

A principal component based  
version of the RTTOV fast  
radiative transfer model

Marco Matricardi

Research Department

April 2010

Submitted for publication in Quarterly J Roy Meteor Soc

This paper has not been published and should be regarded as an Internal Report from ECMWF.  
Permission to quote from it should be obtained from the ECMWF.



Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/publications/>

Contact: [library@ecmwf.int](mailto:library@ecmwf.int)

© Copyright 2010

**European Centre for Medium Range Weather Forecasts  
Shinfield Park, Reading, Berkshire RG2 9AX, England**

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director. Appropriate non-commercial use will normally be granted under the condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

## Abstract

A new version of the RTTOV fast radiative transfer model has been developed that exploits principal component analysis. The model is based on the computation of a database of line-by-line spectra for a large training set of diverse atmospheric situations. The principal component scores obtained from the eigenvectors of the covariance matrix of the simulated radiances are used as input data in a linear regression scheme where they are expressed as a linear combination of profile dependent predictors. The predictors consist of a selected number of polychromatic radiances computed using the standard RTTOV fast transmittance model. The linear regression scheme can then be used to simulate principal component scores and consequently reconstruct radiances for any input atmospheric profile. The principal component based model compares favourably to the conventional RTTOV model both in terms of speed and accuracy: the dimensionality reduction inherent in the use of principal component analysis makes the principal component based RTTOV much more computationally efficient and because of the highly linear relationship between the principal component scores and the independent variables used in the regression scheme, the principal component based RTTOV can reproduce the underlying line-by-line radiances to a much higher degree of accuracy. The availability of a principal component based fast radiative transfer will enable ECMWF to exploit the noise reduction capability of principal component analysis and investigate the direct assimilation of IASI principal component scores in the spectral regions affected by high instrument noise.

KEY WORDS: Satellite data assimilation; Numerical weather prediction; radiative transfer.

## 1 Introduction

The assimilation of high resolution radiances measured by the Atmospheric Infrared Sounder (AIRS) (Aumann et al., 2003) on the Earth Observing System (EOS) Aqua platform and the Infrared Atmospheric Sounding Interferometer (IASI) (Chalon et al., 2001) on the MetOp-A platform has produced a significant positive impact on forecast quality (McNally et al., 2006; Collard and McNally, 2008). The operational use of AIRS and IASI radiances at the European Centre for Medium-Range Weather Forecasts (ECMWF) is currently restricted to a selection of temperature sounding channels in the long-wave region of the spectrum and to a very limited number of humidity sounding channels in the main infrared water vapour band. The short-wave spectral region covered by IASI band 3 (2000-2760  $\text{cm}^{-1}$ ) contains excellent temperature sounding channels which could also be exploited for assimilation into a Numerical Weather Prediction (NWP) system. Channels in the IASI short-wave band are currently underused compared to similar channels in the long-wave region for a number of reasons, which include cloud detection and radiative transfer biases, day-night variations in data usability due to non-local thermodynamic equilibrium (LTE) effects, increased instrument noise. All these issues are under investigation with the aim of optimizing the use of the data.

Regarding the increased instrument noise, we observe that the spectra of high-resolution infrared sounders contain many channels that have very similar spectral signatures. Channels that are similar in content are highly correlated to each other. As discussed in Antonelli et al. (2004), the use of principal component analysis (PCA) can remove a significant fraction of the uncorrelated random error present in the observations by exploiting the high level of correlation among these channels. Based on PCA noise filtering, we are studying the feasibility of mitigating the instrument noise problem in IASI band 3 through the direct assimilation of principal components (PC). The direct use

of PCs in an NWP assimilation system requires an efficient radiative transfer (RT) model (also known as observation operator) that can simulate PCs directly given first-guess fields of temperature, water vapour, ozone, surface properties, and perhaps, at a later time, other variables.

The observation operator used at ECMWF to assimilate satellite radiance data is the Radiative Transfer for TOVS (RTTOV) model (Saunders et al. 1999, Matricardi et al., 2004). It uses a parameterization of the atmospheric transmittances that makes the model accurate and fast enough to fulfill the current NWP requirements of near real-time monitoring and assimilation of satellite radiance data. In this paper, we describe the development of a PCA based version of RTTOV (PC\_RTTOV). Although the rationale behind the development of PC\_RTTOV is the direct simulation of PCs, we observe that PC\_RTTOV also allows the simulation of spectral radiances. Because of the dimensionality reduction inherent to PCA (radiances can be reconstructed using a few linear combinations of PCs), PC\_RTTOV can be used to simulate radiances much more efficiently than the standard RTTOV model. This aspect assumes particular relevance, since the full exploitation of the wealth of information provided by current and future high spectral resolution sounders might require the use of an increased number of channels. In addition, the development at ECMWF of a comprehensive monitoring and forecasting systems for trace atmospheric constituents important for climate and air quality is likely to involve an increasing number of trace gas species in the RTTOV input state vector. The fulfillment of these demands will pose a significant computational burden on the ECMWF Integrated Forecast System (IFS). The availability of a much more computationally efficient RT model could thus become a limiting factor in the assimilation of high-resolution radiance data.

Fast radiative transfer models based on PCA have been described by Liu et al. (2006) and Havemann (2006). Although PC\_RTTOV exploits the same basic principles of PCA, it differs in many aspects from these models, most substantially in the scheme used to predict the PCs. Whereas Liu et al. and Havemann use a multiple linear regression algorithm where the independent variables are monochromatic radiances computed using an appropriate model, the linear regression scheme utilized in PC\_RTTOV uses profile dependent predictors which consist of a selected number of polychromatic radiances computed using the conventional RTTOV fast transmittance model. This approach allows the incorporation of PC\_RTTOV into the well-established RTTOV architecture. This is a very attractive option. In fact, it requires only a limited number of modifications to RTTOV, it has an obvious maintenance advantage, and makes the model suitable for utilization in a scheme where the adjoint physics is used (i.e. the model must give the correct increments when the state vector is perturbed).

The paper is organized as follows: in section 2 we discuss the basic principles of PCA and their application to the simulation of radiance spectra. In section 3 we describe the various stages involved in the training of PC\_RTTOV. The scheme used for the prediction of the PC scores is described in section 4 whereas its accuracy is discussed in section 5. In section 6 we illustrate and discuss the forward model error correlation matrix. Conclusions are given in section 7.

## 2 Basic principles of principal component analysis and its application to radiance spectra

PCA is a method that allows the reduction of the dimensionality of a problem by examining the linear relationship between all the variables contained in a multivariate dataset. The reduction of the dimension of the dataset is obtained by replacing the original set of correlated variables with a smaller number of uncorrelated variables called principal components. Since the new derived variables retain most of the information contained in the original data set, they can be used to efficiently represent the data.

Suppose our dataset consists of  $n$  observations on  $m$  variables arranged into an  $n$  by  $m$  data matrix  $X$ . The dataset can then be represented by the vector population  $\mathbf{x} = (x_1, x_2, \dots, x_m)^T$  (here  $^T$  denotes the transpose). If  $C$  is the  $m$  by  $m$  covariance matrix of the data matrix  $X$ , and  $A$  is the  $m$  by  $m$  matrix formed by the eigenvectors of the covariance matrix arranged as row vectors in descending order according to the magnitude of their eigenvalues, the PCs,  $\mathbf{y}$ , of the vector population can be written as:

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad (1)$$

The eigenvectors represent the directions of maximum variance in the data; consequently, each PC gives the linear combination of the variables that provides the maximum variation. The PCs are orthogonal, hence uncorrelated (this does not imply they are statistically independent), and the values associated to each observation are known as PCs scores. If  $\lambda_i$  is the eigenvalue associated with the  $i^{\text{th}}$  eigenvector, then the value of  $\lambda_i / \sum_{i=1}^m \lambda_i^2$  gives the proportion of variation explained by the  $i^{\text{th}}$  PC.

Since by definition the matrix  $A$  is orthogonal (hence not singular), on the basis that the inverse of an orthogonal matrix is equal to its transpose, we can write:

$$\mathbf{x} = \mathbf{A}^T \mathbf{y} \quad (2)$$

Equations (3) and (4) can be written in discrete notation form as:

$$y_{i,j} = \sum_{k=1}^m A_{i,k} x_{k,j} \quad (3)$$

$$x_{i,j} = \sum_{k=1}^m A_{k,i} y_{k,j} \quad (4)$$

where  $i$  represents the  $i^{\text{th}}$  value and  $j$  is the  $j^{\text{th}}$  observation. A number of PCs,  $p$ , fewer than  $m$  can often represent most of the variation in the data. We can then reduce the dimension of the problem by replacing the  $m$  variables with the first  $p$  PCs. In many applications, the choice of the number of dimensions is based on the total variation accounted for by the leading PCs and it will in general depend on specific aspects of the dataset.

