Efficient radiative transfer simulations for a broadband infrared radiometer—Combining a weighted mean of representative frequencies approach with frequency selection by simulated annealing

S.A. Buehler [a,*], V.O. John [b], A. Kottayil [a], M. Milz [a], P. Eriksson [c]

[a] *Department of Space Science, Luleå University of Technology, Space Campus 1, 98128 Kiruna, Sweden*
[b] *Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK*
[c] *Radio and Space Science, Chalmers University of Technology, 41296 Göteborg, Sweden*

ABSTRACT

We present a method to efficiently simulate the measurements of a broadband infrared instrument. The High Resolution Infrared Radiation Sounder (HIRS) instrument is used as example to illustrate the method. The method uses two basic ideas. Firstly, the channel radiance can be approximated by a weighted mean of the radiance at some representative frequencies, where the weights can be determined by linear regression. Secondly, a near-optimal set of representative frequencies can be found by simulated annealing.

The paper does not only describe and analyze the method, it also describes how the method was used to derive optimized frequency grids for the HIRS instruments on the satellites TIROS N, NOAA 6–19, and Metop A. The grids and weights as well as the optimization algorithm itself are openly available under a GNU public license.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Radiative transfer (RT) simulations for broadband infrared (IR) radiometers are computationally expensive, because of the fine spectral structure inside the instrument bands. A typical example of the class of instruments we are concerned with here is the High Resolution Infrared Radiation Sounder (HIRS), which flies on the NOAA and MetOp series of satellites.

Even for the clear-sky case, simple RT simulations for this instrument are costly, because they require thousands of monochromatic RT calculations for each instrument channel, which are then integrated according to the instrument channel response function. In the presence of clouds this problem is even worse, since the individual monochromatic RT calculations then are computationally much more expensive.

Simulated annealing (SA) is a well-established method for solving global optimization problems, i.e., for finding the global maximum or minimum of a function [1]. It is particularly well suited for functions that depend on many variables and that are strongly nonlinear,

---

a case where many other optimization methods are not applicable.

In this article, we show how SA can be combined with a linear regression method to find representative frequencies and associated weights for HIRS RT simulations. The presented method is an alternative to the established correlated k method for efficient RT simulations in the IR spectral range. To our knowledge, there is only one other article Moncet et al. [2] that describes a similar approach to modeling broadband IR radiances.

We focus on the HIRS instrument to have a well defined test case to illustrate and test the method. However, the method itself is completely general and could be easily applied to other similar instruments. Furthermore, we focus on the clear-sky case, but the method is in principle equally applicable to cloudy radiances, provided that the RT model is sufficiently fast to calculate the required high spectral resolution reference simulations in a reasonable time.

The article is structured as follows: Section 2 describes the RT model that was used for this study, explains the weighted mean of representative frequencies approximation, and describes how simulated annealing can be used to find a near-optimal frequency grid. Section 3 investigates the methods performance and discusses its accuracy and limitations. Finally, Section 4 contains summary and conclusions.

2. Methodology

2.1. The ARTS model

The Atmospheric Radiative Transfer Simulator (ARTS) is a flexible and powerful radiative transfer model for the thermal radiation spectral range (microwave to infrared). It is written in the C++ programming language and freely available under a GNU public license from the web site http://www.sat.ltu.se/arts/. A general description of the clear-sky part of the model can be found in Buehler et al. [3]. For calculations in the presence of clouds, ARTS has two alternative scattering algorithms, a discrete ordinate iterative solution method [4] and a Monte Carlo method [5]. The model has been validated against other RT models in various intercomparison studies, for example Melsheimer et al. [6], Buehler et al. [7], and Saunders et al. [8].

ARTS has already been used in a large and diverse number of applications. It was for example applied to analyze sub-millimeter wave limb sounder data, both for operational instruments [9,10] and for planned future missions [11–13]. It was also used extensively to analyze operational meteorological microwave satellite measurements [14–17], to simulate future sub-millimeter cloud ice measurements [18–20], and to simulate radio occultation measurements [21]. Furthermore, it was used to investigate the impact of clouds on ground based measurements of the cosmic microwave background [22] and to simulate outgoing longwave radiation fluxes [23,24]. Other diverse applications include the analysis of ground based [25,26] and airborne [27] microwave data, the analysis of ground based Fourier transform infrared

spectrometer data [28], the airborne retrieval of snow microwave emissivity [29], and the simulation of the full Stokes vector in inhomogeneous precipitation [30].

The strength of ARTS is that it can be applied to a wide range of flux calculations and sensor simulations, in different viewing geometries. A drawback of this flexibility is that it may require some work to set up the model for a particular application. To ameliorate this, we include setup files for some standard instruments. The basis for this article is the ARTS standard setup for HIRS simulations.

2.2. The HIRS instrument

The High Resolution Infrared Radiation Sounder (HIRS) is described for example in Smith et al. [31]. At present, data are available from 15 different HIRS instruments since 1978. HIRS measures radiation coming from Earth and its atmosphere primarily in the IR region of the spectrum. This includes both thermal IR (Channels 1–12) and near IR (Channels 13–19). Furthermore, there is one channel in the visible spectral range (Channel 20). We focus on the thermal IR channels here. They are widely spaced over a large part of the Planck curve for typical atmospheric temperatures. Fig. 1 shows the HIRS channel positions, relative to the atmospheric IR spectrum.

Figs. 2 and 3 show the spectral response of the different thermal radiation HIRS channels. They also show zenith opacity spectra of various trace gases, so that one can easily see which trace gas is important for which channel. Spectral response functions vary between different HIRS instruments (on different satellites). All figures in this section are for HIRS on NOAA 14.

2.3. Broadband infrared radiative transfer simulations

Broadband infrared radiometers, such as HIRS, have a typical bandwidth of $50\,cm^{-1}$, much wider than the width of individual spectral lines (compare Fig. 2). The instrument measures the integral of the radiance spectrum, weighted by some broad channel response function.

The straightforward approach to simulating such measurements is to calculate the monochromatic radiance spectrum on a high resolution frequency grid, fine enough to resolve all spectral features, and then do the integration as a discrete sum over all frequencies with weights according to the instrument channel response. Explicitly,

$$I_{int} = \int R(v)I(v)\,dv \approx \sum_i R(v_i)I(v_i)\Delta v_i, \tag{1}$$

where I_{int} is the integrated radiance for an instrument channel, v is frequency, $R(v)$ is the instrument channel response function, and $I(v)$ is radiance. The radiances $I(v_i)$ at the individual grid frequencies can be easily calculated by solving the monochromatic radiative transfer (RT) equation. It should be noted here that, to be consistent with the HIRS channel response functions, I should be in radiance units $(W/(m^2\,sr\,Hz))$, and not in brightness temperature units (K).

This approach is robust and exact, provided that the frequency grid is fine enough to resolve all spectral features.
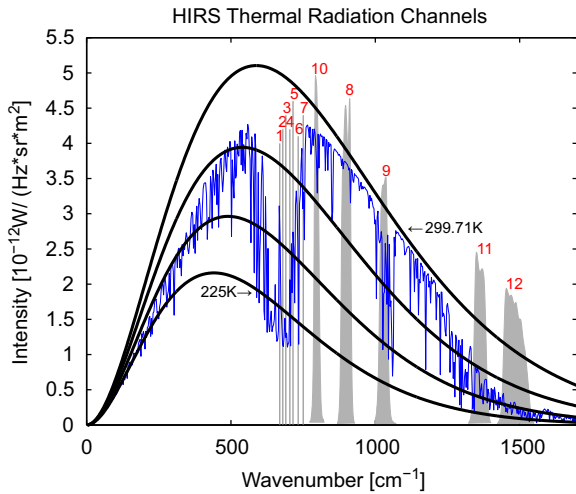
Fig. 1. Positions of NOAA 14 HIRS thermal radiation channels relative to the atmospheric radiance spectrum. Gray bands indicate HIRS channel positions. For Channels 8–12 their shape indicates the shape of the channel response function (with arbitrary scaling). For Channels 1–7 only the center positions are shown, because these channels overlap significantly (see Fig. 2). Smooth black curves show Planck functions for different temperatures. The temperatures are 225, 250, 275, and 299.71 K, where the latter corresponds to the surface temperature. The blue curve shows the atmospheric spectrum for a tropical atmosphere, calculated on a frequency grid with 2.5 cm^{-1} resolution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Characteristics of a particular instrument, such as spectral response function shape and antenna pattern, can be simulated efficiently by multiplying the monochromatic radiances with a pre-calculated sensor response matrix [32].

