solid line), 1.4 km (fig. 4, short dashes) and 1.6 km (fig. 4, long dashes). The dependence on

the phase is most pronounced for a distance of 1.5 km: the oscillation is adapted very well between 20 km and 40 km for a phase of π /2 and 3π/2. For a phase of 0 and π, the contrary holds.

This example directly motivates the application of the ~~iterative~~repeating spline approach on the same test data set (see fig 5 (a)–(f) which can be directly compared to fig 2 (a)–(f): the black line represents the final spline approximation and the different colours refer to the spline approximations during the different iteration steps). In this case, the sum of squared residuals~~approximation error~~ depends much less on the distance between two spline sampling points (fig. 6 (a)) and on the phase of the test data set (fig. 6 (b)). Only for a distance of 1.6 km between two spline sampling points, a slight phase dependence is still visible (fig. 6 (b)).

Until now, we showed only test data which are not superimposed on a larger-scale variation like the atmospheric temperature background. Now, three sinusoidals with vertical wavelengths of 3, 5, and 13 km, phase 0, π/3, and π/5, and amplitude ~~0.5~~2.0 (growing amplitude with height neglected for simplicity reasons) are superimposed on a realistic vertical temperature background (fig. 7 (a)). The background is based on CIRA-86 (COSPAR International Reference Atmosphere, Committee on space Research and NASA National Space Science Data Center, 2006) temperature data for 45° N for January, which is brought on a regular grid using a cubic spline with a distance of 3 km between two spline sampling points. It was checked that no additional signatures are caused thereby. The sum of squared residuals ~~approximation error~~ shows three steps but no superimposed oscillations (fig. 7 (b)): the first step at ca. 6 km to 7 km (distance between two spline sampling points), the second one at ca. 2 km to 3 km and the last one at 1 km to 2 km. Following ~~Shannon~~Nyquist's sampling theorem, this observation can be explained through the ability of the spline to adapt the original wavelengths. The realistic background makes clear why we restrict the calculation of the sum of squared residuals~~approximation error~~ to the height range between 20 km and 40 km: this height interval is chosen in order to exclude especially the stratopause since the fast changing temperature gradient can cause additional problems for the spline approximation. Furthermore, the choice of this interval makes sure that the data used for fig. 7 (b) are part of each iteration step (which is not the case for the data at the beginning and the end of the data series, see section 2).

4 Discussion

In section 3, we showed that the ability of a spline to approximate oscillations and therefore to filter for a specific part of the wave spectrum varies

a) with the number of spline sampling points, and
b) with the exact position (height coordinate) of the spline sampling points.

While the first statement can be explained ~~Shannon~~Nyquist's sampling theorem, the second one is not well-known~~might be surprising~~.

When the distance between two spline sampling points matches exactly half the wavelength of the test data, the approximation is worst for a phase of 0 and π. In this case, the spline sampling points are located exactly between the extrema of the test data. If the height coordinates of the spline sampling points agree with the height coordinates of the extrema of the test data, the contrary holds (in appendix A, we provide a mathematical explanation for this observation). The dependence of the quality of approximation on the phase of the test data decreases with greater / smaller distances between two spline sampling points (fig. 3). These findings directly motivate the use of the presented ~~iterative~~repeating spline approach which is characterized by varying positions (height coordinate) of the spline sampling points.

Furthermore, we showed that if the distance between two spline sampling points is only slightly larger or smaller than half the wavelength present in the data series and if enough wave trains are present (which might not be the case in reality), the (non-~~iterative~~repeating) spline reminds of a beat (see fig. 2 (b) and (c), an explanation is given in appendix B). The subtraction of such a "beat" will lead to an artificial oscillation in the residuals with a periodically increasing and decreasing amplitude reaching ca. 70–80% of the original amplitude at maximum (fig. 2 (e) and (f)). This oscillation must not be interpreted as a gravity wave of varying amplitude, for example, and the described effect has to be taken into account when analysing wavelengths similar to the doubled distance between two spline sampling points.

For our case studies, we used a constant and a realistic CIRA-based temperature background profile. For both background profiles, ~~W~~we showed the sum of squared residuals~~approximation error (the squared differences between the spline and the test data summed up between 20 km and 40 km)~~ decreases much smoother with increasing number of spline sampling points for the ~~iterative~~repeating approach compared to the non-~~iterative~~repeating one (compare fig. 3 to fig. 6 (a)) and the amplitude of the "beat"-like structure is reduced.

However, the motivation for this work was—as already mentioned—the results shown in fig. 1 which are characterized by a strong superimposed oscillation with a wavelength of approximately 10 km for which we do not have a physical explanation. Figure 8 (a) now depicts the mean squared residuals after the application of the ~~iterative~~repeating spline to the same data set, fig. 8 (b) focuses on the year 2014 (the dashed line is based on the application of the ~~iterative~~repeating spline, the solid line

refers to the non-~~iterative~~repeating spline). This year is chosen arbitrarily and allows the direct comparison between the ~~iterative~~repeating and non-~~iterative~~repeating approach. The amplitude of the superimposed oscillation is reduced significantly but the oscillation can still be observed. This supports our hypothesis that the strong superimposed oscillation described in fig. 1 is an artefact of the non-~~iterative~~repeating spline de-trending procedure. Furthermore, it becomes obvious now that gravity wave activity increases less with altitude~~is less variable~~ between approximately 45 km and 60 km height compared to the height range below and above. This is in accordance with literature (e.g. Mzé et al., 2014; Offermann et al., 2009). For most heights, the mean squared residuals are smaller for the ~~iterative~~repeating approach than for the non-~~iterative~~repeating one. At 38 km height for example, the difference reaches ca. 2.5 K², which is approximately 32 % (referring to the mean value of both approaches). Shuai et al. (2014) use an earlier version of TIMED-SABER temperature data (1.07) and a different de-trending procedure as we do in order to derive monthly averages of the squared temperature fluctuations for the years 2002–2010. They provide this parameter in dB ($10 \cdot log_{10}\,(T'^{2}_{GW})$) with the squared temperature fluctuation T'^{2}_{GW} which makes a short calculation necessary.

For 100 km height, we extract a yearly mean of ca. 21 dB for 50°N from their figure 2 that means

$$10 \cdot log_{10}\,(T'^{2}_{GW}) = 21 \Leftrightarrow log_{10}\,(T'^{2}_{GW}) = 2.1 \Leftrightarrow T'^{2}_{GW} = 10^{2.1} \approx 126$$

For 25 km height, we read am mean value of ca. 4 dB:

$$10 \cdot log_{10}\,(T'^{2}_{GW}) = 4 \Leftrightarrow log_{10}\,(T'^{2}_{GW}) = 0.4 \Leftrightarrow T'^{2}_{GW} = 10^{0.4} \approx 2.5$$

These values agree very well with the ones provided here. Furthermore, the overall structure which is characterized by a slow increasing or even a nearly constant gravity wave activity in the upper stratosphere can be observed in their fig. 2 and our fig. 8 (a).