In this paper, the multivariate dataset consists of high-resolution polychromatic radiances calculated using an accurate line-by-line (LBL) model for a sample of diverse atmospheric situations (i.e. the observations). In principle, we can compute PCs scores for any independent observation by projecting

the corresponding LBL radiances on the fixed basis of eigenvectors obtained from the covariance matrix of the training dataset. This algorithm, however, involves the very time consuming calculation of the LBL radiances and consequently we cannot envisage its use in an operational environment where we require the near-real time simulation of PC scores. For this scope, we need to develop a computationally more efficient algorithm. The same algorithm will also allow the reconstruction of the simulated radiances at a fraction of the cost required by a LBL model.

Based on what is discussed above, the main steps involved in the development of a PCA based fast RT model are: the generation of a training database of LBL radiances and the development of an algorithm to predict accurate PCs scores for any atmospheric profile not included in the training dataset. The methodologies adopted in this paper are described in detail in the next two sections.

### **3 The training of the PCA based fast model**

#### **3.1 The atmospheric profile training set**

The profiles used to generate the multivariate dataset of simulated atmospheric radiances should represent the range of variations in temperature, absorber amount and surface parameters found in the atmosphere. To this end, we have utilized the profiles contained in the database described by Chevallier et al. (2006). These profiles were sampled from a large dataset produced using the operational suite of the ECMWF forecasting system (cycle 32R2) for the period July 2006-June 2007. In addition to profiles of temperature, water vapour and ozone, the database contains ancillary information on surface properties.

The Chevallier et al. (2006) database contains 12,564 profiles. As suggested by Chevallier (2002), the reduction of the number of profiles to a size that is manageable for LBL computations can be achieved by randomly sampling the profiles since this would not change the probability distribution of the profiles contained in the original dataset. We have then randomly sampled an arbitrary number of 1000 profiles and have complemented them with a number of profiles selected manually to cover the extreme values in the dataset since they are very difficult to obtain when making a random selection. The inclusion of the additional profiles resulted in a total number of 1039 profiles. In this study, the training of the fast model does not include land surface observations. Consequently, we only selected profiles over the sea surface. All the profiles have been interpolated to the fixed pressure grid of 101 levels specified by Strow et al. (2003) and were used to perform LBL radiance computations using the pressure grid that extends from the top of the atmosphere to the nearest surface pressure value. The statistical characteristics of the training profiles are shown in figures 1, 2, and 3 for temperature, water vapour and ozone respectively.

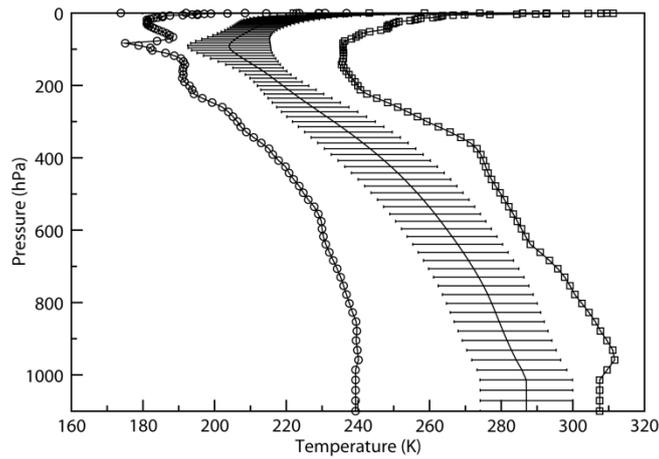


Figure 1. Statistics of the temperature profiles in the 1039 profile training set. The minimum and maximum values are shown as circles and squares respectively. The solid black curve is the mean value and the error bars have a width of twice the standard deviation.

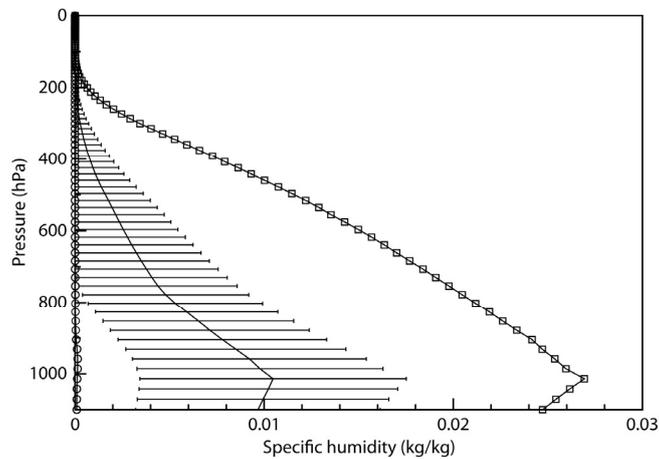


Figure 2. Statistics of the water vapor profiles in the 1039 profile training set. The minimum and maximum values are shown as circles and squares respectively. The solid curve is the mean value and the error bars have a width of twice the standard deviation.

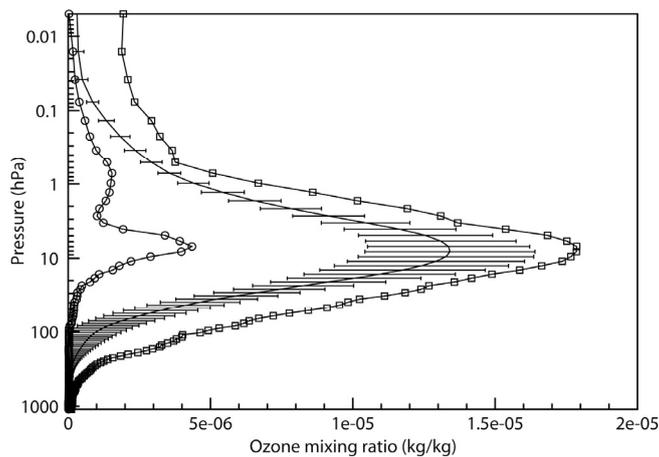


Figure 3. Statistics of the ozone profiles in the 1039 profile training set. The minimum and maximum values are shown as circles and squares respectively. The solid curve is the mean value and the error bars have a width of twice the standard deviation.

In addition to the vertical profiles of temperature and absorber amount, the LBL computations require the values of surface temperature, skin temperature, surface pressure and wind speed at 10 m. In our 1039 profile dataset, values of the surface pressure vary typically between 990 and 1020 hPa with minimum and maximum values attained at 940 and 1040 hPa respectively. Regarding surface temperature and skin temperature, they vary between 273 and 303 K with values as small as 240K observed mostly over sea ice. The computation of the emissivity used in the LBL computations requires the knowledge of the 10 m wind speed. In our dataset this quantity varies between 0 and 40 m/s. However, since it is estimated (Watts et al., 1996) that for wind speeds greater than 15 m/s the emissivity model used in our LBL computations can be invalidated by the presence of foam, any value greater than 20 m/s was scaled down to a values smaller than 20 m/s.

Finally, for testing purposes, we have generated a similar database of 1039 profiles independent of the profiles used to train PC\_RTTOV.

### 3.2 The LBL database of atmospheric radiances

The database of LBL radiances used in this study was calculated using version 11.1 of the LBLRTM (LBLRTM\_v\_11.1) LBL model developed at AER (Clough et al., 2005). Our computations have been carried out assuming clear sky conditions over the sea surface in absence of solar radiation. Monochromatic radiances were computed at a resolution of  $0.001 \text{ cm}^{-1}$  for each atmospheric profile and five values of the zenith angle that cover the viewing geometry of IASI and AIRS (namely the angles for which the secant has equally spaced values from 1 to 2). This choice corresponds to a total number of 5195 observations. To compute the spectral emissivities used in the LBL computations we have modified the model by Masuda et al. (1988) by including the effect of surface reflected emission using the methodology described in Wu and Smith (1995) and applying the salinity correction by Pinkley and Williams (1976).

