Atmos. Chem. Phys. Discuss., 5, S4640–S4647, 2005
www.atmos-chem-phys.org/acpd/5/S4640/
European Geosciences Union

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Technical note: analytical estimation of the optimal parameters for the EOF retrievals of the IASI Level 2 Product Processing Facility and its application using AIRS and ECMWF data" *by* X. Calbet and P. Schlüssel

X. Calbet and P. Schlüssel

Received and published: 21 December 2005

article

Response to referee's review of ACPD-2005-0202

21 December 2005

Minor comments:

Page 2: line 23. Section 1. References introduced (Zhou et al.) in the paper as suggested by the referee.

Page 3: line 9. Section 1 and 7. Some comments and clarifications have been added in the final part of the introduction and the conclusions to make the goal of the paper clearer. It should be noted that the aim of this paper is not to compare retrievals using brightness temperatures with the ones using radiances. The aim of this paper is to have a method that will give the optimal parameters for any of them. The formulas to calculate these parameters have been obtained analytically and they can be applied to ANY EOF retrieval with the same algorithm as the one presented here regardless of whether you are using brightness temperatures, radiances, radiosondes or analyses. An example, where this technique can be applied is presented (AIRS brightness temperature and ECMWF analyses). I hope I have made all this clear in the introduction and the final conclusions. Revised Introduction and Conclusions are attached at the end of this document.

Page 4: line 7. Section 2. Some comments and a table (Table 2) with the scene selection criteria have been included.

ACPD

5, S4640–S4647, 2005

Interactive Comment

Figure 1: It has been explained in the text (Section 2) why results up to 100 hPa are shown. This is our main area of interest, the troposphere. In reality the retrievals have been performed up to 0.1 hPa but the results are not shown.

Figure 14 and 16: the axis have been extended to show the whole profile.

Major comments:

"While this results in closer agreement between AIRS and ECMWF, there needs to be some scientific discussion about what is the logic to justify this step of removing the bias".

There is not only a scientific reason but also a mathematical proof. The main reason is that measurements and "calculated" spectra are derived from slightly different sources and have different statistics. This leads to a bias. All this is shown in the mathematical derivation: how the bias is formed (Eq. 16) and how it is resolved (Eq. 38). Again, this concept has been hopefully expanded in the introduction and conclusions.

"Can the authors account for why the bias is present?" Yes, read the above and the modified introduction and conclusion.

"Can a bias, if not explainable, be removed from a data set so simply?" Ideally the bias should not exist. But the confrontation of the real radiative measurements with the rest of the system (the calculated radiative measurements: ECWMF+RTTOV8) unfortunately shows a bias. The sources of biases is explained in the modified introduction and conclusions (basically they come from instrument errors, radiative transfer model errors and poor representativenes of ECMWF analyses). Once the bias is detected it must be corrected to do proper retrievals. Again, there is a mathematical proof behind it.

"Does the bias correction of real observed radiances with model derived radiances, and validation of retrievals against model data only, represent an optimal approach to high resolution data inversion?" No. The bias and noise corrections are optimal in the

ACPD

5, S4640–S4647, 2005

sense that they are the best for this particular algorithm and data sets. Naturally, some analytical results from this paper should be valuable for other EOF retrieval methods. Other EOF algorithms and data sets can behave better than this one in this or other circumstances. Each method has its own pros and cons. See the added discussion at the end of the conclusion where I comment pros and cons of EOF retrievals using direct measurements as training.

1. Revised Introducion

A series of European satellites, known as Metop, will be launched in the frame of the EUMETSAT Polar System (EPS) in low Earth orbits. The first launch of the Metop satellites is planned for 2006 and will carry the Infrared Atmospheric Sounding Interferometer (IASI). IASI is a high-spectral-resolution infrared sounding instrument developed by the Centre National d'Etudes Spatiales (CNES) and based on a Fourier transform spectrometer. IASI spectra are represented by 8461 spectral samples, between 3.62 and $15.5\,\mu\mathrm{m}$, with a spectral resolution of $0.5\,\mathrm{cm}^{-1}$ after apodisation. Its spatial resolution is 25 km at nadir with an IFOV (Instantaneous Field of View) size of 12 km at a satellite altitude of 819 km. As part of EPS, EUMETSAT is developing the operational IASI Level 2 Product Processing Facility (IASI L2 PPF), which will generate atmospheric state retrievals from the IASI radiance spectra (Schlüssel 2005).

One of the retrieval techniques available in the IASI L2 PPF is based on Empirical Orthogonal Functions (EOF), which is a valuable and very computer efficient method. It consists in performing a linear regression of the principal components or EOF of the measured brightness temperature spectra and the atmospheric state parameters. In this paper, the particular EOF retrieval method developed for the IASI L2 PPF will be reviewed analytically and tested with real data available from the AIRS instrument.

AIRS is a high-spectral-resolution infrared sounder launched in May 2001 on board

ACPD

5, S4640–S4647, 2005

the NASA Aqua satellite (Aumann et al. 2003). It has a spectral coverage from 3.7 to $15.4\,\mu\mathrm{m}$ with a spectral resolution of 1200 ($\lambda/\Delta\lambda$) and a total of 2378 channels. Its spatial resolution is about 28 km at nadir with an IFOV size of 14 km.

The EOF retrieval method has been studied before with synthetically generated data (Huang and Antonelli 2001), but further problems arise when used with real data as is acknowledged by (Zhou et al. 2002). Namely, the existence of a significant bias between the measured and modeled derived radiance and the dominant influence of the radiative transfer model errors on the observational error analysis.