In order to give a comprehensive comparison of the ~~iterative~~repeating and the non-~~iterative~~repeating spline algorithm, we calculate also the mean (non-squared) residuals. In this case, the results look very similarly; in both cases, they again show an oscillation with a vertical wavelength of 10−20 km (fig. 8 (c) for the non-~~iterative~~repeating approach, fig. 8 (d) for the ~~iterative~~repeating spline approach). We can explain this in the following way: when calculating the mean (non-squared) residuals and the mean squared residuals at a specific height, one refers to two different parameters of the distribution of residuals at that specific height. While the mean (non-squared) residuals estimate the mean of the distribution, the mean squared residuals refer to the variance of the distribution. We conclude: at a defined height, the ~~iterative~~repeating approach changes the mean of the distribution of the residuals only slightly, but it reduces its spread significantly. For individual profiles, the approximation through the ~~iterative~~repeating approach is therefore less variable in average and can be recommended. The ~~iterative~~repeating approach can also be recommended if squared residuals are needed for further analysis (e.g. for the calculation of the wave potential energy). If non-squared residuals will be analysed, it does not make a

difference in average which approach is applied. In this case, only waves with amplitudes larger than 0.5 K in the stratosphere and 1.0 K in the mesosphere (fig. 8 (c) and (d)) should be taken seriously.

It is known that the tropo-, strato-, and mesopause ~~height regions~~ where the temperature gradient changes are challenging for approximation methods. The same holds for the beginning and the end of a data series. Since the beginning and the end of the data series do not very much (we analysed the SABER data between ca. 12–14 km and 110 km height), possible side effects might not cancel out.~~, and w~~We speculate that these~~is~~ ~~is~~ are at least ~~one~~two of the reasons for the superimposed oscillation in fig. 8 (c) and (d). However, analysing this hypothesis is beyond the scope of this manuscript.

There exist many methods to approximate / de-trend / filter time series (see e.g Baumgarten et al. (2015), and references therein) and we do not claim that the presented ~~iterative~~repeating cubic spline is the best method for every purpose and every data series. It is just one possible algorithm which reduces disadvantages of the non-~~iterative~~repeating cubic spline routine as it was proposed by Lawson and Hanson (1974) like the dependence of approximation on the exact position (height coordinate) of the spline sampling points. However, it comes along with enhanced computational effort which is of special importance when analysing large data sets.

5 Summary

It is essential for the analysis of atmospheric wave signatures like gravity waves that these fluctuations are properly separated from the background. Therefore, particular attention must be attributed to this step.

Cubic splines with equidistant sampling points are a common method in atmospheric science for the approximation of superimposed, large-scale structures in data series. The subtraction of the spline from the original time series allows the investigation of the residuals by means of different spectral analysis techniques. However, splines can generate artificial oscillations in the residuals. The ability of a spline to approximate oscillations varies not only with the number of spline sampling points, but also with their exact position. When the distance between two spline sampling points equals exactly or approximately half the wavelength of the waves present in the data, the last-mentioned effect is most pronounced.

Since knowledge about the wavelengths present in the data set is normally not available in advance, this directly motivates the use of an ~~iterative~~repeating spline which is based on changing starting points. It comes along with enhanced computational effort but it can be recommended for the approximation / de-trending of individual profiles and if squared residuals are needed for further analysis (e.g. for the calculation of the wave potential energy).

Acknowledgement

We would like to thank the TIMED-SABER team for their great work in providing an excellent data set.

We also thank the Bavarian Ministry for Environment and Consumer Protection for financially supporting our work: V. Wendt was paid by the Bavarian project BHEA (Project number TLK01U-49580, 2010–2013). The work of S. Wüst was subsidised in parts by this project.

At least, we thank Julian Schmoeckel, University of Augsburg, for helping to produce the test data sets and the figures.

Appendix A

Between two spline sampling points, a spline is equal to a cubic polynomial of the form

$f(z) = az^3 + bz^2 + cz + d$ with $a, b, c, d \in \mathbb{R}$. (2)

Its derivatives are

$f^{(1)}(z) = 3az^2 + 2bz + c,$ (3)

$f^{(2)}(z) = 6az + 2b,$ (4)

$f^{(3)}(z) = 6a.$ (5)

Between two spline sampling points, the second derivative of a spline depends linearly on the height coordinate z. That means the curvature of the spline can change from negative to positive or vice versa between two spline sampling points but it can only increase or decrease linearly or it can stay constant. At the spline sampling points, all derivatives of the two adjacent polynomials must agree. For example, a spline cannot form two parabolas with different signs in two adjacent intervals in order to approximate a sine / cosine since the second derivative (curvature) would be positive constant in one interval and negative constant in the other. If the spline sampling points are not distributed in a way such that the curvature of the original function increases or decreases linearly between two spline sampling points, the spline cannot approximate the original function properly.

Therefore, the ability of the spline to reproduce a sine / cosine does not only depend on the number of spline sampling points, it varies also with their position.

Appendix B

The optimal spline parameters are determined through a least-square approach: depending on the spline parameters, the squared differences between the spline and the original data set are minimized. The maximum wavelength which a spline can approximate in principle is equal to two times the distance between two spline sampling points.

Let us denote the oscillation which has to be approximated with $f_1(z)$ and the spline with $f_2(z)$.

If those two oscillations which will be subtracted from each other are characterized by very similar wave numbers $k_1$and k_2, then a beat with the following wave numbers will occur.

$$f_1(z) - f_2(z) = \sin k_1 z - \sin k_2 z \overset{(*)}{=} 2\cos\left(\frac{k_1+k_2}{2}z\right)\sin\left(\frac{k_1-k_2}{2}z\right) \quad (6)$$

where

$\frac{k_1+k_2}{2}$ is the wave number of the beat, which is very similar to the original wave number, and

$\frac{k_1-k_2}{2}$ is the wave number of the envelope.

$(*)$: application of an addition theorem

References

Baumgarten, G., Fiedler, J., Hildebrand, J., and Lübken, F.-J.: Inertia gravity wave in the stratosphere and mesosphere observed by Doppler wind and temperature lidar, Geophys. Res. Lett., 42, 10,929–10,936, doi: 10.1002/2015GL066991, 2015.