The LBL calculations include 17 atmospheric constituents. The only species that can vary are  $\text{H}_2\text{O}$  and  $\text{O}_3$ . The other species are held constant and include  $\text{CO}_2$ ,  $\text{N}_2\text{O}$ ,  $\text{CO}$ ,  $\text{CH}_4$ ,  $\text{NO}_2$ ,  $\text{SO}_2$ ,  $\text{NO}$ ,  $\text{N}_2$ ,  $\text{O}_2$ ,  $\text{HNO}_3$ ,  $\text{OCS}$ ,  $\text{CCl}_4$ ,  $\text{CF}_4$ ,  $\text{CCl}_3\text{F}$  (CFC-11) and  $\text{CCl}_2\text{F}_2$  (CFC-12). For  $\text{CO}_2$ ,  $\text{N}_2\text{O}$ ,  $\text{CO}$  and  $\text{N}_2\text{O}$  we use the climatological profiles described in Matricardi (2008) whereas for the other species we use the AFGL atmospheric constituent profiles with the exception of  $\text{HNO}_3$  and  $\text{NO}$  for which we have utilized profiles generated using the MOZART chemical transport model (Hauglustaine et al., 1998). The profiles of  $\text{CCl}_4$ ,  $\text{CF}_4$ ,  $\text{CCl}_3\text{F}$  (CFC-11),  $\text{CCl}_2\text{F}_2$  (CFC-12) were scaled to reflect present-day concentrations using the values tabulated in Table 1.

Table 1. Present-day concentrations for CFC's and Halons included in the LBLRTM computations.

CFC-11 (ppbv)	CFC-12 (ppbv)	CCl4 (ppbv)	CF4 (ppbv)
246	540	74	90

For the LBL computations, we have created a molecular database that blends line parameters obtained from different sources. This choice is based on the results shown in Matricardi (2007) which suggest that in a number of spectral regions the use of alternative line parameters can result in a better agreement between observations and simulations. The blended database is based on the AER file supplied with the LBLRTM\_v11.1 package. This file is largely drawn from HITRAN2004 and includes updates up to 1/1/2007. However, for the ozone region between 1000 and 1080  $\text{cm}^{-1}$  we use ozone line parameters from HITRAN\_2000 (Rothman et al., 2003) whereas in the spectral region between 1700 and 2400  $\text{cm}^{-1}$  we use water vapour line parameters from GEISA\_2003 (Jacquinet-Husson et al., 2005).

Finally, we have computed the polychromatic radiances for the channels of IASI and AIRS by convolving the monochromatic radiances with the specification of the spectral response function available at the time.

### 3.3 The radiance reconstruction error

As discussed in section 2, the development of PC\_RTTOV requires the computation of the eigenvectors of the covariance matrix of the training dataset and the generation of the PCs scores for each observation in the same dataset. We can then reconstruct any spectrum to a given accuracy using a truncated number of principal components,  $p$ . This number will in general depend on the number of observations,  $n$ , and variables,  $m$ , in the dataset. In practical terms, we require that the value of  $p$  is such that for each channel,  $i$ , the radiance reconstruction error,  $\varepsilon_i$ , is below the instrument noise,  $\eta_i$ :

$$\varepsilon_i = \sqrt{\frac{\sum_{j=1}^n \left[ \sum_{k=1}^m A_{k,i} \mathcal{Y}_{k,j} - \sum_{k=1}^p A_{k,i} \mathcal{Y}_{k,j} \right]^2}{n}} < \eta_i \quad \forall i \quad (5)$$

Note that in this section (and throughout the remainder of the paper) radiances are expressed in units of black body equivalent brightness temperature (K).

We have computed the radiance reconstruction error,  $\varepsilon_i$ , for different values of  $p$ . Figure 4 shows how the radiance reconstruction error for IASI varies with the number of PCs. The solid line represents the error whereas the dotted line is the instrument noise. Note that for each channel the instrument noise has been computed at a scene temperature corresponding to the mean radiance. For all cases that we have tested, the reconstruction error is below the instrument noise for all channels and decreases steeply with the number of principal components. Results in Figure 4 refer to the observations in the training set. The reconstruction of the radiances for the profiles in the independent set yields almost identical results. In PC\_RTTOV the number of principal components is a tunable parameter. Although it can be shown that the use of a number of principal components less than 100 can still fulfill Eq. (5), we do not want reconstruction errors to become a significant fraction of the LBL model errors (see Matricardi (2009) and Shepard et al. (2009) for details on the LBLRTM spectroscopic errors), above all in IASI band 3. Consequently, we will restrict the use of the principal components to a number not less than 100.

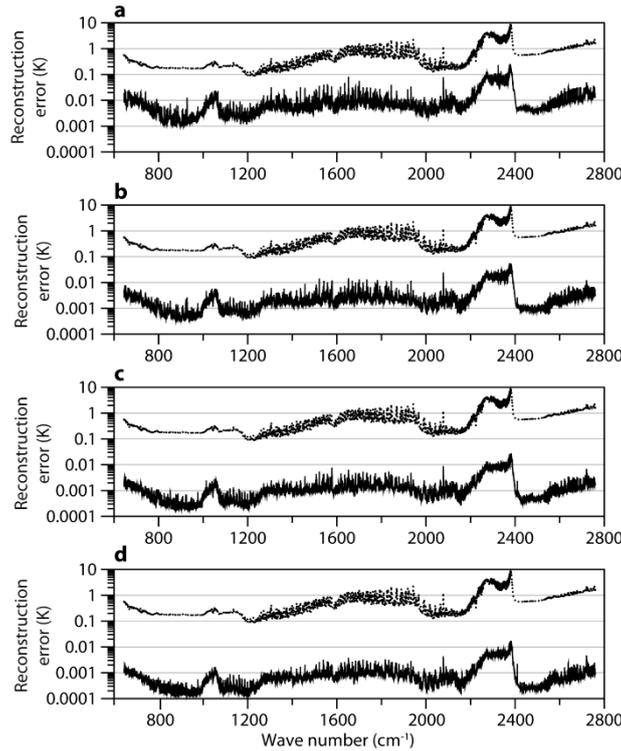


Figure 4: The reconstruction error for the IASI channels using (a) 100 principal components, (b) 200 principal components, (c) 300 principal components, and, (d) 400 principal components. The dotted line is the instrument noise computed at a scene temperature corresponding to the mean radiance.

#### 4 The PCs scores statistical regression algorithm

The use of the PC scores of a radiance spectrum in NWP assimilation requires the computation of the scores for any input profile. The scope of this section is to describe the development of an algorithm that can perform this task very efficiently. To this end, we have fitted an equation to the values of the PC scores obtained from the training dataset of LBL radiances described in the previous section. In principle, our equation could relate the principal component scores to independent variables (or predictor variables) that consist of a combination of profile dependent quantities such as temperature and absorber amount. However, the theory of PCA briefly discussed in section (2) indicates that there is a linear functional relation between PC scores and polychromatic radiances. It is then natural that we try to fit an equation in the form of a polynomial that is linear in the polychromatic radiances of the  $m$  channels of the selected sensor. The efficiency of the algorithm hinges on the rapid computation of the polychromatic radiances. In this paper, we perform this task using the conventional RTTOV fast radiative transfer model. We should note that RTTOV has been trained using exactly the same LBL model, molecular database, and vertical pressure grid as presented in this paper. The implications of using a different pressure grid for the radiance computation are discussed in section 5. Our problem then reduces to the computation of coefficient estimates in the multiple linear regression model defined by:

$$\mathbf{y} = \tilde{\mathbf{X}}\mathbf{b} \quad (6)$$

where  $\mathbf{y}$  is the vector of  $n$  observation of the dependent variable (i.e. the principal component scores of the LBL spectra),  $\tilde{\mathbf{X}}$  is a  $n$  by  $m$  matrix of the independent predictors variables (i.e. the RTTOV radiances) and  $\mathbf{b}$  is a vector of length  $m$  of unknown coefficients that must be estimated.

In a linear regression model, the coefficient estimates describe the characteristics of the population from which the sample is taken and are interpreted as a measure of the true characteristics of a population. However, predictor variables can be linearly related to another predictor variable or to a subset of variables and this can result in a value of the estimated coefficients in a sample that can differ significantly from the true value in the population. Consequently, the findings for a sample can be difficult to replicate in an independent sample from the same population. It is then important that for the estimation of the coefficients, the effect of co-linear relations between the predictor variables is reduced as much as possible.

In our regression model, the predictor variables are potentially the many thousands of channels in the spectra of IASI and AIRS (8461 and 2378 channels respectively). These spectra contain many channels that have very similar spectral signatures. As discussed above, channels that are similar in content can cause problems in the regression since they can be very highly correlated. To reduce the effect of co-linear relations between the predictor variables, we have arranged the channels in clusters that contain highly correlated channels. To calculate the clusters we have followed the approach by McMillin and Goldberg (1997). The first step in the calculation of the clusters involves the computation of the correlation matrix of the RTTOV channel radiances for the training profile dataset. We then search for the channel with the largest standard deviation (lead channel) and find all the channels that are correlated to the lead channel above a given threshold. The channel cluster thus obtained is removed from the selection pool and the process is repeated until no more clusters can be found. At the end of this process, the first,  $d$ , lead channels become the,  $d < m$ , predictors variables in the regression.