To make this paper more readable, the real world example data is presented throughout the analytical demonstrations, but conceptually this paper could be divided in two separate parts. The first one (sections 2 to 5) deals with the analytical derivation of the best parameters to be used in EOF retrievals. The demonstration is general enough to account for different types of EOF retrievals using the same algorithm as shown in this paper. It can be applied whether radiances or brightness temperature measurements are used. The method can also be applied whether it is calibrated and validated using numerical model analyses or using radiosonde data. The first condition to apply the analytical results is that it is only calibrated and validated with one set of atmospheric profiles, that is, either radiosondes or numerical model analyses, but not both at once. The second condition is that the "total" noise of the measurements has gaussian statistics. By "total" noise it is meant the observed minus "calculated" measurement standard deviation as shown in Fig. **??**. This "total" noise includes the instrumental noise, the forward radiative transfer model errors and the representativeness of the data used as the real atmospheric profiles. Once these two conditions are met, the analytical results show which bias corrections and noise figures are the optimal ones in the EOF retrievals.

The second part of the paper (section 6 and throughout sections 2 to 5) verifies the analytical results with a real world example, the EOF retrievals of the IASI L2 PPF using real AIRS spectra. In this particular example, AIRS brightness temperatures are the measured quantities and the atmospheric profiles are calibrated and validated against ECMWF analyses. It has been verified (not shown in this paper) that the noise of the observed minus calculated brightness temperatures do show gaussian statistics, and hence the analytical optimal bias and standard deviation corrections can be applied.

2. Revised conclusions

2.1. General

Given the specific algorithm shown in this paper, which consists of fitting a linear regression to the EOF components of synthetic spectral data, and a given amospheric data set, it has been proven analytically that the optimal retrieval is obtained by performing the following steps:

1. Obtain from the real atmospheric profiles and the radiative transfer model (in our case ECMWF analyses and RTTOV-8) the “calculated” spectra. These spectra are then subtracted from the observed measured spectra (AIRS). Finally the mean of this difference and its standard deviation is calculated.

2. When performing the linear regression of the training data a Gaussian noise component should be added to the training spectra with a standard deviation that matches the one above, that is, the one obtained from the difference of observations minus “calculated” spectra.

3. When performing the retrievals, the measured spectra (AIRS) should be bias corrected with the aforementioned value, that is, the average of the difference between the observation minus the “calculated” spectra.

The reason for the existence of a bias arise from the fact of using different sets of data for training and retrieval and from a divergence between observed and calculated radiative measurements with differing statistics. The origin of this "total" noise, and thus its bias and standard deviation, can come from instrument noise, errors in the radiative transfer model and poor representativeness of the atmospheric states (ECWMF analyses in this case). It is very difficult, if not impossible, to discriminate between these three and to detect which source is the most significant one with the data used in this paper.

The bias correction is critical for the success of the EOF retrievals. If these bias corrections are not applied, significant biases appear in the retrievals degrading them significantly (compare Figs. **??** and **??**).

Adding the optimal noise to the EOF retrievals is not critical and reasonable retrievals can be obtained without it (Fig. **??**). Although its addition improves the retrieval by a noticeable amount (compare with Fig. **??**). An added benefit to the use of the optimal noise is that the number of eigenvectors is not critical as long as it is high enough to reach the plateau observed in Fig. **??**. This is not the case when a smaller than optimal noise is added and thus the optimal number of eigenvectors must be found (Eq. **??** and Fig. **??**).

The optimal bias corrections and added noise that have been derived in this paper imply that to improve the EOF retrievals one must resort to either changing the overall algorithm or using other datasets, like for example, training the retrievals with latitude classified data or obtaining the real atmospheric profiles from another source such as radiosondes.

One drawback of this technique is that the retrievals will be fined tuned to whatever data we have used as real world atmospheric profiles (ECMWF in this case). The retrievals will try to resemble this real world data set.

ACPD

5, S4640–S4647, 2005

Interactive Comment

This technique has been tested with real data from 24 h of a randomly chosen data set (namely 8650 clear sky spectra on 6 October 2003 during nighttime and over ocean) and it has been optimized for this same data set. It is not exactly known how this technique can be extended to other dates, in the case that, for example, the biases change slightly with time. This effect could lead in the end to final biases when using the data for climatological purposes. This effect could be specially difficult to solve if the bias changes occur because of real atmospheric variability.

2.2. Other algorithms

To overcome the problem of the bias and noise corrections altogether an alternative EOF technique could be used, by using the same training data set as the one to be retrieved. The EOF retrieval could be trained with direct radiative measurements and radiosonde profiles for example. In this case the statistics of the training and retrieved data sets should be the same showing none of the problems studied in this paper. But if this ideal situation is not met and there is a statistical difference between the training dataset and the one used for retrievals a bias will show up (Eq. **??**). In this case part of the theoretical analysis derived in this paper could be used. Biases corrections could be derived in a similar way as shown here (Eq. **??**). If the standard deviations are also different, there will be a noise mismatch degrading the retrievals (Eq. **??**). Standard deviation corrections could be applied by adding noise to one of the real measurements until both covariances are matched (Eq. **??**).

Using real measurements for training is not exempt of drawbacks. The first one of them is that normally the set of satellite radiative data with collocated radiosondes measurements is usually scarce. This will give rise to probable differences between the statistics of the training data set and the retrieved one. Another one is that if the training data set is obtained in a specific region of the planet, it will not be global enough to perform universal retrievals, leading again to biases.

ACPD
5, S4640–S4647, 2005

Interactive Comment

EGU