Committee on Space Research; NASA National Space Science Data Center: COSPAR International Reference Atmosphere (CIRA-86): Global Climatology of Atmospheric Parameters. NCAS British Atmospheric Data Centre. http://catalogue.ceda.ac.uk/uuid/4996e5b2f53ce0b1f2072adadaeda262, 2006.

Eckermann, S. D., Hirota, I., and Hocking, W. K.: Gravity wave and equatorial wave morphology of the stratosphere derived from long-term rocket soundings. Q.J.R. Meteorol. Soc., 121, 149–186, doi: 10.1002/qj.49712152108, 1995.

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

~~Gubbins, D.: Time series analysis and inverse theory for geophysicists, Cambridge University Press, 2004.~~

Kramer, R., Wüst, S., and Bittner, M.: Climatology of convectively generated gravity waves at Prague based on operational radiosonde data from 13 years (1997–2009). J. Atmos. Sol.-Terr. Phys., 140, 23–33, doi: 10.1016/j.jastp.2016.01.014, 2016.

Krebsbach, M., and Preusse, P.: Spectral analysis of gravity wave activity in SABER temperature data, Geophys. Res. Lett., 34, L03814, doi: 10.1029/2006GL028040, 2007.

Lawson, C. L., and Hanson, R. J.: Solving least squares problems, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, U.S.A. 1974.

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

Offermann, D., Gusev, O., Donner, M., Forbes, J. M., Hagan, M., Mlynczak, M. G., Oberheide, J., Preusse, P., Schmidt, H. and Russell III, J. M.: Relative intensities of middle atmosphere waves, J. Geophys. Res., 114, D06110, doi: 10.1029/2008JD010662, 2009.

Shuai, J., Zhang, S., Huang, C., YI, F., Huang, K., Gan, Q., and Gong, Y.: Climatology of global gravity wave activity and dissipation revealed by SABER/TIMED temperature observations. Science China Technological Sciences, 57, 998–1009, doi: 10.1007/s11431-014-5527-z, 2014.

Wright, C. J., Rivas, M. B., and Gille, J. C.: Intercomparisons of HIRDLS, COSMIC and SABER for the detection of stratospheric gravity waves, Atmos. Meas. Tech., 4, 1581-1591, doi: 10.5194/amt-4-1581-2011, 2011.

Wüst, S., and Bittner, M.: Gravity wave reflection: case study based on rocket data. J. Atmos. Sol.-Terr. Phys., 70, 742–755, doi:10.1016/j.jastp.2007.10.010, 2008.

Wüst, S., and Bittner, M.: Resonant interaction between two planetary waves with zonal wave number two? J. Atmos. Sol.-Terr. Phys., 73, 771–778, doi:10.1016/j.jastp.2011.01.004, 2011.

Wüst, S., Wendt, V., Schmidt, C., Lichtenstern, S., Bittner, M., Yee, J.-H., Mlynczak, M. G., and Russell III, J. M.: Derivation of gravity wave potential energy density from NDMC measurements. J. Atmos. Sol.-Terr. Phys., 138–139, 32-46, doi:10.1016/j.jastp.2015.12.003, 2016.

Young, L. A., Yelle, R. V., Young, R., Seiff, A., and Kirk, D. B.: Gravity Waves in Jupiter's Thermosphere, Science, 276, 108–111, doi:10.1126/science.276.5309.108, 1997.

Zhang, Y., Xiong, J., Liu, L., and Wan, W.: A global morphology of gravity wave activity in the stratosphere revealed by the 8-year SABER/TIMED data, J. Geophys. Res., 117, D21101, doi:10.1029/2012JD017676, 2012.

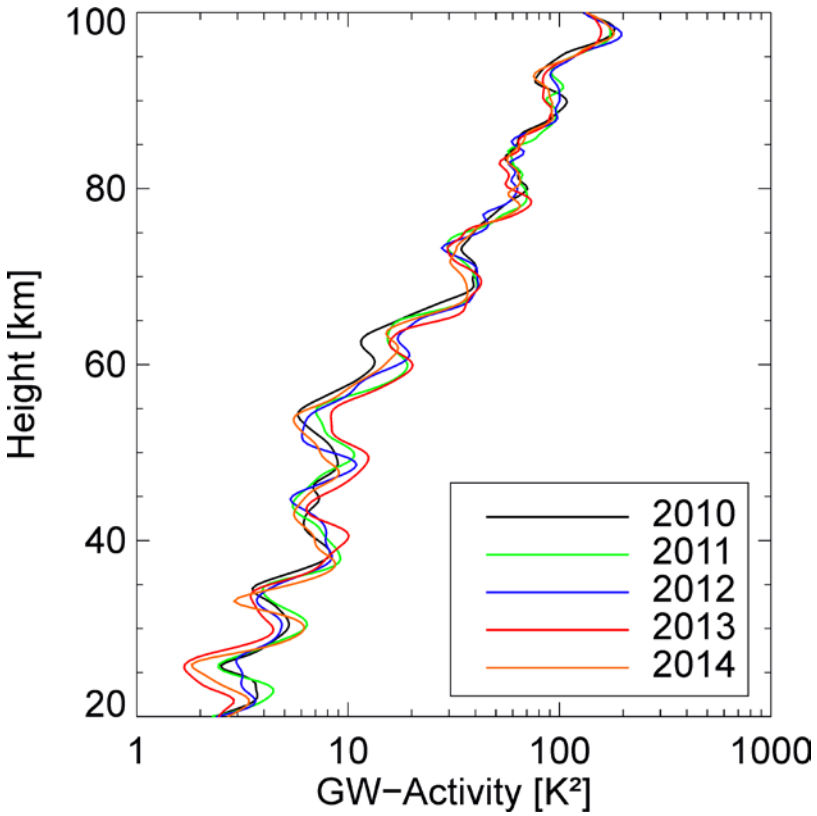


Figure 1: Mean squared temperature residuals for the years 2010 to 2014 (colour-coded): they are derived from TIMED-SABER, data version 2.0 by using a cubic spline routine with equidistant sampling points for de-trending. The distance between two spline sampling points is 10 km. All vertical SABER temperature profiles which were retrieved between 44° N and 48° N and 5° O and 15° O are used (that means approximately 30-50 profiles per month and approximately 500 profiles per year).