If we fix the number of predictor variables, we can estimate the PC\_RTTOV gain in computational efficiency (i.e. the ratio between the RTTOV and the PC\_RTTOV computational time),  $g$ , as follows. For the standard RTTOV model, the bulk of the radiance computation consists of the calculation of the layer optical depths. The calculation of the optical depths for a single channel requires a number of linear operations equal to  $l \cdot s \cdot pr$ , where  $l$  is the number of vertical layers,  $s$  is the number of absorbing species, and  $pr$  is the number of predictors per layer. Typically,  $l=100$ ,  $s=7$ , and,  $pr=12$ , therefore, the computation of the optical depths for  $c$  channels requires no less than  $o = c \cdot 8400$  operations. In the case of PC\_RTTOV, the radiance simulation for  $c$  channels involves the computation of the values of the independent variables ( $d \cdot 8461$  operations) and the radiance reconstruction ( $p \cdot (c + d)$  operations) for a total of  $opc = d \cdot 8461 + p \cdot (c + d)$  operations. We can then give a conservative estimate of  $g$  dividing  $o$  by  $opc$ :

$$g \cong \frac{c}{d + p(c + d) / 8400} \quad (7)$$

Results for IASI are shown in Figure 5 where values of  $g$  are plotted for  $c = 8461$  (i.e. the whole spectrum) and a choice of 300, 400, 500, and, 600 predictor variables. We see that for this choice of parameters, we estimate that the gain in computational efficiency varies between 8 and 21.

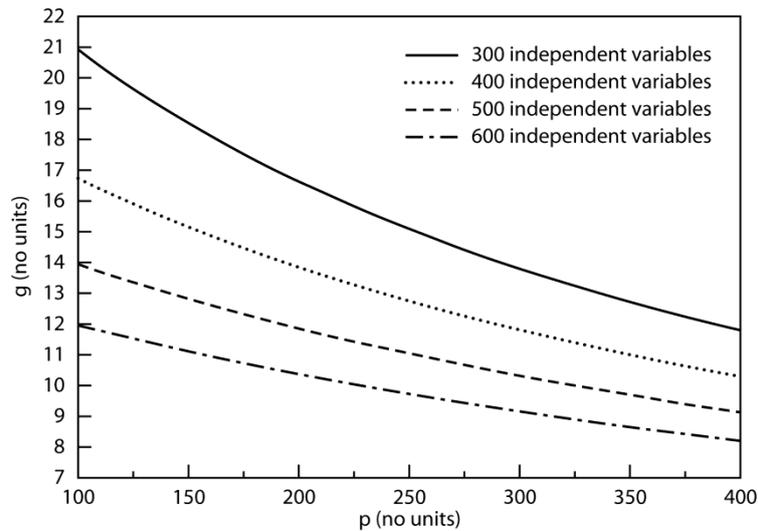


Figure 5: The gain in computational efficiency of PC\_RTTOV over RTTOV.

Note that  $g$  decreases with  $c$  and reaches the value of 1 for a number of channels,  $\tilde{c}$  equal to:

$$\tilde{c} = \frac{d(1 + \frac{p}{8400})}{1 - \frac{p}{8400}} \quad (8)$$

The number of lead channels and the number and distribution of channels in each cluster depend on the threshold value used to specify the level of correlation. A too high threshold can result in a large number of highly correlated lead channels whereas a too low threshold can result in too few lead channels. To address the problem posed by the choice of the optimal value of the correlation threshold, we have used an objective criterion that consists in the selection of the value of the threshold that minimizes the fitting root-mean-square (rms) error defined as:

$$rms = \left[ \frac{\sum_{i=1}^m \sum_{j=1}^n (R_{i,j}^{LBL} - R_{i,j}^{PC})^2}{m n} \right]^{1/2} \quad (9)$$

where  $R_{i,j}^{LBL}$  are the LBL radiances and  $R_{i,j}^{PC}$  are the PC\_RTTOV radiances. To minimize the contribution due to the radiance reconstruction error, we have computed the PC\_RTTOV radiances using 400 principal components.

We have evaluated the quantity in Eq. (9) for various numbers of predictor variables and threshold values. In particular, we have selected 300, 400, 500, and, 600 predictor variables for IASI and 200, 300, and, 400 predictor variables for AIRS and have varied the value of the threshold between 0.9972 and 0.9995 in increments of 0.001. To fix the maximum number of predictor variables for IASI, we have used Eq. (7) and found the value that results in at least a ten-fold gain in computational efficiency using 200 principal components for the radiance reconstruction. For AIRS, we have sought at least a five-fold gain using 100 principal components. Results are plotted in Figure 6, which shows

how the fitting error varies with the value of the threshold and with the number of predictor variables. It is evident that the optimal value of the threshold (i.e. the absolute minimum of the fitting error curve) varies with the number of predictor variables in the regression. The optimal threshold values used in this paper are tabulated in Table 2. Also noticeable is the fact that the fitting error decreases with the number of predictors variables in the regression. This aspect will be discussed in more detail in the next section. Although not shown here, the clusters with the largest population are generally found in the more transparent regions of the spectrum whereas the more opaque regions see the presence of clusters with a smaller number of channels. In many instances, the clusters contains only one channel.

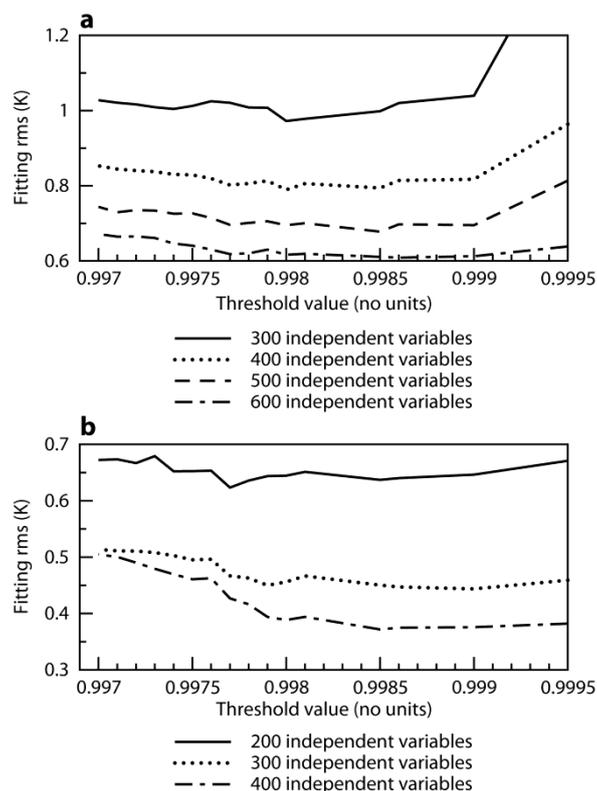


Figure 6: The fitting rms error for (a) IASI and (b) AIRS.

Table 2. The optimal value of the correlation threshold used in the selection of the independent variables.

Number of independent variables	200	300	400	500	600
Value of the threshold for IASI		0.998	0.998	0.9985	0.9986
Value of the threshold for AIRS	0.9977	0.999	0.9985		

## 5 Performance of PC\_RTTOV for simulations of IASI and AIRS radiances

In section 3.3 we have discussed the radiance reconstruction error by varying the number of PCs used to reconstruct the radiances. The spectra shown in Figure 4 were obtained using exact PC scores. Now that we have determined an algorithm to predict the PCs scores, we can assess the accuracy of PC\_RTTOV by a comparison of the radiances computed by the fast model with reference radiances obtained in different ways. Firstly, the PC\_RTTOV radiances computed for the dependent set of profiles that were used to train the fast model can be compared with the LBL model equivalents to determine the accuracy of the fast model itself. In the same context, we shall also use a set of profiles independent of the regression coefficients to allow uncertainties from different type of profiles to be included. Thirdly, the fast model radiances can be compared with measured radiances to allow spectroscopic and other uncertainties to be included.