Unfortunately, however, the straightforward approach typically requires thousands of individual RT calculations, one for each point in the frequency grid. For broadband IR instruments it is therefore only suitable for reference calculations, not for operational use.

2.4. The correlated k method

The correlated k method [33–35] is a method that is often used to reduce the number of frequency grid points for which monochromatic RT calculations have to be done. The basic idea of the method is to sort the frequency grid according to the absorption coefficient at the different frequencies. On the re-ordered frequency grid the spectrum is smooth and monotonic, so that it can be approximated with very few frequency grid points and linear interpolation between the grid points.

The advantages of the method are that it is efficient and conceptually easy. The disadvantage is that the exact sorting depends on pressure, temperature, and trace gas concentration. It is thus not straightforward to determine the optimal compromise for the frequency grid sorting, optimal in the sense that it minimizes the error over a
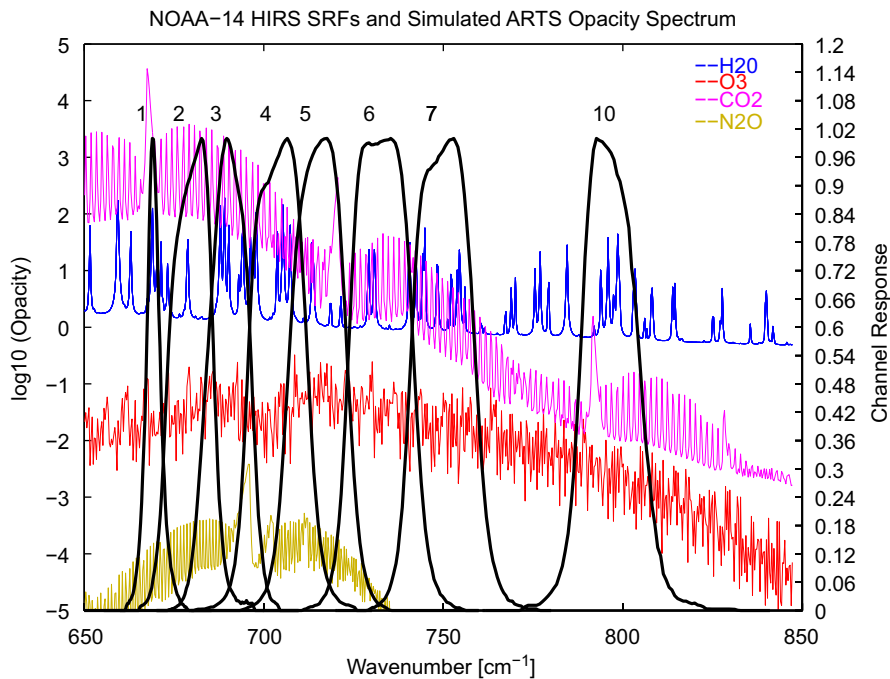

Fig. 2. Atmospheric zenith opacity spectra and HIRS spectral response functions. The left y-axis is for the opacity spectra, the right y-axis is for the channel response functions. Colored lines show the zenith opacity for key trace gas species for a midlatitude summer atmospheric scenario. The opacity calculation was done on a 0.02 cm^{-1} resolution frequency grid, but subsequently binned to 0.3 cm^{-1} resolution to get smoother curves. Black curves show channel response functions for the HIRS instrument on NOAA 14. They were normalized such that the peak of each function is 1. The figure shows Channels 1–7 and 10, see Fig. 3 for the other thermal radiation channels. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
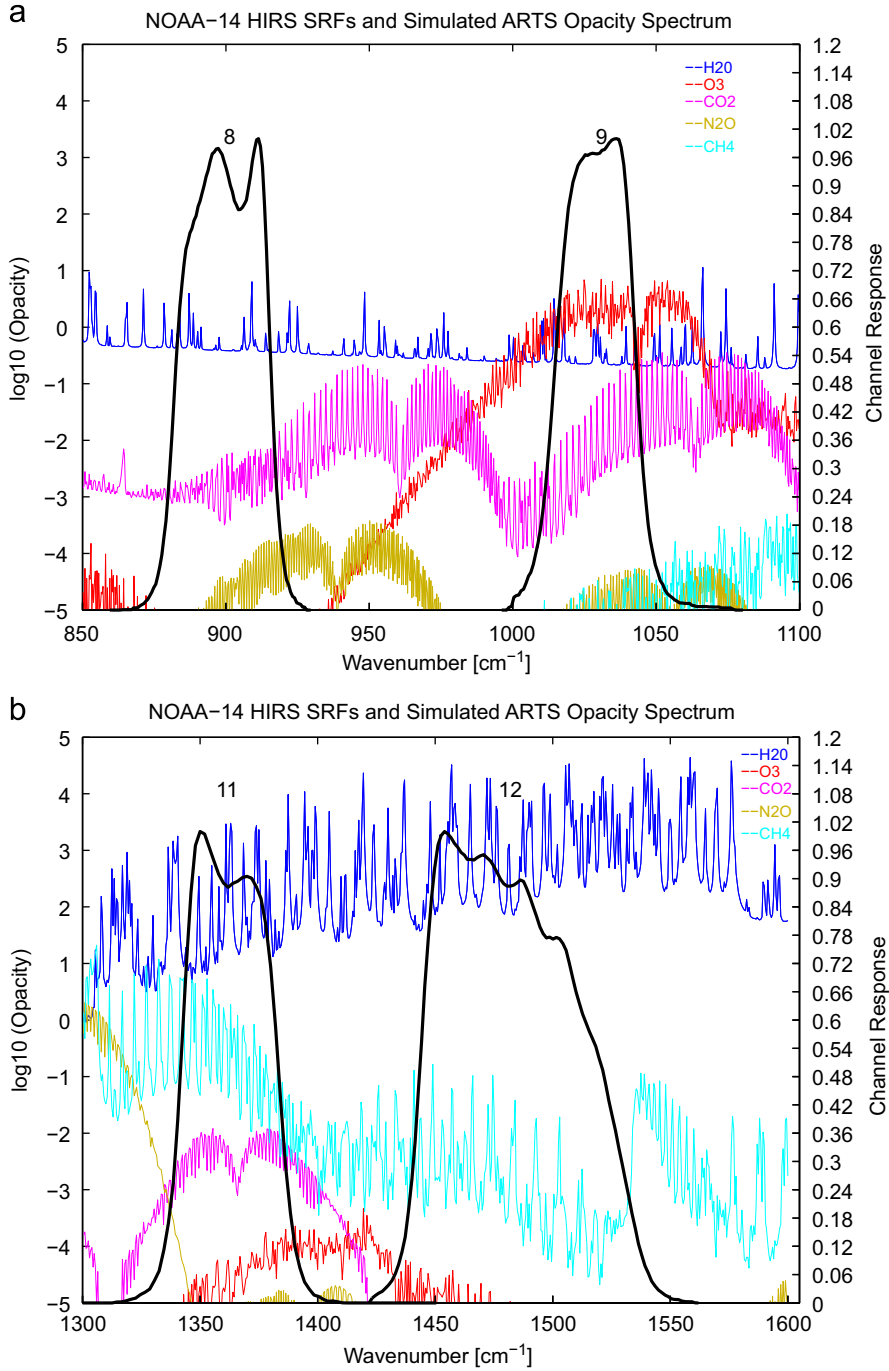
Fig. 3. As Fig. 2, but for Channels 8 and 9 (top) and 11 and 12 (bottom).

large ensemble of atmospheric cases and for different instrument viewing angles.