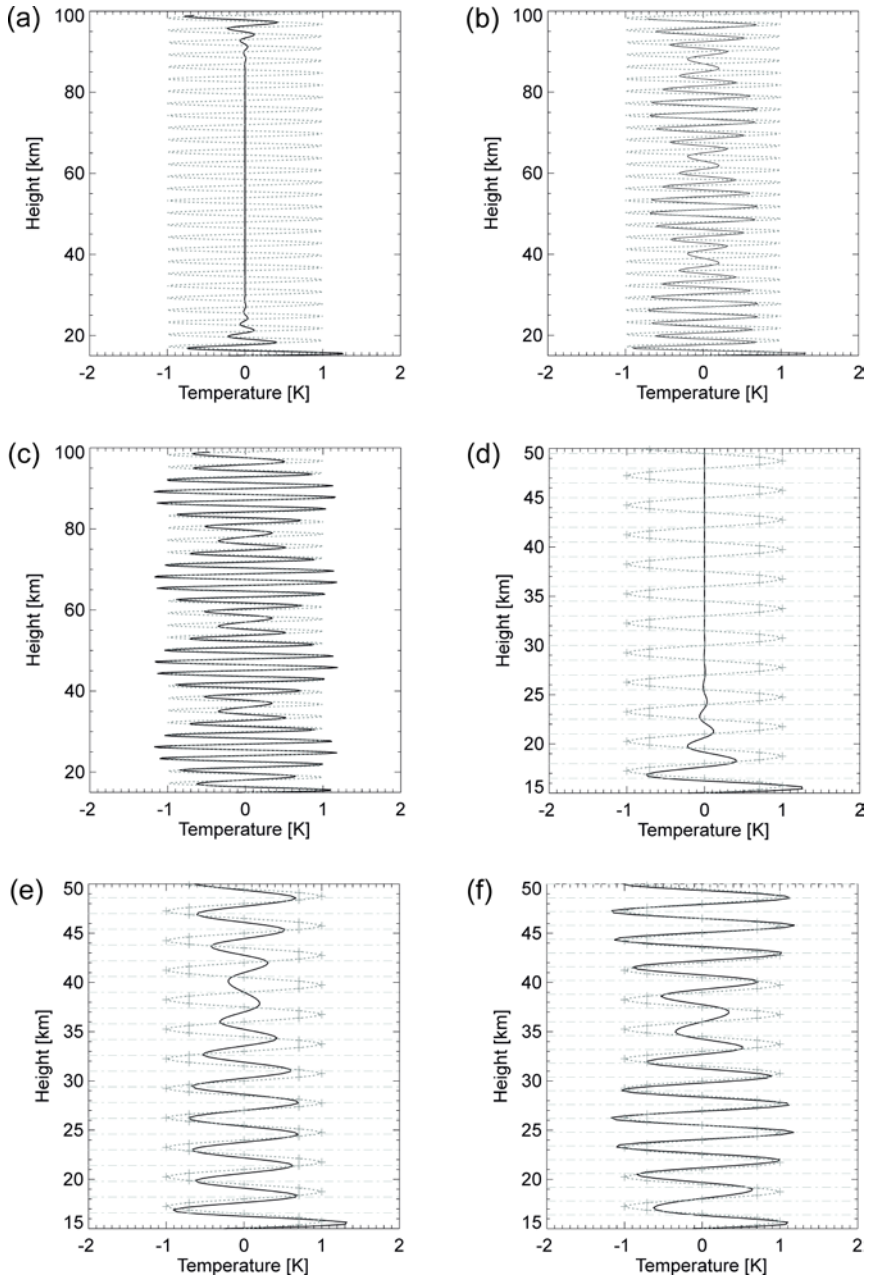

(a)
(b)
(c)
(d)
(e)
(f)
Height [km]
Temperature [K]

Figure 2: This figure shows the approximation of a cubic spline using different numbers of spline sampling points:

(a) A spline with a distance of 1.5 km between two spline sampling points is fitted (solid line) to the test data (dotted line).

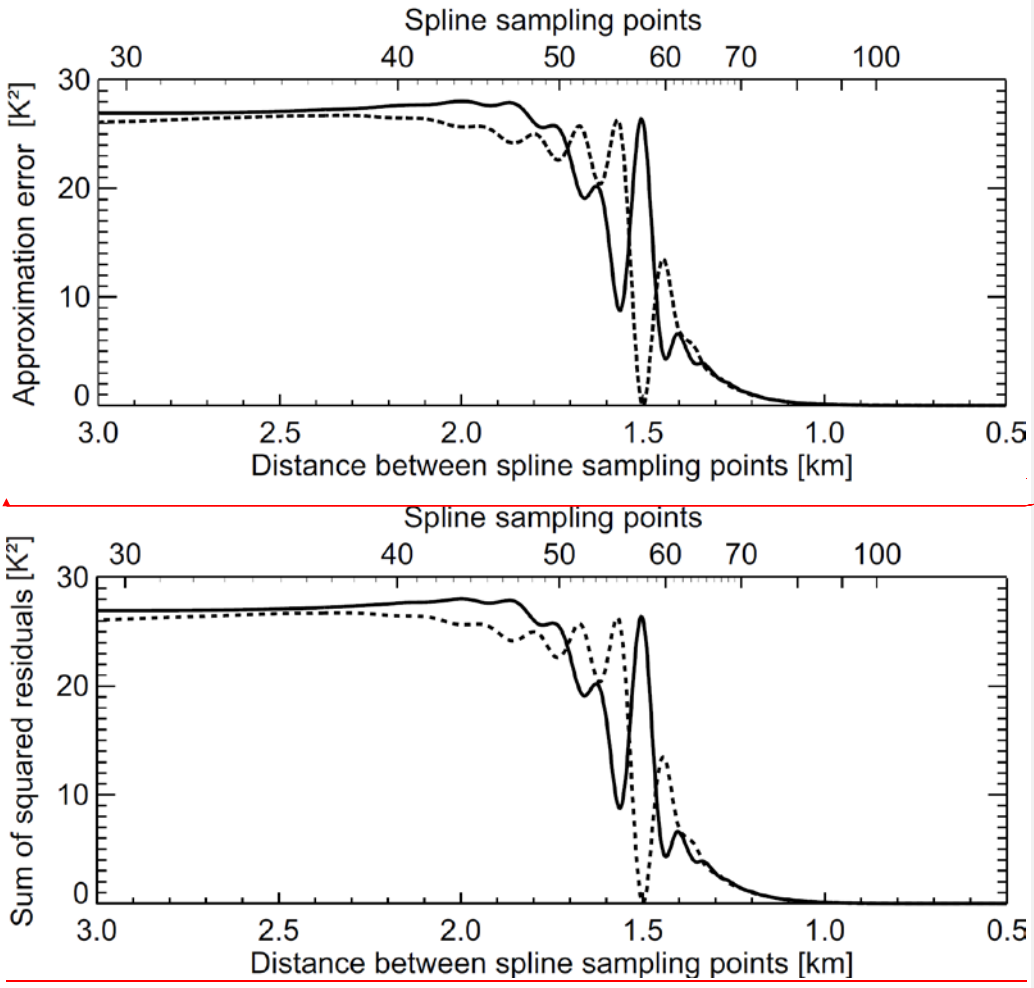
(b) Same as (a) but the distance between two spline sampling points is 1.6 km.