### 5.1 The fit of PC\_RTTOV to the LBL radiances

The PC\_RTTOV radiances can be computed using a variable number of predictors and PCs. Since the number of predictors and the number of PCs can be combined in many different ways, it is impractical to present results on a single channel basis for all the possible combinations. Consequently, we will first concentrate on a discussion of the results in terms of the overall rms fitting error computed using Eq. (9) and then present results on a single channels basis only for a selected number of cases. The rms fitting error for IASI and AIRS is plotted in Figure 7 for the dependent and independent profiles as a function of the number of PCs and the number of predictor variables. Results for the independent set give a more realistic estimate of the error in that they tell us how good we are in replicating the findings for the training sample in an independent sample from the same population. All the panels in Figure 7 show that an increase in the number of predictors results in a

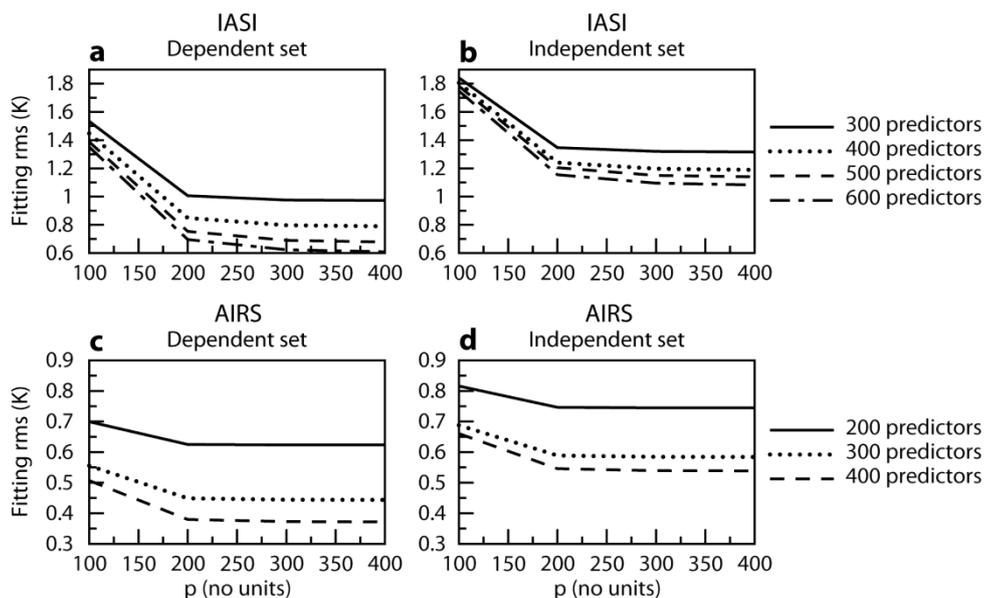


Figure 7. The fitting rms error as a function of the number of predictors and the number of PCs for IASI and AIRS.

reduction of the rms fitting error. Although not shown here, the reduction of the error is most noticeable in the CO<sub>2</sub> temperature sounding bands and in the water vapour band. The error spread is less accentuated for the independent case suggesting that errors related to the representatives of the training set are more important than errors determined by the choice of the number of predictors. This is also reflected in the fact that when moving to the independent dataset, the errors obtained using more predictor variables tend to increase more than the errors obtained using less predictor variables. As shown in Figure 7, the use of 300 or, indeed, 400 principal components, give little advantage over the use of 200 principal components. This is especially true for AIRS for which the error exhibits a less steep dependence on the number of PCs used in the radiance reconstruction.

Results for the independent set on a single channels basis (i.e. in Eq. (9) we do not perform the summation over the number of channels,  $i$ ) are presented in Figures 8 and 9 for IASI and AIRS respectively. For each instrument we have selected two cases, i.e. the combination of the number of predictor variables and PCs that corresponds to the minimum and maximum absolute value of the computational gain,  $g$ . Results are presented in terms of bias, rms error, and, maximum error. As discussed above, results for IASI are more influenced by the number of principal components. As Figure 4 shows, this is clearly the case for the channels in the CO<sub>2</sub>  $\nu_3$ -band (the region around 2350 cm<sup>-1</sup>) where errors depend markedly on the number of principal components. In contrast, the larger error in the centre of the fundamental CO<sub>2</sub>  $\nu_2$ -band at 667 cm<sup>-1</sup> appears to be a genuine feature of the regression algorithm. The rms error depends on the choice of parameters. For the vast majority of the IASI and AIRS channels it does not exceed 0.05 K. Outliers are the IASI channels in the CO<sub>2</sub>  $\nu_3$ -band and a very few AIRS channels in the ozone band. For both instruments the bias makes a marginal contribution to the rms error. It should be stressed that for any choice of parameters, the rms error is still below the noise of the two instruments. We should note that although in this study we have limited the minimum number of PCs to 100 and used no less than 300 IASI and 200 AIRS predictor variables, we can still reduce these numbers for any specific application that requires an even higher computational gain. Of course we must trade off any choice versus the loss in radiance accuracy.

Results in Figures 8 and 9 can be compared to equivalent results obtained using the conventional RTTOV model. These are shown in Figure 10. It is clear that the improvement in radiance accuracy that can be achieved using PC\_RTTOV is dramatic. This is true above all for the channels characterized by the presence of absorption lines due to ozone and water vapour. Also noticeable are the larger RTTOV errors in the CO<sub>2</sub>  $\nu_3$  band. Although the conventional RTTOV model does not compare favourably to PC\_RTTOV, we should note that the rms errors observed for the large sample of 5191 observations used in this paper are indeed very close to the rms errors obtained for the smaller sample of 258 independent observations used by Matricardi (2008). This result attests the robustness of the fast algorithm used to predict the RTTOV transmittances.

So far we have discussed the statistics of the error using predictor variables computed using the same 101 level pressure grid used in the LBL computations. The RTTOV model uses an internal interpolation scheme that allows the levels in the input profile to be defined by the user. The user defined profile is interpolated to the levels specified in the coefficient file (101 levels in our case) to perform the computations of the optical depths. To ensure that the Jacobians are properly mapped back on the user levels, the integration of the RT equation is performed on the same levels specified by the user by interpolating the optical depths into these levels. The predictor variables computed in

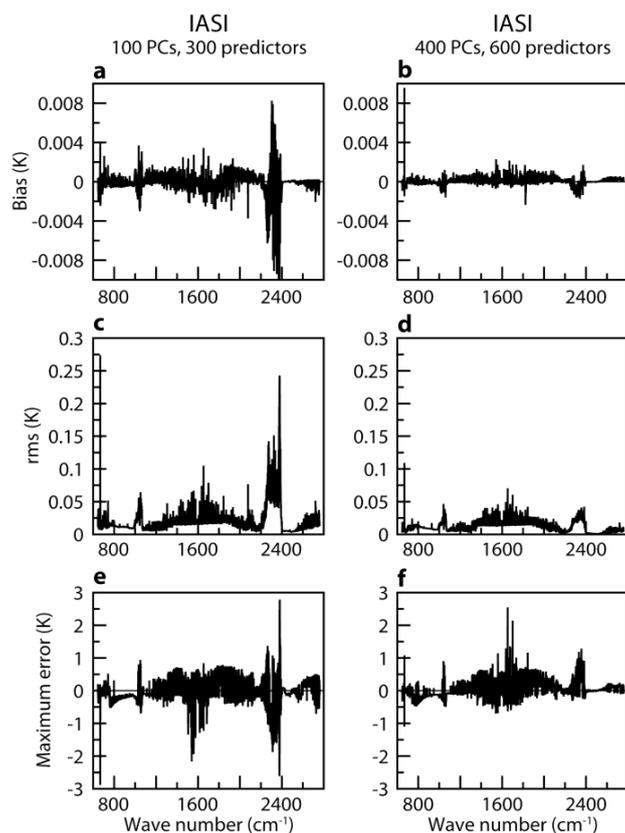


Figure 8. The mean value (bias), root mean square and maximum value of the difference between PC\_RTTOV and LBL radiances for the IASI instrument.