2.5. The weighted mean of representative frequencies (WMRF) approximation

The WMRF approximation introduced here is numerically quite similar to the correlated k method, but its derivation and theoretical justification are different. The simple assumption underlying this method is that the radiance at a single frequency will be representative for the radiance at some other frequencies (where atmospheric optical properties are similar). Thus, the integrated radiance for a broad instrument channel can be approximated as

$$I_{int} \approx \tilde{I}_{int} = \sum_{j=1}^{n} w_j I(v_j), \qquad (2)$$

with a set of n appropriately selected frequencies v_j and associated weights w_j. In contrast to the v_i in Eq. (1), the v_j here can be reduced, so that there can be a considerable saving of computational cost. (According to our results presented later, less than approximately 20 v_j are sufficient to accurately reproduce a reference calculation with several thousand v_i.)

We will call Eq. (2) the Weighted Mean of Representative Frequencies (WMRF) approximation. The authors of [2] also used Eq. (2) as a starting point for an RT algorithm that they call Optimal Spectral Sampling (OSS). We are not using that name here, since OSS refers to the entire RT calculation strategy, including the frequency selection. However, our method shares several key properties with OSS, as will be discussed further below.

The obvious questions about Eq. (2) are how to select the v_j and how to determine the associated w_j. One possible strategy would be to pick regular intervals in an absorption-sorted frequency grid and assume linear interpolation, in which case the method would be very similar to the correlated k method.

If a set of frequency grid points v_j is given, then one can easily determine the optimal weights w_j by multiple linear regression over a set of reference calculations for an ensemble of atmospheric states a_s. The optimal weights are those that minimize a cost function that is some measure for the error of the approximation.

Natural choices for the cost function are the absolute RMS error E_{abs} or the relative RMS error E_{rel}, defined as

$$E_{abs}(v_j, w_j) = \mathrm{rms}(\tilde{I}_{int}(a_s, v_j, w_j) - I_{int}(a_s)),$$

$$E_{rel}(v_j, w_j) = \mathrm{rms}\left(\frac{\tilde{I}_{int}(a_s, v_j, w_j) - I_{int}(a_s)}{I_{int}(a_s)}\right), \tag{3}$$

where rms() indicates the root mean square. As the radiance varies strongly with frequency, the relative error was chosen as the appropriate quantity to minimize.

Concerning the weights w_j, we demand them to be positive, consistent with our physical understanding of the problem. Negative weights would mean partial cancelation of terms, which is undesirable if the approximation is to generalize well. In practice, this is implemented by simply removing frequencies with negative weights from the list of active frequencies.

As pointed out by [2], it is desirable that the sum of all weights equals exactly one, so that Eq. (2) corresponds to a true weighted mean. The actual sum of the weights determined by regression is usually close to one, but not exactly one. (The maximum deviation that we encountered for the final weights without normalization was 0.0002.) To force the weight sum to be exactly one, we normalize the weights by dividing them by the actual weight sum value. This differs from the approach of [2], who do the regression only for frequencies 1 to $n-1$, and assign weight $w_n = 1 - \sum_{j=1}^{n-1} w_j$ to frequency n.

Since we know how to find optimal weights w_j for a given set of frequencies v_j, we are left with the problem of how to find the v_j. Unfortunately, the number of possible sets is combinatorially large, since there are

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} \tag{4}$$

ways to select a set of n frequencies from a total of N, where $\binom{N}{n}$ indicates the binomial coefficient. For example, there are approximately 2.6×10^{23} possibilities to select 10 frequencies out of a set of 1000. Hence, trying all possible sets is impractical. At the same time, the cost function $E(v_j, w_j)$ is very nonlinear, so that analytical methods to find the minimum are not applicable. Hence, a stochastic method is required.

One possibility is to guess appropriate frequencies, based on opacity, list of absorbing species, and other factors, and then fine tune the initial selection by trial and error. This was termed 'pseudo-k' approach by Frank Evans, and used for example to create training data for the proposed SIRICE and CIWSIR [18,19] sub-millimeter ice cloud satellite missions.

A more formal approach is described in Wiscombe and Evans [36], who apply the method to transmission instead of radiance. The problem to select representative frequencies arises similarly in that case. Their algorithm starts with $n = 1$ and determines the best frequency v_1 by an exhaustive search. Then more frequencies are added successively, each time doing an exhaustive search for the frequency that gives the largest improvement when added to the existing set. Some special measures are necessary to prevent the algorithm from being trapped in local minima.

In the OSS article, [2] describe an algorithm which they call Monte Carlo search, and which shares key properties with our algorithm. We shall come back to it below.

2.6. Simulated annealing

The simulated annealing method is described for example in Kirkpatrick et al. [1]. It is a method of randomly moving a simulated system through different states, in order to find a global minimum in some cost function, such as Eq. (3).

Assume that the system is initially in a state S_0 (for example a random state). In our application of the method each state S corresponds to a selection of frequency grid points. To be fully explicit here, we first calculate simulated spectra on a high resolution frequency grid for an ensemble of atmospheric conditions. The state S of the system is given by a selection of n frequencies among all the available frequencies. We also call the frequencies in S the active frequencies, and all others the inactive frequencies.

The number n of frequency grid points v_j is fixed and much smaller than the number N of frequency grid points v_i in the high resolution reference calculations. The optimization problem to be solved is to find the optimum set of frequencies for a given n. (Later, one can then increase n until a desired accuracy is reached.)

The system is fully described by the set of active frequencies S, because the associated weights can be calculated directly by linear regression over the ensemble

of atmospheric conditions. In practice the regression is also over the set of possible instrument viewing angles, since the goal is to find a set of frequencies that works for all viewing angles.

To implement the simulated annealing method, one needs to define a function

$$S' = \text{neighbor}(S), \tag{5}$$

which returns a new state S' of the system that is in some sense a neighbor of the original state S. In our implementation of the method this function exchanges one of the active frequencies with a random inactive frequency. We will call the operation of changing S to S' a move of the system.

For the new frequency set S' the weights w'_j are determined by multiple linear regression, and the cost E' is calculated according to Eq. (3). If E' is lower than E then the new frequency set S' is better than the original one S, and the move is therefore accepted. However, if one were to accept only moves where E is decreased, the algorithm would run into a local minimum of E very quickly, instead of finding a global minimum.

The key feature of the simulated annealing method is that sometimes the new state is accepted even if it results in an increase of E. This step is controlled by an acceptance probability function $p(E - E', T)$. The classical choice of this function is

$$p(E - E', T) = \begin{cases} 1 & E' < E, \\ \exp((E - E')/T) & E' \geq E. \end{cases} \tag{6}$$

This choice of p function, and the notation used, suggest an analogy with the physical process of annealing in metallurgy, where heating and slow cooling is used to create a more regular crystal structure in a material. In the analogy, E corresponds to the energy of the system, T to the temperature, and p resembles the Boltzmann probability factor. Note that the 'temperature' T as defined by Eq. (6) has the same unit as the error (or energy) E. Technically, the p function is used in our annealing code as follows: We do a test move and calculate p. Then we create a uniformly distributed random number x in the open interval $]0,1[$ with the Matlab 'rand' function. The move is accepted if $x < p$.

The algorithm starts with a high T, so that nearly all moves are accepted. Then T is gradually reduced, so that the system is pushed towards states with lower E, but without getting stuck in local minima. (The exact procedure is described in more detail below.) At the end, when T is very small, the system is effectively 'frozen in' and no more moves are possible. The algorithm will always lead to a state with very low E, compared to the entire state space. We call this unrigorously 'a global minimum' here. It should be noted, however, that it is unlikely to find the true global minimum. But for the application described here (and many others) the true global minimum is only of theoretical interest, while in practice any state with very low E is acceptable.

Our algorithm shares key properties with the OSS algorithm described by [2]. In OSS, new frequency sets are also accepted with a certain probability that is adjusted as the algorithm progresses. However, OSS retains from the Wiscombe and Evans [36] algorithm the idea to use the state $S(n)$ for n frequencies as the starting point for the search for the best state $S(n + 1)$. Our algorithm does not use this concept. For each n we start with a random state and a high 'temperature', and derive the solution by 'cooling down' the system, as described above.