(c) Same as (a) but the distance between two spline sampling points is 1.4 km.

(d) Same as (a) but restricted to the height range between 15 km and 50 km. The dashed-dotted lines refer to the height-coordinate of the spline sampling points. The asterisks show the sampling points of to the original sine.

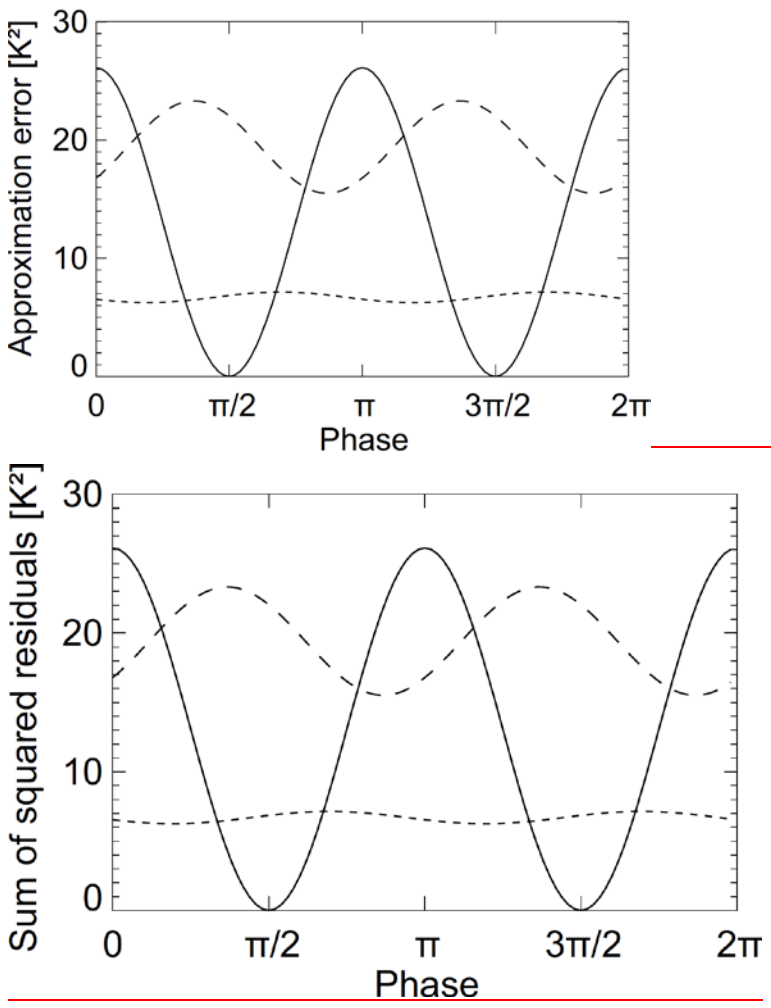
(e) Same as (b) but restricted to the height range between 15 km and 50 km.

(f) Same as (c) but restricted to the height range between 15 km and 50 km.



Formatiert: Englisch (Großbritannien)

Figure 3: This figure shows the differences between the spline and the approximated test data (solid line: phase of 0, dashed line: phase of π/2) which are summed up between 20 km and 40 km. They are plotted versus the distance between the number of spline sampling points / distance between spline sampling points. The number of spline sampling points / distance between spline sampling points refers to the whole height range between 15 km and 100 km.



Formatiert: Englisch (Großbritannien)

Figure 4: Dependence of the sum of squared residuals ~~approximation error~~ on the phase of the wave with a wavelength of 3.0 km and a distance of 1.4 km (short dashes), 1.5 km (solid line) and 1.6 km (long dashes) between two spline sampling points.