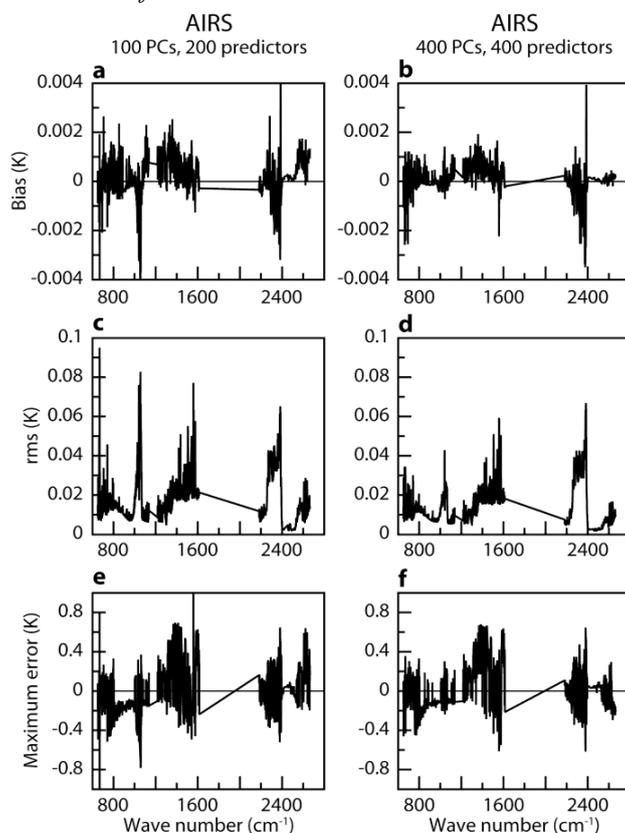


Figure 9. The mean value (bias), root mean square and maximum value of the difference between PC\_RTTOV and LBL radiances for the AIRS instrument.

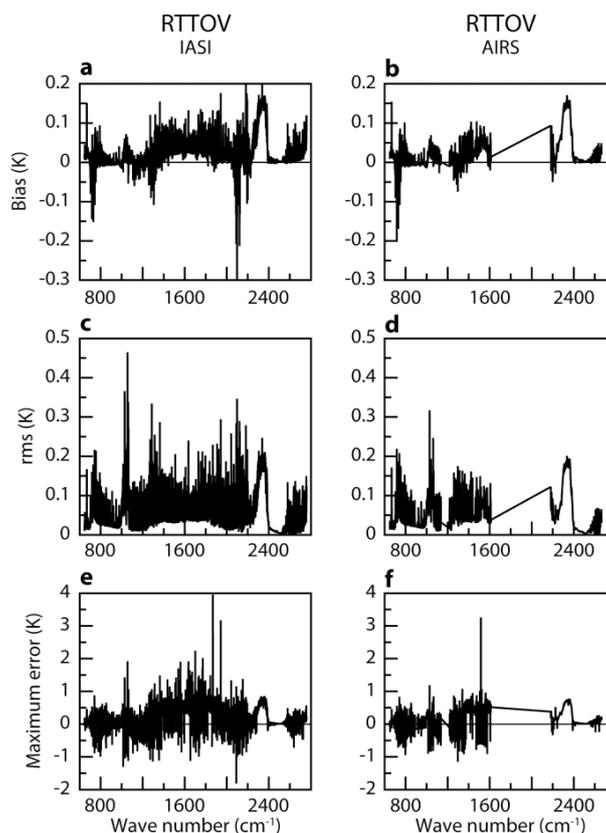


Figure 10. The mean value (bias), root mean square and maximum value of the difference between RTTOV and LBL radiances for the IASI instrument.

this way are inconsistent with those utilized during the training process and their use will produce errors in the simulated radiances. To estimate these errors, we have utilized the profiles in the independent set to compute IASI predictor variables obtained by interpolating the RTTOV optical depths to the 43 level pressure grid used in many RTTOV coefficients and to the 90 level pressure grid used in the RTIASI fast RT model (Matricardi, 2003). The rationale behind the choice of the 90 RTIASI levels is that in the current ECMWF assimilation system the RTTOV internal interpolation is carried out on 91 vertical levels. The radiances simulated using the new specification of the predictor variables have been compared to the original LBL radiances. Results (not shown here) tell us that by using the error-prone predictor variables, the radiances simulated by PC\_RTTOV have errors whose magnitude is comparable (in rms terms) to the magnitude of the errors introduced in the computation of the predictor variables (i.e. the error due to the change in the resolution of the vertical pressure grid used in the RT computations). For the examples presented in this paper, these errors are of the order of a few mK for the 90 level pressure grid whereas for the 43 level grid they vary between 0.1 K and 0.2 K.

## 5.2 Comparison with observed radiances

In this section we compare IASI spectra measured over sea with simulations carried out using PC\_RTTOV. The simulations are performed using fields of temperature, water vapour and ozone obtained from the ECMWF short-range forecasts during the period 1 April-10 April 2008. Only channels detected as clear by the ECMWF cloud detection algorithm are processed. To avoid reflected

solar radiation and non-LTE effects, in the spectral region between 2000 and 2760  $\text{cm}^{-1}$  we have sampled only night-time spectra. Since the ECMWF cloud detection algorithm (McNally and Watts 2003) finds clear channels rather than clear locations, the size of the sample varies with the sensitivity of the channel to clouds and amounts to several thousands of spectra for the channels peaking at middle and high altitudes to a few thousand for the channels peaking at low latitudes or at the surface.

By comparing clear sky observations with simulations we allow LBL spectroscopic errors and uncertainties in the specification of the input state vector to be included in the error statistics alongside with errors due to residual cloud contamination. This aspect of the problem has been discussed in detail in Matricardi (2009) and Shepard et al. (2009) with particular emphasis on the LBLRTM LBL model and any further examination is outside the scope of this paper. The results presented here serve as a sanity check to demonstrate that PC\_RTTOV does not give unexpected results. To this end, in Figure 11 we have plotted the mean value of the difference (bias) between the observed spectra and the spectra simulated by PC\_RTTOV (panel (a)) and RTTOV (panel (b)). Figure 11 clearly shows that the PC\_RTTOV biases are very similar to those obtained using the conventional (and well established) version of RTTOV used in the ECMWF operations.

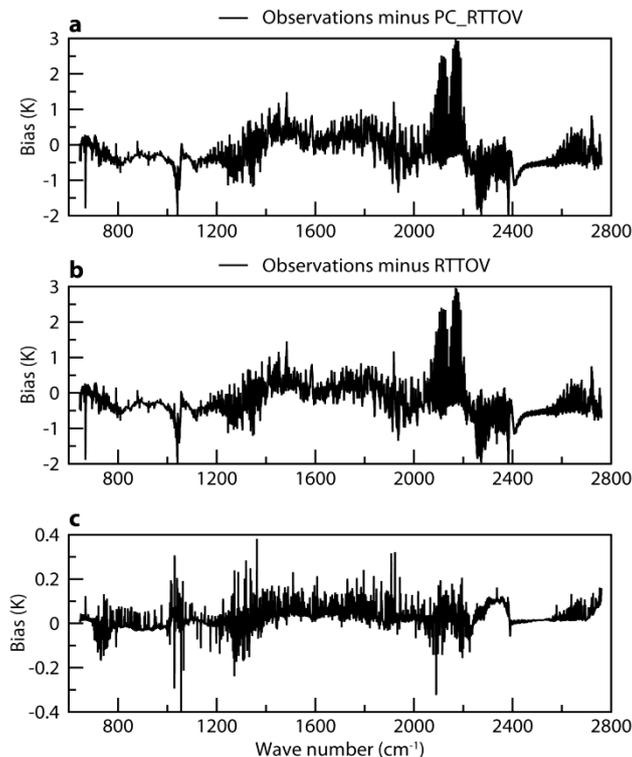


Figure 11. The mean value (bias) of the difference between observed radiances and radiances computed using a) PC\_RTTOV, b) RTTOV. In panel c) we plot the difference between the quantities plotted in panels a) and b).

It is generally assumed that fast model errors are dominated by the contribution given by spectroscopic errors in the underlying LBL models. If we assume that results shown in Figure 8 are representative of the errors associated with the PC\_RTTOV fast parameterization of the PC scores, then the curve plotted in panel (c) of Figure 11 (i.e. the difference between the curve in panel (a) and the curve in panel (b)) should give a reasonable estimate of the contribution given by the RTTOV fast transmittance parameterization scheme to the total error. It is interesting to see that results shown in Figure 11, panel (c), are in line with what we would expect based on the errors plotted in Figure 10. In

fact the channels that exhibit the largest differences are exactly the same channels that exhibit the largest rms errors in Figure 10.