For the study described here, the algorithm was implemented as a set of simple Matlab functions. The code is publicly available from our web page as part of the Atmlab software package (see Section 4 for address).

2.7. Example case HIRS

The HIRS instrument ideally serves to illustrate the method. High resolution reference spectra for this instrument were calculated with the RT software ARTS for a set of 42 different atmospheric states. The atmospheric states were those from Garand et al. [37]. The Garand data set contains pressure, temperature, and concentrations of the trace gases H_2O, O_3, CO_2, N_2O, CO, and CH_4. The spectra were simulated for five different viewing angles, covering the viewing angle range of HIRS.

The frequency grid for the reference calculations contained 72 717 frequencies in total, equally spaced in frequency (but only within the instrument bands). The HIRS instrument setup was the one that is distributed along with ARTS. It contains channel response functions for all different versions of HIRS (on different satellites) as distributed by NOAA/NESDIS (http://www.star.nesdis.noaa.gov/smcd/spb/calibration/hirs/hirssrf.html). Figs. 2 and 3 show these functions for NOAA 14.

Since some channels overlap, it is theoretically most efficient to simulate all channels simultaneously, so that some frequencies are used for more than one channel. We can take this into account in the optimization process, so the system state S is a selection of frequencies for all channels, not just for one. However, in that case we impose the constraint that the frequencies contribute to the sum of Eq. (2) only if they are inside the same channel. This constraint is believed to increase the robustness and generalizability of the method.

The example results shown in Sections 3.1 and 3.2 use the simultaneous optimization approach. However, for the final set of optimizations for operational use, discussed in Sections 3.3 and 3.4, we decided to use the more primitive approach of optimizing the grid separately for each channel. That option leads to faster convergence of the algorithm, at the cost of slightly more frequencies in the final grid.

3. Results and discussion

3.1. Annealing algorithm performance

Fig. 4 shows how the error E between the fast calculation and the reference calculation develops with time (iteration count). This example calculation tries to

find the best combination of 133 frequencies to cover all HIRS bands for NOAA 14.

The chosen number of frequencies here is arbitrary, since the aim is only to illustrate how the method works. The annealing algorithm will optimize the frequency selection for any (fixed) number of frequencies. This implies no loss of generality, since the whole procedure can be easily repeated with a different number of frequencies. Later on (in Section 3.3) we present calculations where frequencies have been successively added until a desired accuracy was reached.

The annealing run contained almost 4 million iterations, which were executed in blocks of 10 000. After each block, the RMS error for the block was calculated, together with the maximum and minimum errors. The blue shaded region in the figure is the area between the minimum and maximum errors; the RMS error lies in between.

The evolution of the block RMS error is used to adjust the 'temperature' T. If the error has not decreased between the last block and the block before, then T is reduced (in this case multiplied by 0.9). The idea of this cooling scheme is to only cool if the system has reached an equilibrium, marked by a steady error level. This avoids a too rapid cooling which would lead to the algorithm being trapped prematurely in a local minimum.

One practical problem is how to select the initial T at the start of the annealing run. We solve this by simulating one block of 10 000 iterations where we accept all moves and save their errors. Then we choose the initial T such that almost all of the moves would have been accepted.

Specifically, it was set here so that 99% of the moves would have been accepted.

Choosing a too low initial T will lead to the algorithm not really exploring the full range of possibilities, and should therefore be avoided. Choosing a too high initial T, on the other hand, will not affect the annealing result, but will waste computation time during which T is reduced to a more appropriate value. Our choice is quite conservative, which can be seen from the fact that at the beginning (at low iteration number) the error E decreases rapidly. This is due to the elimination of obviously bad frequency combinations, where for example some instrument channel has no frequency at all.

After this initial drop E changes only very slowly up to approximately iteration block 120, although T decreases steadily. From approximately iteration block 120 onwards the system 'condenses', because T is now so low that only rather good frequency combinations (that lead to low E) have a chance to be accepted.

Towards the end of the run the accepted moves get fewer and fewer. The frequency combination is near optimal, so it is unlikely to find a better one by random changes. At the same time T is now very small, so it is unlikely that a move that increases E is accepted. The algorithm is stopped if there are no more successful moves inside one entire block.

We find the convergence behavior that is seen in Fig. 4 to be general for all annealing runs that were carried out. It leads to a characteristic S shape in a log–log plot of E versus T (Fig. 5). The entire run took approximately 15 h on a PC.
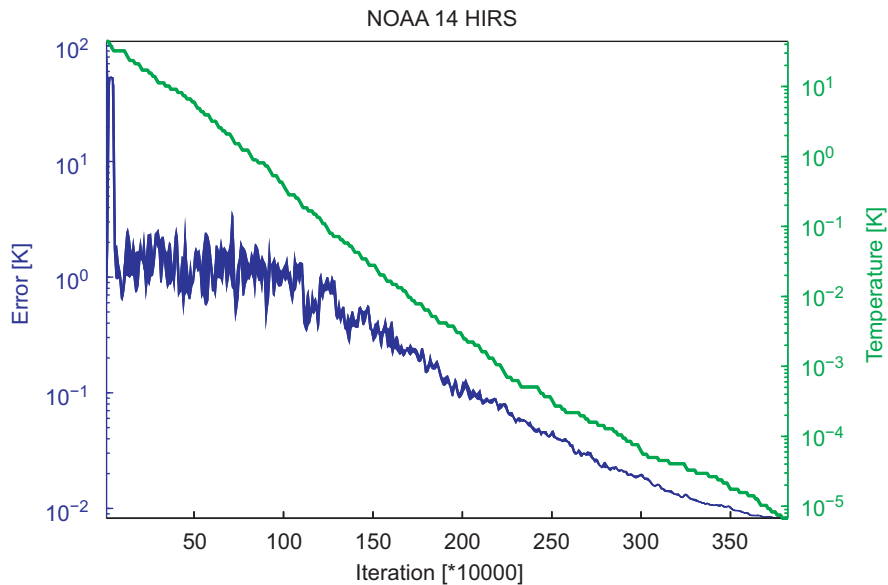

Fig. 4. Time development of 'temperature' T (green) and error E (blue) during one annealing run. Time is measured by iteration number. Parameters are evaluated and plotted once for every block of 10 000 iterations. The blue shaded region marks the area between maximum and minimum errors during each block. As an exception, the minimization was here done for the brightness temperature error in Kelvin, in order to have a more intuitive error measure. Due to this choice, the 'temperature' is also in Kelvin, since it has the same unit as the error that is minimized. Note that operational optimizations in Section 3.3 did not use this unit, but used the relative error in radiance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
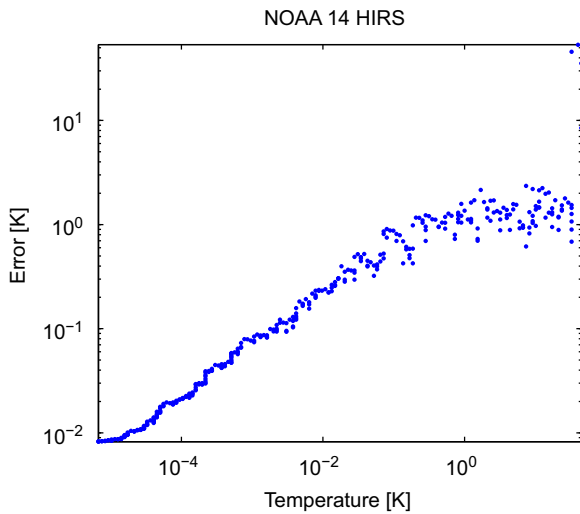
NOAA 14 HIRS

Fig. 5. Error E versus 'temperature' T. Each dot is for one simulation block of 10 000 iterations. Note the typical S shape, the curve is flat for very large or very small T, but steeper in the intermediate T range.