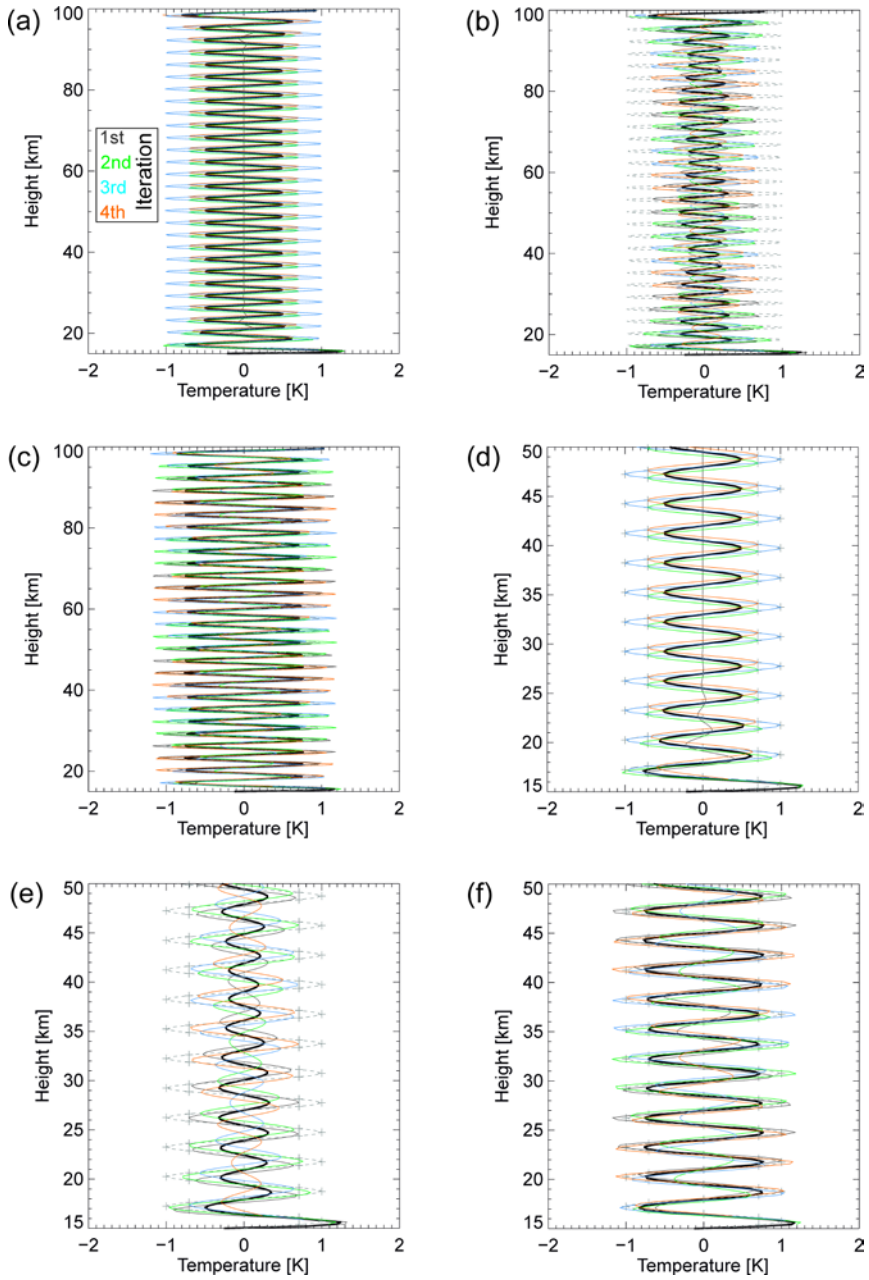
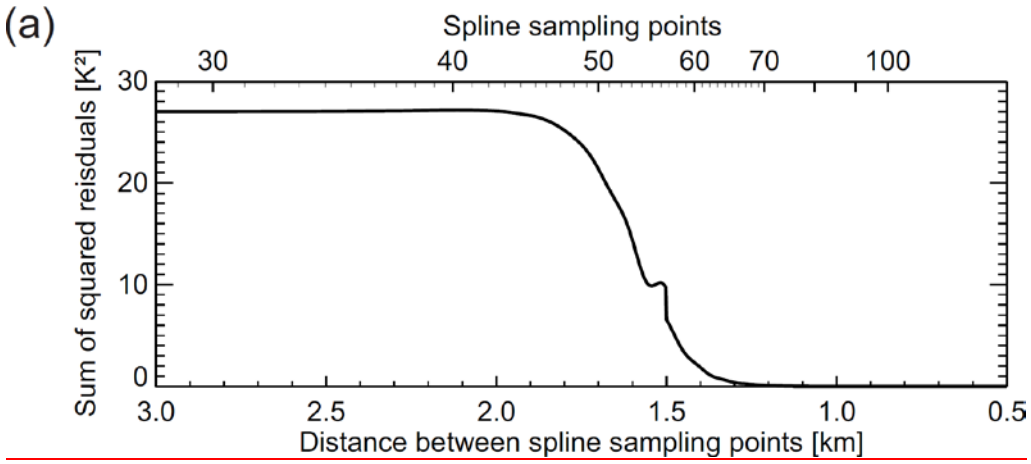
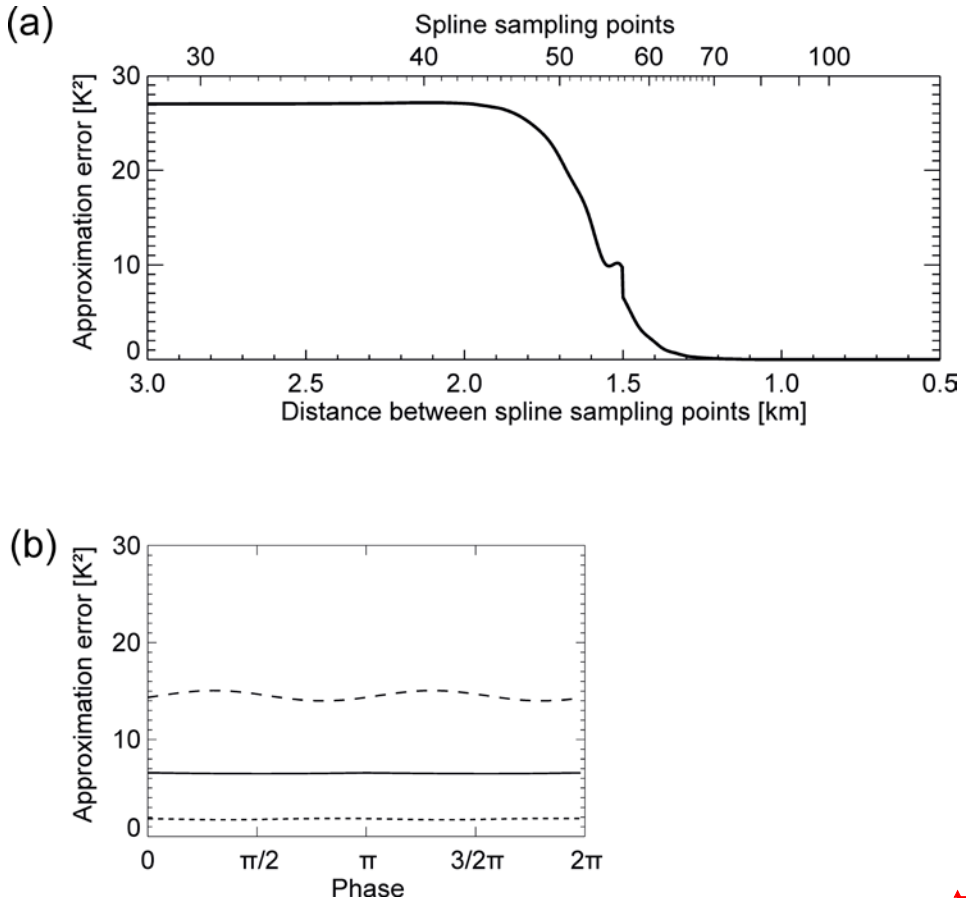
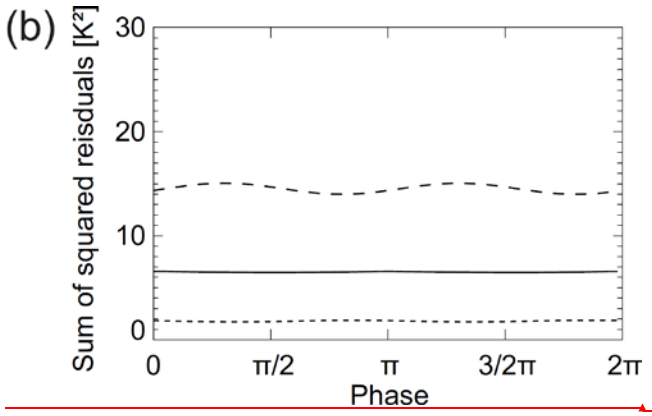

(a)
(b)
(c)
(d)
(e)
(f)
Height [km]
Temperature [K]
1st
2nd
3rd
4th
Iteration

Figure 5: Here, the results based on the ~~iterative~~repeating spline approach are shown. The different colours refer to the different spline approximations (to keep it as clear as possible, we only show the first four iterations, a fifth one exists for case (b) and (e), see step 4 of the algorithm). The black line represents the final spline approximation. The distance between two spline sampling points in part (a) to (f) agrees with the respective values in figure 2 part (a) to (f). While part (a) to (c) show the height range between 15 km and 100 km, part (d) to (f) focus on the height range between 15 km to 50 km. The asterisks ~~and dashed dotted lines~~ have the same meaning as in figure 2