## 6 The correlation matrix of the fast forward model error for IASI

The assimilation of satellite radiance data into a NWP system involves the definition of the observation error covariance matrix that is used to specify errors associated with radiance data. If we assume that the instrument noise is independent of the forward model errors, the observation error covariance matrix,  $\mathbf{O}$ , is the sum of the instrumental error covariance matrix,  $\mathbf{E}$ , and the forward model error covariance matrix,  $\mathbf{F}$ , which is based on the estimation of errors associated with fast RT models. The IASI radiances used in NWP assimilation are Level 1C apodised radiances. The effect of the apodisation is to introduce short range correlations in the instrumental noise linking each channel with its four nearest neighbours (Amato et al. 1999). Thus, long range correlations can only be introduced through the contribution given by the forward model error correlations which are usually neglected. In PC\_RTTOV the radiance reconstruction involves the use of variables (the PCs) that are linear combinations of channel radiances. Consequently we can expect substantial inter-channel error correlations. The significance of these correlations depends on the relative magnitude of forward model and instrumental errors. If we consider only the errors due to the fast parameterization of the PC scores, results shown in section 5.1 suggest that for PC\_RTTOV the instrumental noise is by far the dominant term in the observation error covariance matrix. It would seem then that a diagonal matrix gives an appropriate representation of the forward model error covariance matrix. However, if we consider the overall forward model error, a diagonal approximation could not be appropriate. For instance, in some spectral regions the overall forward model error can be dominated by the contribution due to spectroscopic errors in the underlying LBL model. Although the systematic component of the error is removed by the bias correction scheme used in the assimilation system, any random component might result in additional correlated errors. The validity of a diagonal approximation will again depend on the relative magnitude of forward model and instrumental errors and consequently some treatment of the error correlation could be required.

An in-depth discussion of the observation error matrix is outside the scope of this paper (the reader can refer to Bormann et al. (2010) for more details). Nevertheless, we consider it is important to document the fast model error correlation matrix since this aspect is too often overlooked in the literature. To this end, in Figure 12 we present the error correlation matrix for PC\_RTTOV obtained by computing the brightness temperature differences between fast and LBL model for the testing set of independent profiles. The details of the correlation matrix depend on the number of predictor variables used in the regression and on the number of PCs used in the radiance reconstruction. Results in Figure 12 refer to the 600 predictor-400 PC case since it represents one of the most complex cases. The colour scale varies from white (very high correlation) to black (very high anticorrelation). We also show (Figure 13) the equivalent correlation matrix for RTTOV although we will discuss in detail only the features of the PC\_RTTOV correlation matrix.

If we start our examination of the correlation matrix from low wave numbers, we see that there are high correlations between the channels in the CO<sub>2</sub> Q-branches in the 15  $\mu\text{m}$  absorption band. A moderately high anticorrelation also exists between these channels and the channels in the window region around 2400  $\text{cm}^{-1}$ . Channels in the window region between 800  $\text{cm}^{-1}$  and 1200  $\text{cm}^{-1}$  are also highly correlated and a high level of correlation exists between these channels and the channels in the

window regions between  $2000\text{ cm}^{-1}$ - $2150\text{ cm}^{-1}$  and between  $2400\text{ cm}^{-1}$ - $2760\text{ cm}^{-1}$ . The star-like feature around  $1050\text{ cm}^{-1}$  represents a strong correlation structure within the  $\text{O}_3$  band. Channels in the  $\text{O}_3$  band exhibits a moderate anticorrelation with channels in the  $\text{H}_2\text{O}$  band between  $1400$  and  $1700\text{ cm}^{-1}$  and a high anticorrelation with channels in the window region between  $2400\text{ cm}^{-1}$  and  $2760\text{ cm}^{-1}$ . The cross-like and the star-like structure centred in the middle of the  $\text{H}_2\text{O}$   $\nu_2$  band at  $1600\text{ cm}^{-1}$  represents a very significant correlation structure within the  $\text{H}_2\text{O}$  band. Regarding the  $\text{H}_2\text{O}$  band, we can also observe a moderate anticorrelation structure between the channels in the  $1500$ - $1700\text{ cm}^{-1}$  region and the window channels between  $800\text{ cm}^{-1}$  and  $1200\text{ cm}^{-1}$ . Very high correlations structures appear to be present within the window region between  $2000\text{ cm}^{-1}$  and  $2200\text{ cm}^{-1}$ , the  $\text{CO}_2$   $\nu_3$ -band around  $2300\text{ cm}^{-1}$  and the window region between  $2400\text{ cm}^{-1}$  and  $2760\text{ cm}^{-1}$ . Finally, various moderately high anticorrelation structures are present between  $2200\text{ cm}^{-1}$  and  $2760\text{ cm}^{-1}$ . If we compare the PC\_RTTOV with the RTTOV correlation matrix we note that in PC\_RTTOV the average level of correlation has increased (i.e. regions that in RTTOV appear in red/orange-yellow now appear in deep red/blue-white). Also noticeable is the much bolder correlation structure in the  $\text{H}_2\text{O}$  band and the presence of high correlations between channels in different window regions.

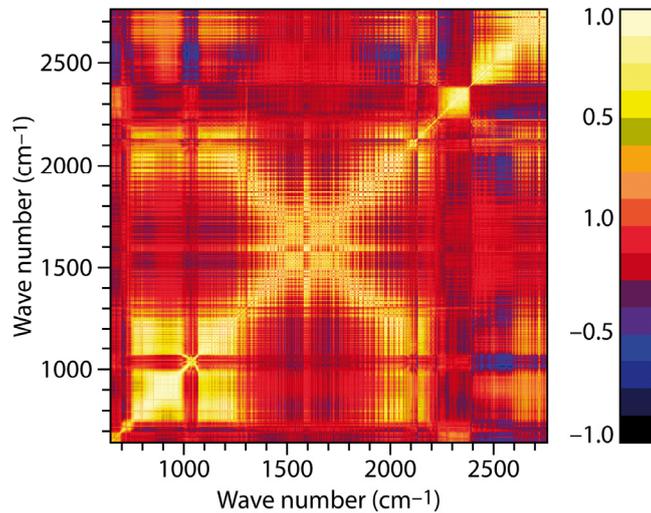


Figure 12. The forward model error correlation matrix for PC\_RTTOV.

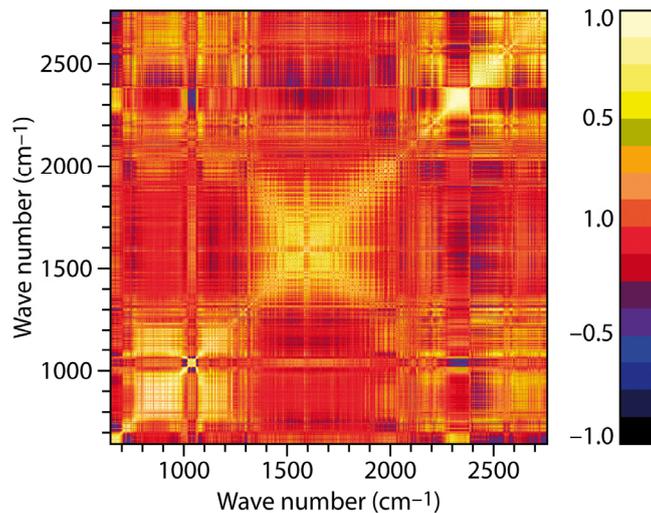


Figure 13. The forward model error correlation matrix for RTTOV.

## Conclusions

In this paper we have described the development of a new version of the RTTOV fast radiative transfer that exploits PC analysis (PC\_RTTOV). The model is based on the computation of a database of line-by-line spectra for a large training set of diverse atmospheric situations. The PC scores obtained from the eigenvectors of the covariance matrix of the simulated radiances are used as input data in a multiple linear regression scheme where they are expressed as a linear combination of profile dependent predictors which consist of a selected number of polychromatic radiances computed using the standard RTTOV fast transmittance model. The PC scores can be predicted for any state vector that includes variable profiles of temperature, water vapour, ozone, and, surface parameters.

The number of predictor variables used in the regression algorithm and the number of PCs used for the radiance reconstruction are tunable parameters in the model. By varying these parameters, we can change the computational efficiency and the accuracy of the model. The use of the default parameters specified in this paper allows the simulation of an IASI spectrum typically 8-21 times faster than RTTOV. Further improvements can be achieved by using a custom specification of the tunable parameters trading off accuracy versus computationally efficiency.

In terms of radiance computation, PC\_RTTOV can reproduce a large dataset of independent LBL radiances typically within 0.05K rms. For many channels, the actual value of the error is well below this figure. This is a dramatic improvement over the conventional RTTOV model. This is mainly due to the inherent linear relationship between the dependent variables and the independent predictor variables used in the PCs regression algorithm. The comparison of IASI radiances simulated with PC\_RTTOV with a large dataset of radiances measured by IASI gives results that are very similar to those obtained using the conventional (and well established) version of RTTOV used in the ECMWF operations. This is a further confirmation of the robustness of the PCs scores prediction algorithm.