3.2. Example results for HIRS channel 12

We use the example of Channel 12 for HIRS on the NOAA 14 satellite to discuss the results of the annealing optimization. The top plot of Fig. 6 shows the high resolution atmospheric spectrum in blue. This is the spectrum for the near-nadir looking direction and the first of the 42 Garand atmospheric scenarios. The spectrum is dominated by water vapor absorption features in this frequency region. The spectral response function of HIRS Channel 12 on NOAA 14 is superimposed in green. The nine black vertical lines indicate the selected representative frequencies.

The selected frequencies cover nearly the entire range of observed brightness temperatures (T_b). What is more, they cover the T_b range quite uniformly, which is more apparent if the spectrum is sorted by increasing T_b, as shown in the bottom plot of the figure. Numbers mark identical frequencies.

Here the relationship to a correlated k method is clearly visible. Both methods lead to regularly spaced representative frequencies in the re-sorted frequency grid. The difference is that in our case the spacing is the outcome of the optimization procedure, whereas for a correlated k method it is prescribed.

An obvious question about Fig. 6 is why the frequency pairs 3–8 and 4–9 are so close together in the bottom plot. Would it not be better to spread out the frequencies more homogeneously? The answer is that the sorting depends on the atmospheric conditions. Fig. 7 shows that for Garand scenario 2 the sorting is quite different from scenario 1 (and the other 40 are again different). The position of the WMRF frequencies is derived by minimizing the RMS error over all 42 scenarios, so they are a reasonable compromise for each individual scenario.

However, there is something else that is odd about frequencies 8 and 9. Their weights (shown by the green bars in the bottom plots of Figs. 6 and 7) are very small compared to the other weights. If the frequency set is really optimal in some sense, there must be a reason for the presence of these frequencies despite their small weights. The reason must be that these frequencies represent important aspects of the spectrum for certain atmospheric conditions. Indeed, a closer analysis reveals that they play an important role in some of the atmospheric scenarios, as shown in Fig. 8. For the conditions of Garand scenario 7, frequency 9 alone represents the most transparent regions of the spectrum, where T_b is the highest. For the conditions of scenario 12, frequency 8 alone represents the most opaque regions of the spectrum, where T_b is the lowest.

3.3. Operational optimization runs

WMRF frequency grids and weights were derived for a variety of HIRS sensors and are included with the ARTS model distribution. Specifically, we have calculated them for the HIRS instruments on the satellites TIROS N, NOAA 6–19, and Metop A.

The training data were based on the 42 Garand profiles, but contained each profile for three different surface emissivities, 1.0, 0.8, and 0.6, so that there are 126 training cases in total. Five different HIRS viewing angles were simulated for each training case.

Each channel of each instrument was optimized separately. The optimization was done iteratively, starting with 5 frequency grid points for each channel. An annealing run was carried out to minimize the relative error in radiance. If the final fractional relative error was above 10^{-4}, the procedure was repeated with one frequency more. In other words, frequencies were added, until the desired accuracy was reached. For a scene brightness temperature of 300 K, the choice of 10^{-4} for the desired fractional relative accuracy corresponds to an absolute accuracy in brightness temperature of $10^{-4} \times 300\,K = 0.03\,K$, sufficient for accurate instrument simulation.

This iterative optimization procedure results in 5–22 frequencies for each channel. In the case of NOAA 14, the average number is 17 frequencies per channel.

3.4. Error analysis and test of generality

The derived operational frequency and weight sets for HIRS on NOAA 14 were used for testing the accuracy of the WMRF fast calculations. Over the 42 Garand scenarios, the RMS error of the WMRF fast calculations compared to the reference calculations in brightness temperature units was found to be 0.014 K. This very good agreement for the Garand scenarios is expected, since these are the scenarios used for training, so this is really only a test of our method's self consistency.

A more ambitious test is to compare reference and WMRF calculations for an independent set of atmospheric scenarios. The diverse data set by Chevallier et al. [38] was
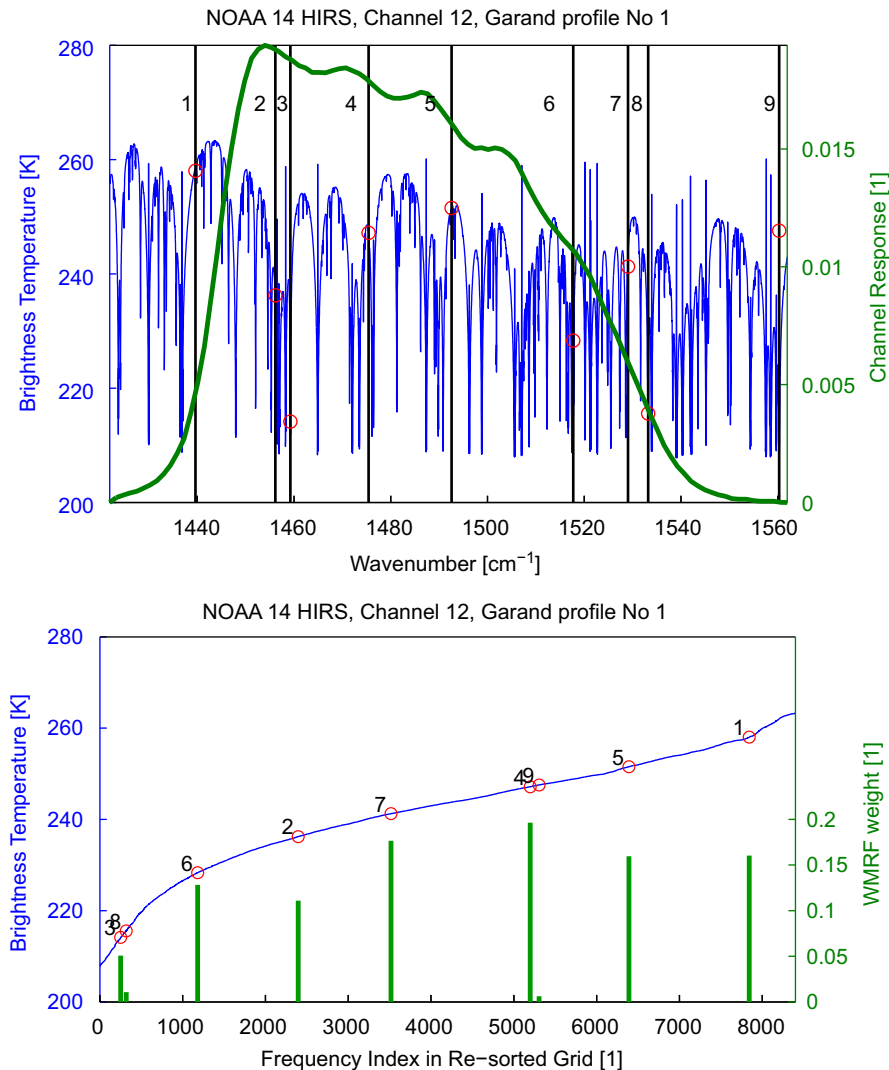
Fig. 6. Radiative transfer simulation for HIRS Channel 12 on NOAA 14, Garand atmospheric scenario 1. Top: The blue line shows the high resolution atmospheric brightness temperature spectrum for a near-nadir viewing direction. The green line shows the response function of HIRS Channel 12. Black vertical lines with red circles indicate the positions of the nine frequencies that were selected to represent this channel. Each frequency is labeled by a number. Bottom: The blue line shows the brightness temperature spectrum as a function of frequency index on the re-sorted frequency grid. The frequency is sorted, so that brightness temperature is monotonically increasing. Red circles mark the nine selected representative frequencies. Numbers can be used to identify the same frequency in top and bottom plots. (E.g., representative frequency 1 is on the left edge of the top plot, but on the right edge of the bottom plot.) Green bars indicate the weight for each frequency, which is found by linear regression as described in Section 2.5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

used for this purpose. More precisely, we started with the diverse Chevallier data sets for temperature, humidity, and ozone, which are a total of 15 000 atmospheric scenarios. From these 15 000 scenarios, 1000 were drawn randomly for the test.