(a)
Spline sampling points
30
40
50
60
70
100
Sum of squared reisduals [K²]
30
20
10
0
3.0
2.5
2.0
1.5
1.0
0.5
Distance between spline sampling points [km]



Formatiert: Englisch (Großbritannien)

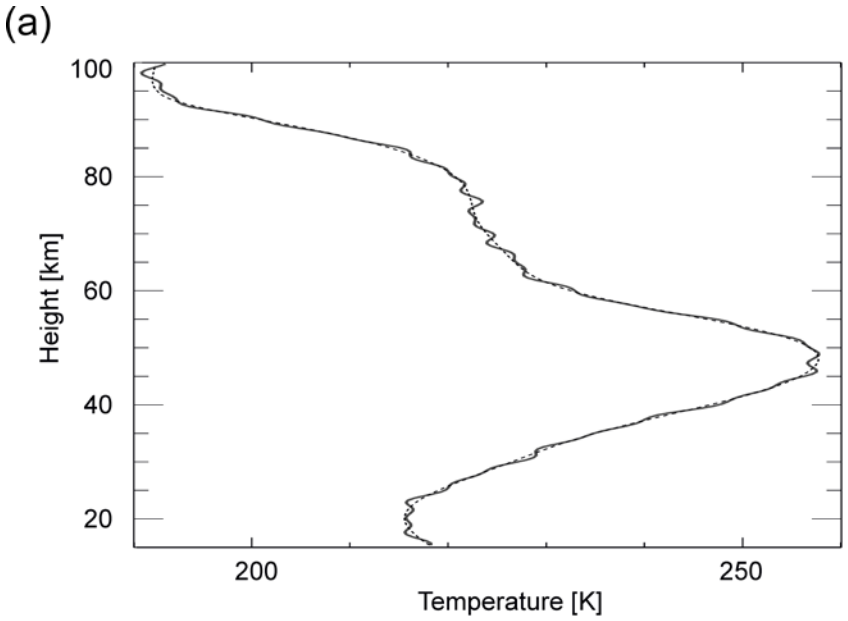


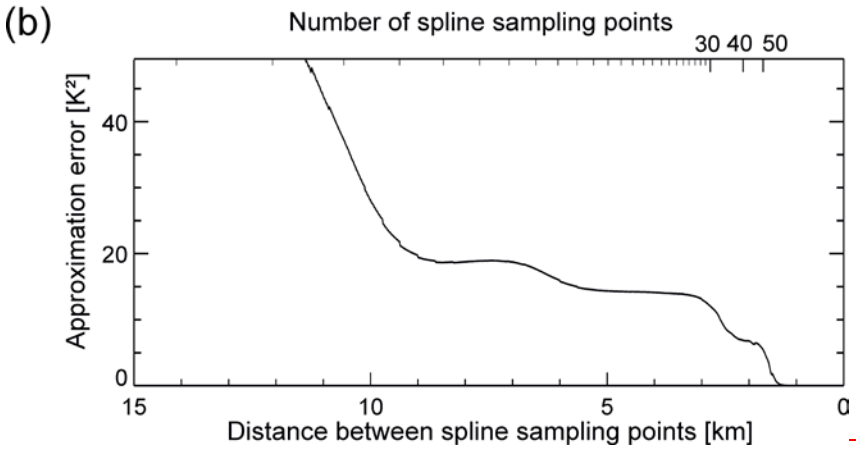
Formatiert: Englisch (Großbritannien)

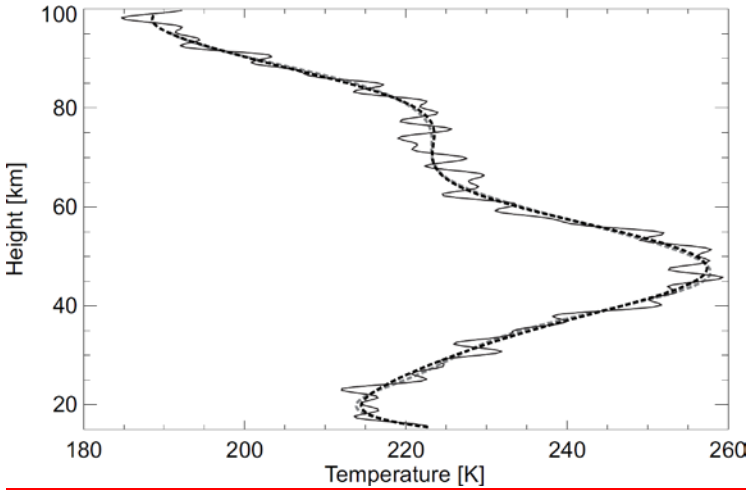
Figure 6: Part (a) is equivalent to figure 3, part (b) is equivalent to figure 4, but here the ~~iterative~~repeating spline approach is used.

Formatiert: Englisch (Großbritannien)

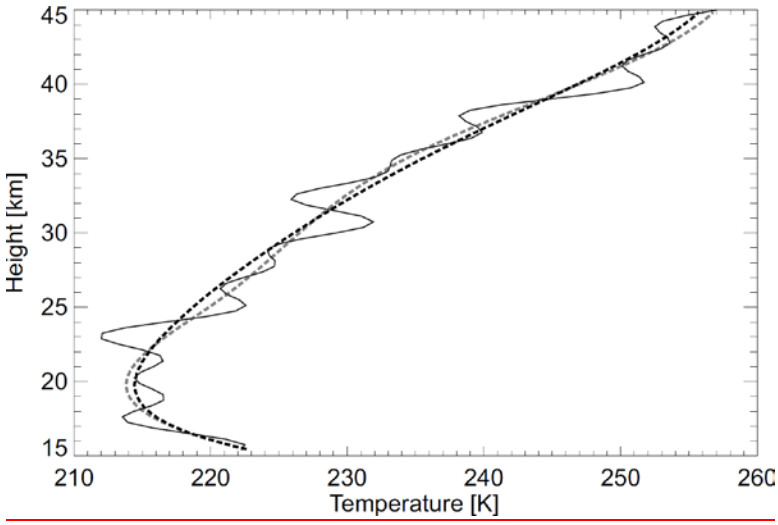
(a)


(a)
100
80
60
40
20
Height [km]
200
250
Temperature [K]


(b)
Number of spline sampling points
30 40 50
Approximation error [K²]
40
20
0
15
10
5
0
Distance between spline sampling points [km]


100
80
60
40
20
Height [km]
180
200
220
240
260
Temperature [K]

(b)


45
40
35
30
25
20
15
Height [km]
210
220
230
240
250
260
Temperature [K]

(c)

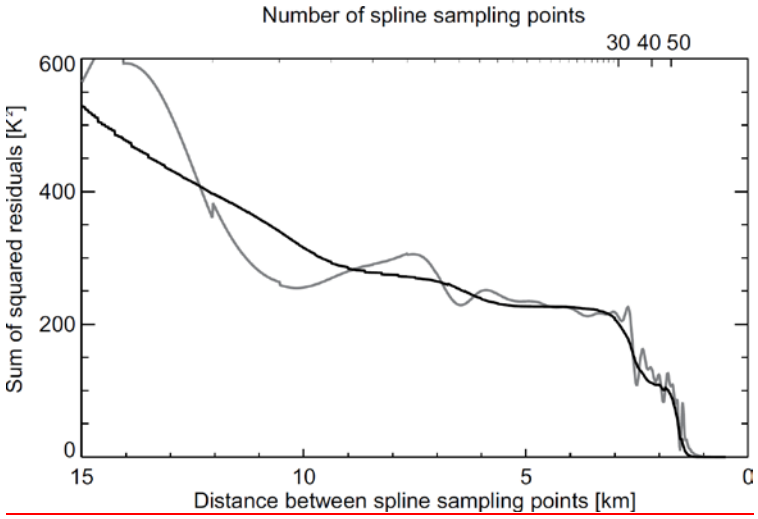


Figure 7: a) The solid line depicts three sinusoidals with vertical wavelengths of 3, 5 and 13 km, phase 0, π/3, and π /5, and amplitude ~~0.5~~2.0 ~~which are superimposed on a realistic temperature profile~~and the spline approximations for the repeating (black) and non-repeating (grey) spline approach (dashed line~~, sampling point distance 10 km~~). Part b) as part a) but focussing on the height interval between 20 and 40 km for which the sum of squared residuals is calculated in part c). The range of the x- and y-axis differs from the ones used in fig. 3 and 6 (a).~~Part b) shows the approximation error depending on the distance between two spline sampling points.~~

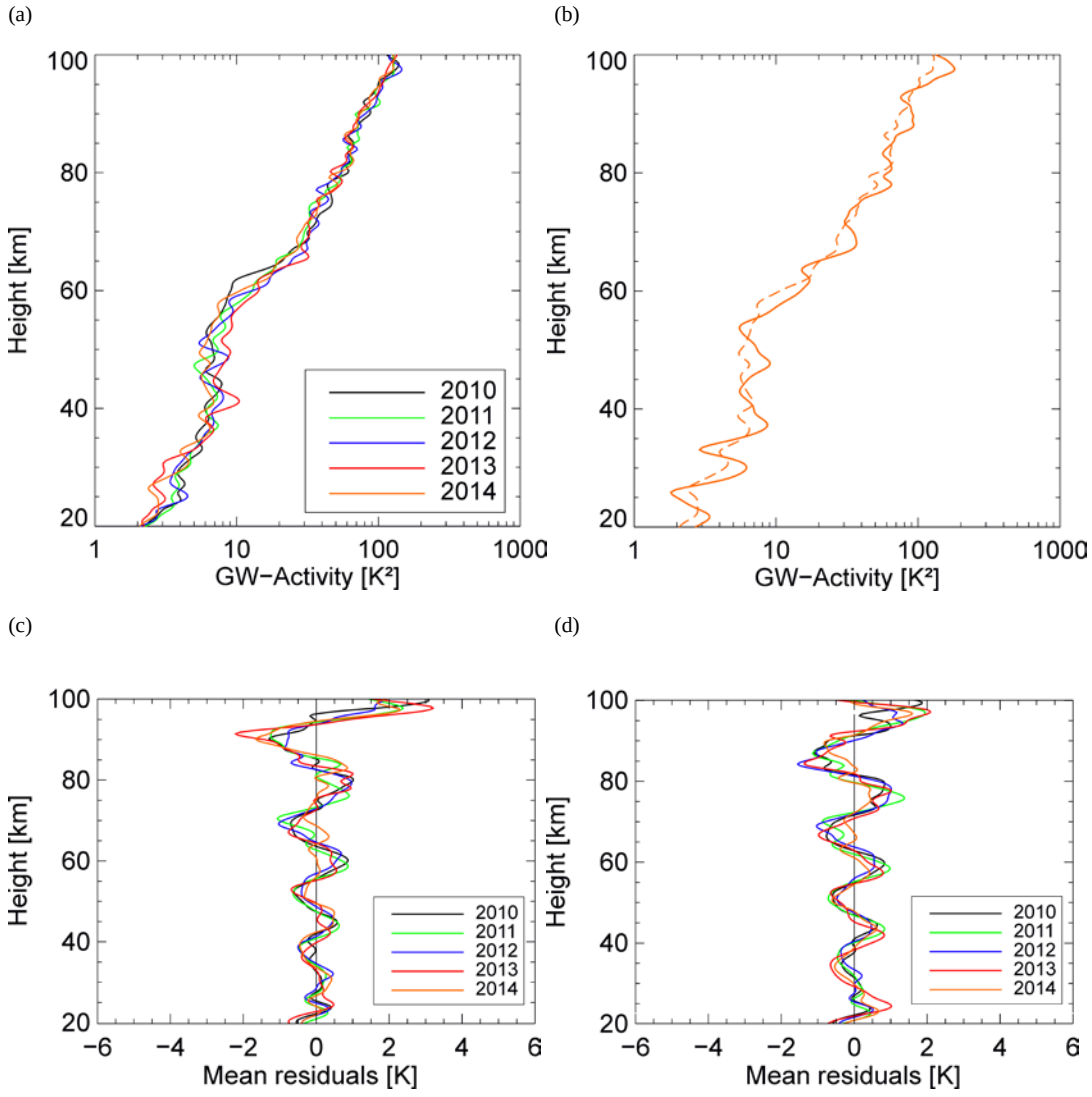


Figure 8: a) As figure 1, for de-trending the ~~iterative~~repeating cubic spline routine with equidistant sampling points as it is described in section 2 is used. b) Mean squared residuals for the ~~iterative~~repeating (dashed line) and the non-~~iterative~~repeating (solid line) approach for the year 2014. c) Mean (non-squared) residuals for the *non-~~iterative~~repeating* approach for the years 2010 to 2014, and d) mean (non-squared) residuals for the ~~*iterative*~~*repeating* approach.