The Chevallier scenarios do not contain all trace gases of the Garand scenarios: they are missing CO_2, CH_4, N_2O, and CO. For the test calculations, these were set to $VMR(CO_2) = 3.60e - 04$, $VMR(CH_4) = 1.7e - 6$, $VMR(N_2O) = 3.2e - 7$, and $VMR(CO) = 0$. These volume mixing ratios are typical tropospheric values for these species in the Garand scenarios, except for CO. For CO the assumption of

a constant VMR value is not appropriate, so this species was simply neglected. This has no impact on the results, since CO is only spectroscopically active in the short wave HIRS channels (mostly Channels 13 and 14), but has no transitions within the thermal HIRS channels.

ARTS was used to simulate HIRS measurements for five different viewing angles for the 1000 selected Chevallier scenarios. Both WMRF simulations and high resolution reference simulations were done. The WMRF simulations use on average 17 frequency grid points per channel, the reference simulations 3827. The total run time on Intel Core2 8-core CPUs was approximately 6 min for the
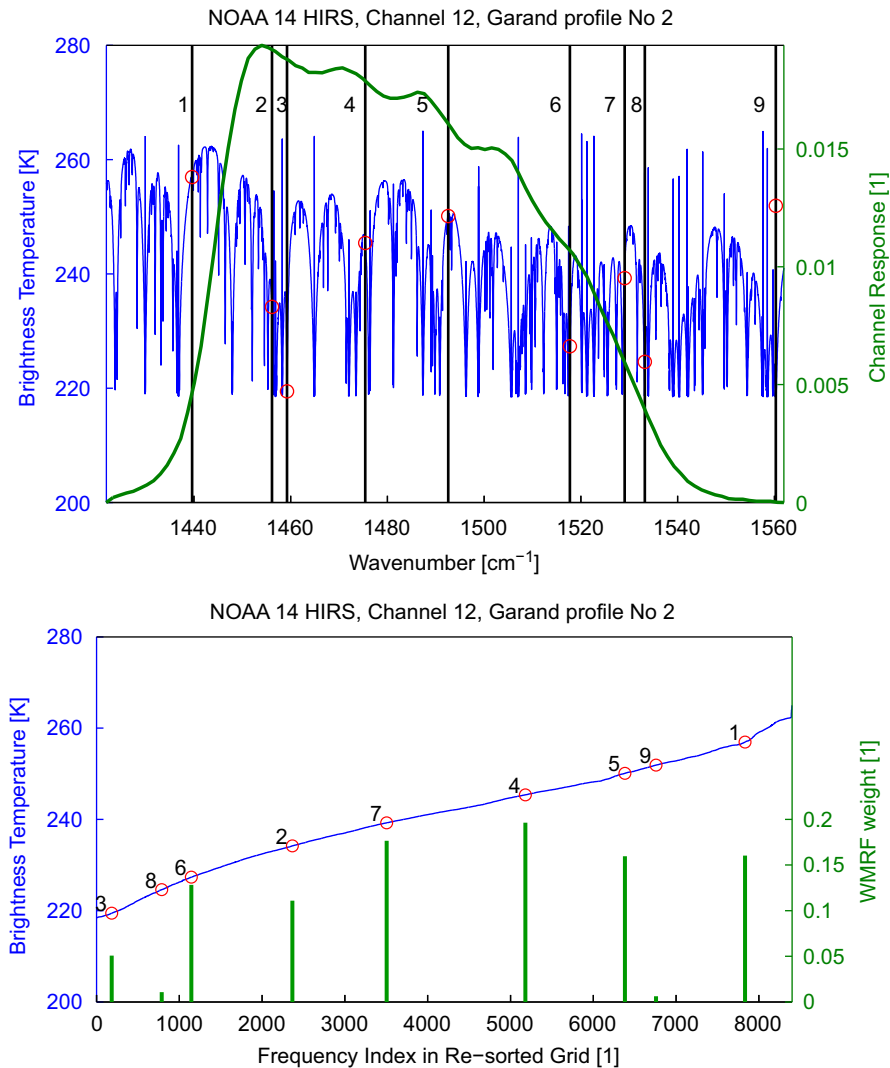
Fig. 7. As Fig. 6, but for a different atmospheric scenario, Garand scenario 2. Note that a slightly different sorting is necessary to make brightness temperature increase monotonically, hence some of the representative frequencies have changed their position in the bottom plot.

WMRF simulations, and 150 h for the reference simulations. (These run times include the line-by-line calculation of absorption coefficient tables, which is done once at the beginning of each batch run. The fast calculation was done in a single batch run, the reference calculation was broken down into 10 batch runs so that it could be run on a computer cluster.)

Fig. 9 summarizes the results of this exercise. It shows for each channel the mean and the standard deviation of the difference between fast calculation and reference calculation, taken over the 1000 atmospheric cases and all viewing angles. Differences for most channels are small, below 0.05 K. Only for Channel 1 the mean difference is as large as 0.3 K. The reason for this discrepancy is that this channel has a very high opacity (compare Fig. 2), so that it receives radiation from the upper stratosphere. The

Garand and Chevallier profiles are less consistent at these high altitudes than at lower altitudes.

3.5. Jacobians

The ARTS program can calculate analytical or semi-analytical Jacobians for various quantities, such as trace gas concentrations or temperature [3]. Each monochromatic radiance can be accompanied by monochromatic Jacobians, which are subject to the same post processing as the radiances, in order to simulate instrument antenna pattern and frequency response. In this context, it is interesting whether the selected WMRF frequency grid is suitable for Jacobian calculation, in addition to radiance calculation.
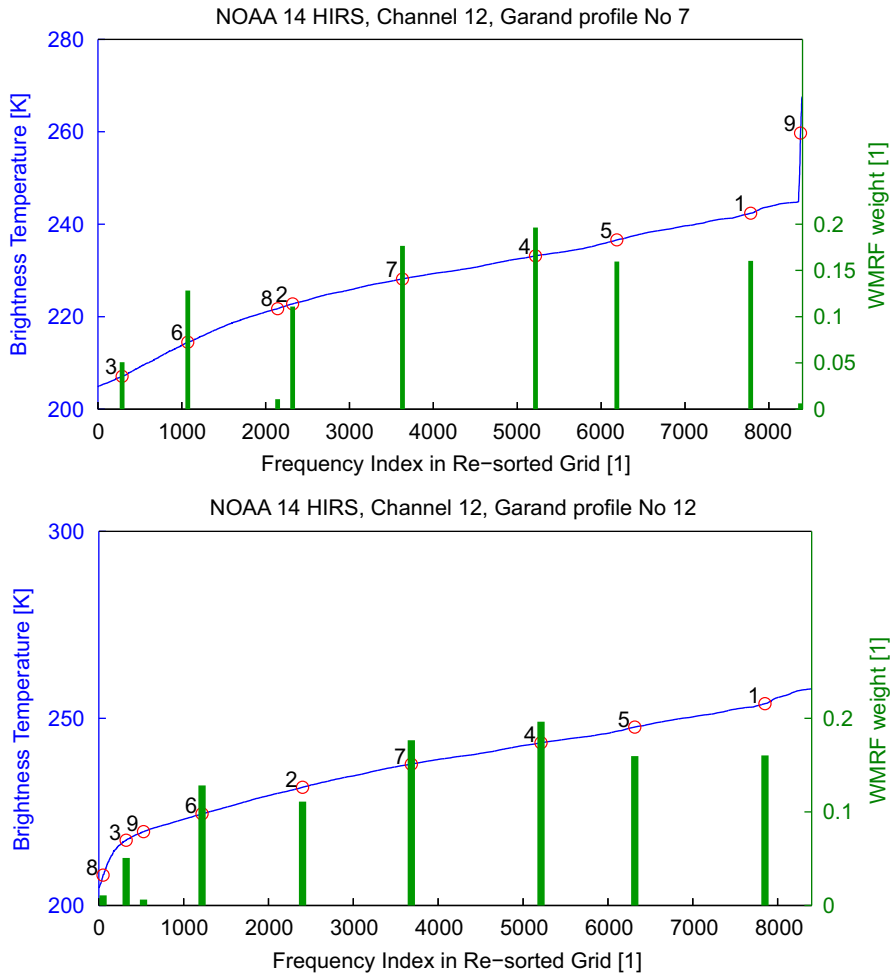
Fig. 8. Explanation for the presence of frequencies 8 and 9 in the selection. These plots are similar to the bottom plots of Figs. 6 and 7, but for Garand scenarios 7 (top) and 12 (bottom). In the top case frequency 9 is essential, in the bottom case frequency 8.
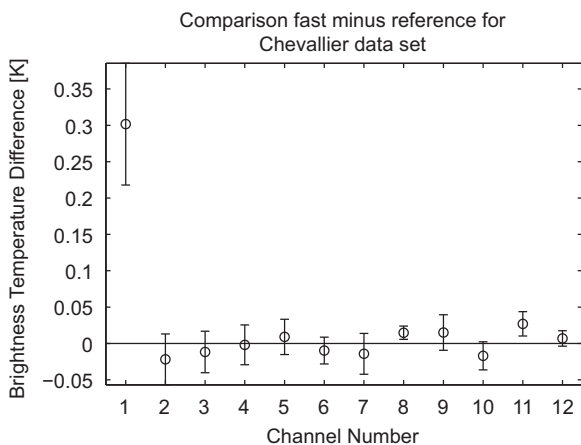

Fig. 9. Mean and standard deviation of the difference between fast calculation and reference calculation. Instrument characteristics are for NOAA 14 HIRS. The statistics are taken over 1000 randomly selected cases from the Chevallier data sets for humidity, temperature, and ozone variations. They are also taken over different viewing angles, representing the full HIRS scan range.

To check this we examine the Jacobians for water vapor for the first of the Garand scenarios. The left plot in Fig. 10 shows the Jacobians as calculated on the reference grid (solid curves) and on the WMRF grid (circles). Only HIRS channels sensitive to water vapor are shown. On first sight the agreement between fast and reference calculations appears excellent.

For a closer look, the right plot in Fig. 10 shows the relative difference ((fast-reference)/reference in percent). This plot reveals that at altitudes where Jacobians are small, relative differences approach 100%.

This result is not completely unexpected. Implicitly, the WMRF grid is optimized to represent those altitudes where changes have the strongest impact on the radiances. (Since the grid is selected to minimize the radiance error.) These are the altitudes where Jacobians are large. Frequencies, that represent altitudes where the Jacobian is small, tend to be optimized away.

One can conclude that the WMRF method as presented here is suitable for Jacobian calculations only if large relative errors outside the Jacobian peak area are acceptable. This issue should not be a problem
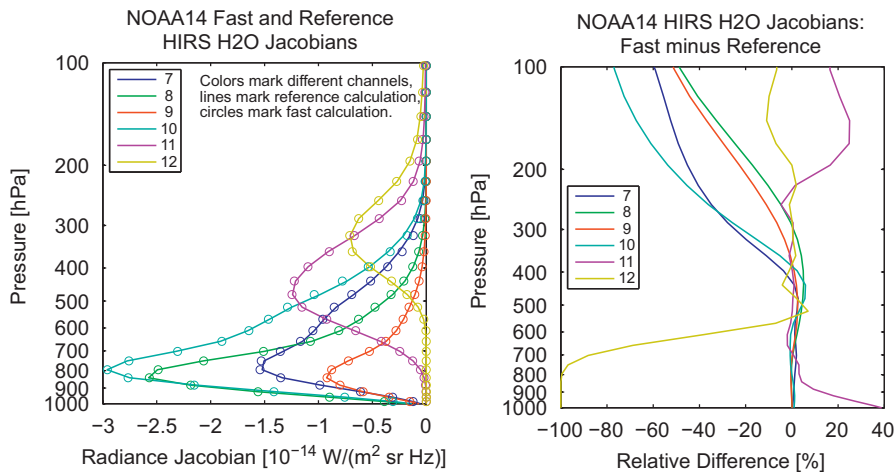
Fig. 10. Water vapor Jacobians for first Garand scenario for HIRS Channels 7–12. Left: Calculation on the reference grid (solid curves) and on the WMRF grid (circles). Right: Relative difference between the two calculations ((fast-reference)/reference in percent).

for most retrieval algorithms, but the user should be aware of it.

A possible solution to get precise Jacobians throughout would be to include the Jacobians in the cost function (Eq. (3)) for the simulated annealing algorithm. However, this has obvious drawbacks: Firstly, it is expected to lead to a larger number of frequencies in the WMRF grid (for the same radiance accuracy). Secondly, different Jacobian species would have to be taken into account for the different instrument channels, since not all species are active in each channel.

3.6. WMRF versus correlated k

A quantitative comparison of the new method to a traditional correlated k method is beyond the scope of this article. However, a few qualitative remarks are in order.

The WMRF method was shown to give good calculation accuracy at significantly reduced computational cost, compared to reference calculations. On theoretical grounds, we believe it to yield a speed increase that is at least comparable to a traditional correlated k method, but we have not shown this explicitly.

It is likely that for most users practical considerations drive the decision on what method to use. Compared to the correlated k method and the pseudo-k method, our new method has the advantage that the procedure is completely automatic. However, it has the disadvantage that the simulated annealing procedure requires considerable computation time. The calculation of high resolution reference spectra also requires considerable computation time, but is necessary for both methods.

4. Summary, conclusions, and outlook

We presented a method to efficiently simulate broadband IR radiometer measurements. It uses two basic ideas. Firstly, the channel radiance can be approximated by a weighted mean of the radiance at some representative frequencies, where the weights can be determined by linear regression. (We term this the Weighted Mean of Representative Frequencies (WMRF) approximation.) Secondly, a set of representative frequencies can be found by simulated annealing.

The method was shown to yield good accuracy (better than 0.05 K) with only few representative frequencies (5–22 per HIRS channel).

Representative frequencies and weights for the HIRS instrument are available under a GNU public license as part of the ARTS radiative transfer program at http://www.sat.ltu.se/arts/. Currently included are data for the instruments on TIROS N, NOAA 6–19, and Metop A. We plan to add further instruments, as their characteristics become available.

The optimization method itself was implemented as a set of simple Matlab functions. The only inputs required are instrument channel response characteristics, and high resolution reference spectra for a set of atmospheric scenarios. Hence, the functions can be used together with any radiative transfer program, and for any sensor with broadly similar characteristics as HIRS. The method should also be suitable for fast radiation flux or cooling rate calculations, but we have not yet tried this.

The functions are distributed under a GNU public license as part of the Atmlab software package, and can be downloaded from http://www.sat.ltu.se/arts/tools/.

An interesting open issue is, whether the presented method performs equally well for the calculation of cloudy spectra, as for the investigated clear sky spectra. Intuitively we believe so, but it has not yet been demonstrated. To study this is beyond the scope of the present article, since we are lacking at the moment a good set of high resolution reference calculations for a diverse set of atmospheric states. This issue will be the subject of further study.

Acknowledgments

References

[1] Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. Science 1983;220:671–80.

[2] Moncet J-L, Uymin G, Lipton AE, Snell HE. Infrared radiance modeling by optimal spectral sampling. J Atmos Sci 2008;65: 3917–34.

[3] Buehler SA, Eriksson P, Kuhn T, von Engeln A, Verdes C. ARTS, the atmospheric radiative transfer simulator. J Quant Spectrosc Radiat Transfer 2005;91(1):65–93. doi: 10.1016/j.jqsrt.2004.05.051.

[4] Emde C, Buehler SA, Davis C, Eriksson P, Sreerekha TR, Teichmann C. A polarized discrete ordinate scattering model for simulations of limb and nadir longwave measurements in 1D/3D spherical atmospheres. J Geophys Res 2004;109(D24):D24207, doi:10.1029/2004JD005140.

[5] Davis C, Emde C, Harwood R. A 3D polarized reversed Monte Carlo radiative transfer model for mm and sub-mm passive remote sensing in cloudy atmospheres. IEEE Trans Geosci Remote 2005;43(6):1096–101.

[6] Melsheimer C, Verdes C, Buehler SA, Emde C, Eriksson P, Feist DG, et al. Intercomparison of general purpose clear sky atmospheric radiative transfer models for the millimeter/submillimeter spectral range. Radio Sci 2005;RS1007, doi:10.1029/2004RS003110.

[7] Buehler SA, Courcoux N, John VO. Radiative transfer calculations for a passive microwave satellite sensor: comparing a fast model and a line-by-line model. J Geophys Res 2006;111:D20304, doi:10.1029/2005JD006552.

[8] Saunders R, Rayer P, Brunel P, von Engeln A, Bormann N, Strow L, et al. A comparison of radiative transfer models for simulating atmospheric infrared sounder (AIRS) radiances. J Geophys Res 2007;112:D01S90, doi:10.1029/2006JD007088.

[9] Eriksson P, Ekström M, Rydberg B, Wu DL, Austin RT, Murtagh DP. Comparison between early Odin-SMR, Aura MLS and CloudSat retrievals of cloud ice mass in the upper tropical troposphere. Atmos Chem Phys 2008;8(7):1937–48.

[10] Davis CP, Wu DL, Emde C, Jiang JH, Cofield RE, Harwood RS. Cirrus induced polarization in 122 GHz aura microwave limb sounder radiances. Geophys Res Lett 2005;32:L14806, doi:10.1029/2005GL022681.

[11] Perrin A, Puzzarini C, Colmont J-M, Verdes C, Wlodarczak G, Cazzoli, G, et al. Molecular line parameters for the MASTER (Millimeter wave Acquisitions for Stratosphere/Troposphere Exchange Research) database. J Atmos Chem 2005;50(2):161–205. doi: 10.1007/s10874-005-7185-9.

[12] Verdes CL, Buehler SA, Perrin A, Flaud J-M, Demaison J, Wlodarczak, G, et al. A sensitivity study on spectroscopic parameter accuracies for a mm/sub-mm limb sounder instrument. J Molec Spectro 2005;229(2):266–75. doi: 10.1016/j.jms.2004.09.014.

[13] Verdes C, von Engeln A, Buehler SA. Partition function data and impact on retrieval quality for a mm/sub-mm limb sounder. J Quant Spectrosc Radiat Transfer 2005;90(2):217–38. doi: 10.1016/j.jqsrt.2004.03.012.

[14] John VO, Buehler SA. The impact of ozone lines on AMSU-B radiances. Geophys Res Lett 2004;31:L21108, doi:10.1029/2004GL021214.

[15] Buehler SA, John VO. A simple method to relate microwave radiances to upper tropospheric humidity. J Geophys Res 2005;110:D02110, doi:10.1029/2004JD005111.

[16] John VO, Buehler SA. Comparison of microwave satellite humidity data and radiosonde profiles: a survey of European stations. Atmos Chem Phys 2005;5:1843–53 sRef-ID:1680-7324/acp/2005-5-1843.

[17] Melsheimer C, Heygster G. Improved retrieval of total water vapor over polar regions from AMSU-B microwave radiometer data. IEEE Trans Geosci Remote 2008;46:2307–22. doi: 10.1109/TGRS.2008.918013.

[18] Buehler SA, Jimenez C, Evans KF, Eriksson P, Rydberg B, Heymsfield, AJ, et al. A concept for a satellite mission to measure cloud ice water path and ice particle size. Q J R Meteorol Soc 2007;133(S2): 109–28. doi: 10.1002/qj.143.

[19] Jimenez C, Buehler SA, Rydberg B, Eriksson P, Evans KF. Performance simulations for a submillimetre wave cloud ice satellite instrument. Q J R Meteorol Soc 2007;133(S2):129–49. doi: 10.1002/qj.134.

[20] Rydberg B, Eriksson P, Buehler SA. Prediction of cloud ice signatures in sub-mm emission spectra by means of ground-based radar and in-situ microphysical data. Q J R Meteorol Soc 2007;133(S2): 151–62. doi: 10.1002/qj.151.

[21] Eriksson P, Jimenez C, Murtagh D, Elgered G, Kuhn T, Buehler S. Measurement of tropospheric/stratospheric transmission at 10–35 GHz for H_2O retrieval in low Earth orbiting satellite links. Radio Sci 2003;38(4):8069, doi:10.1029/2002RS002638.

[22] Pietranera L, Buehler SA, Calisse PG, Emde C, Hayton D, John, VO, et al. Observing CMB polarisation through ice. Mon Not R Astron Soc 2007;376:645–50. doi: 10.1111/j.1365-2966.2007.11464.x.

[23] Buehler SA, von Engeln A, Brocard E, John VO, Kuhn T, Eriksson P. Recent developments in the line-by-line modeling of outgoing longwave radiation. J Quant Spectrosc Radiat Transfer 2006;98(3): 446–57. doi: 10.1016/j.jqsrt.2005.11.001.

[24] John VO, Buehler SA, von Engeln A, Eriksson P, Kuhn T, Brocard, E, et al. Understanding the variability of clear-sky outgoing long-wave radiation based on ship-based temperature and water vapor measurements. Q J R Meteorol Soc 2006;132(621):2675–91. doi: 10.1256/qj.05.70.

[25] Deuber B, Morland J, Martin L, Kämpfer N. Deriving the tropospheric integrated water vapor from tipping curve-derived opacity near 22 GHz. Radio Sci 2005;40:RS5011, doi:10.1029/2007RG000233.

[26] Golchert SHW, Buschmann N, Kleindienst A, Palm M, Schneider N, Jønch-Sørensen, H, et al. Starting long-term stratospheric observations with RAMAS at Summit, Greenland. IEEE Trans Geosci Remote 2005;43(5):1022–7. doi: 10.1109/TGRS.2004.840660.

[27] Müller SC, Kämpfer N, Feist DG, Haefele A, Milz M, Sitnikov, N, et al. Validation of stratospheric water vapour measurements from the airborne microwave radiometer AMSOS. Atmos Chem Phys 2008;8:3169–83.

[28] Fiorucci I, Muscari G, Bianchi C, Girolamo PD, Esposito F, Grieco G, et al. Measurements of low amounts of precipitable water vapor by millimeter wave spectroscopy: an intercomparison with radiosonde, Raman lidar, and Fourier transform infrared data. J Geophys Res 2008;113:D14314, doi:10.1029/2008JD009831.

[29] Harlow RC. Airborne retrievals of snow microwave emissivity at AMSU frequencies using ARTS/SCEM-UA. J Appl Meteorol Clim 2007;46:23–35. doi: 10.1175/JAM2440.1.

[30] Adams IS, Gaiser P, Jones WL. Simulation of the Stokes vector in inhomogeneous precipitation. Radio Sci 2008;43:RS5006, doi:10.1029/2007RS003744.

[31] Smith WL, Woolf HM, Hayden CM, Wark DQ, McMillin LM. The TIROS-N operational vertical sounder. Bull Am Meteorol Soc 1979;60:1177–87.

[32] Eriksson P, Ekstroem M, Melsheimer C, Buehler SA. Efficient forward modelling by matrix representation of sensor responses. Int J Remote Sensing 2006;27(9–10):1793–808. doi: 10.1080/01431160500447254.

[33] Goody R, West R, Chen L, Crisp D. The correlated-k method for radiation calculations in nonhomogeneous atmospheres. J Quant Spectrosc Radiat Transfer 1989;42:539–50.

[34] Lacis AA, Oinas V. A description of the correlated k distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres. J Geophys Res 1991;96:9027–63.

[35] Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J Geophys Res 1997;102: 16663–82.

[36] Wiscombe WJ, Evans JW. Exponential-sum fitting of radiative transfer functions. J Comp Phys 1977;24:416–44.

[37] Garand, L, et al. Radiance and Jacobian intercomparison of radiative transfer models applied to HIRS and AMSU channels. J Geophys Res 2001;106(D20):24017–31.

[38] Chevallier F, Di Michele S, McNally AP. Diverse profile datasets from the ECMWF 91-level short-range forecasts. Technical Report, NWP SAF Satellite Application Facility for Numerical Weather Prediction, Document No. NWPSAF-EC-TR-010, Version 1.0, 2006.