A study of the PC\_RTTOV error correlation matrix shows that there are a number of long range correlation structures that are not present in the correlation matrix of the conventional RTTOV. The significance of these correlations depends on the relative magnitude of the forward model and instrumental errors. If we consider only the errors due the fast parameterization of the PC scores, the results presented in this study suggest that a diagonal matrix gives an appropriate representation of the forward model error covariance matrix. However, the validity of a diagonal approximation to the error covariance will depend on the magnitude of the spectroscopic errors introduced by the underlying LBL model and consequently some treatment of the error correlation could be required.

Finally, the availability of a principal component based fast radiative transfer will enable ECMWF to exploit the noise reduction capability of PCA and investigate the direct assimilation of IASI PC scores in the spectral regions affected by high instrument noise.

## Acknowledgments

This work has been funded by EUMETSAT, contract No. EUM/CO/06/1504/PS. We want to thank T. McNally, A. Collard, and, P. Bauer, all at ECMWF, for the useful discussions we had during the course of the work and for the suggestions to improve the quality of the paper.

## References

- Amato U, de Canditis D, Serio C. 1999. Effect of apodisation on the retrieval of geophysical parameters from Fourier transform spectrometers. *Appl. Optics*, **37**: 6537, 6543.
- Antonelli P, Revercomb HE, Sromovsky LA, Smith WL, Knuteson RO, Tobin DC, Garcia RK, Howell HB, Huang H-L, Best FA. 2004. A principal component noise filter for high spectral resolution infrared measurements. *Journal of Geophysical Research*, **109**: D23102.
- Aumann HH, Chahine MT, Gautier C, Goldberg MD, Kalnay E, McMillin LM, Revercomb H, Rosenkranz PW, Smith WL, Staelin DH, Strow LL, Susskind J. 2003 : AIRS/AMSU/HSB on the Aqua mission: design, science objectives, data products, and processing systems. *Geoscience and Remote Sensing, IEEE T.*, **41**, 253–264.
- Bormann N, Collard A, Bauer P. 2010: Estimates of spatial and inter-channel observation error characteristics for current sounder radiances for NWP, part II: Application to AIRS and IASI data. *Q.J.R.Meteorol.Soc.*, **136**, in press.
- Chalon G, Cayla F, Diebel D. 2001: 'IASI: An Advanced Sounder for Operational Meteorology'. Proceedings of the 52th Congress of IAF, Toulouse, France.
- Chevallier F. 2003. 'Sampled database of 60-level atmospheric profiles from the ECMWF analyses'. NWP SAF Report No. NWPSAF-EC-TR-001.
- Chevallier F, Di Michele S, McNally AP. 2006. 'Diverse profile datasets from the ECMWF 91-level short-range forecasts'. NWP SAF Report No. NWPSAF-EC-TR-010.
- Clough SA, Shepard MW, Mlawer EJ, Delamere JS, Iacono MJ, Cady-Pereira KE, Boukabara S, Brown PD. 2005. Atmospheric radiative transfer modelling: a summary of the AER codes. *J. Quant. Spectrosc. Radiat. Transfer*, **91**: 233–244.
- Collard AD, McNally AP. 2009. Assimilation of IASI radiances at ECMWF, *Q. J. Roy. Meteorol. Soc.*, **135**: 1044-1058.
- Hauglustaine DA, Brasseur GP, Walters S, Rasch PJ, Muller J-F, Emmons LK, Carrol MA. 1998. MOZART, a global chemical transport model for ozone and related chemical tracers, 2, Model results and evaluation. *J. Geophys. Res.*, **103**: 28,291-28,335.
- Havemann S. 2006. 'The development of a fast radiative transfer model based on an empirical orthogonal functions (EOD) technique'. Proceedings of the SPIE, Volume **6405**.
- Jacquinet-Husson N, Scott NA, Chedin A, Garceran K, Armante R, Chursin AA, Barbe A, Birk M, Brown LR, Camy-Peyret C, Claveau C, Clerbaux C, Coheur PF, Dana V, Daumont L, Debacker-Barilly MR, Flaud JM, Goldman A, Hamdouni A, Hess M, Jacquemart D, Kopke P, Mandin J-Y, Massie S, Mikhailenko S, Nemtchinov V, Nikitin A, Newnham D, Perrin A, Perevalov VI, Regalia-Jarlot L, Rublev A, Schreier F, Schult I, Smith KM, Tashkun SA, Teffo JL, Toth RA, Tyuterev VIG, Vander Auwera J, Varanasi P, Wagner G. 2005. The 2003 edition of the GEISA/IASI spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer*, **95**: 429-467.

Liu X, Smith WL, Zhou DK, Larar A. 2006. Principal component-based radiative transfer model for hyperspectral sensors: theoretical concept. *Appl. Opt.* **45**: 201-209.

Masuda K, Takashima T, Takayama T. 1988. Emissivity of pure sea waters for the model sea surface in the infrared window regions. *Remote Sens. Environ.*, **24**: 313-329.

Matricardi M. 2003. 'RTIASI-4, a new version of the ECMWF fast radiative transfer model for the infrared atmospheric sounding interferometer'. ECMWF Technical Memo No. 425. European Centre for Medium-Range Weather Forecasts: Reading, UK. Available at <http://www.ecmwf.int/publications/>.

Matricardi M, Chevallier F, Kelly G, Thepaut J-N. 2004. An improved general fast radiative transfer model for the assimilation of radiance observations. *Q. J. R. Meteorol. Soc.*, **130**: 153-173.

Matricardi M. 2007. 'An inter-comparison of line-by-line radiative transfer models'. ECMWF Technical Memo No. 525. European Centre for Medium-Range Weather Forecasts: Reading, UK. Available at <http://www.ecmwf.int/publications/>.

Matricardi M. 2008. 'The generation of RTTOV regression coefficients for IASI and AIRS using a new profile training set and a new line-by-line database'. ECMWF Technical Memo No. 564. European Centre for Medium-Range Weather Forecasts: Reading, UK. Available at <http://www.ecmwf.int/publications/>.

Matricardi M. 2009. Technical Note: An assessment of the accuracy of the RTTOV fast radiative transfer model using IASI data. *Atmos. Chem. Phys.*, **9**: 6899-6913.

McMillin LM, Goldberg MD. 1997. 'The use of super channels for high resolution soundings'. Proceedings of the ninth international TOVS conference, Igls, Austria, 20-26 February 1997.

McNally AP, Watts PD. 2003. A cloud detection algorithm for high-spectral-resolution infrared sounders. *Q. J. Roy. Meteorol. Soc.*, **129**: 3144-3423.

McNally AP, Watts PD, Smith JA, Engelen R, Kelly GA, Thepaut J-N, Matricardi M. 2006. The assimilation of AIRS radiance data at ECMWF. *Q. J. Roy. Meteorol. Soc.*, **132**: 935-957.

Pinkley LW, Williams D. 1976. Optical properties of sea water in the infrared. *J. Opt. Soc. Am.*, **66**: 554-558.

Rothman LS, Barbe A, Chris Benner D, Brown LR, Camy-Peyret C, Carleer MR, Chance K, Clerbaux C, Dana V, Devi VM, Fayt A, Flaud J-M, Gamache RR, Goldman A, Jacquemart D, Jucks KW, Lafferty WJ, Mandin J-Y, Massie ST, Nemtchinov V, Newnham DA, Perrin A, Rinsland CP, Schroeder J, Smith KM, Smith MAH, Tang K, Toth RA, Vander Auwera J, Varanasi P, Yoshino K. 2003. The HITRAN molecular spectroscopic database: edition of 2000 including updates through 2001. *J. Quant. Spectrosc. Radiat. Transfer*, **82**: 5-44.

Rothman LS, Jacquemart D, Barbe A, Chris Benner D, Birk M, Brown LR, Carleer MR, Chackerian Jr. C, Chance K, Coudert LH, Dana V, Devi VM, Flaud J-M, Gamache RR, Goldman A, Hartmann J-M, Jucks KW, Maki AG, Mandin J-Y, Massie ST, Orphal J, Perrin A, Rinsland CP, Smith MAH, Tennyson J, Tolchenov RN, Toth RA, Vander Auwera J, Varanasi P, Wagner G. 2005. The HITRAN 2004 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer*, **96**: 139-204.

Saunders R, Matricardi M, Brunel P. 1999. An improved fast radiative transfer model for assimilation of satellite radiance observations. *Quart. J. Roy. Meteor. Soc.* **144**: 1547-1558.

Shepard MW, Clough SA, Payne VH, Smith WL, Kireev S, Cady-Pereira KE. 2009. Performance of the line-by-line radiative transfer model (LBLRTM) for temperature and species retrievals: IASI case studies from JAIVEx. *Atmos. Chem. Phys.* **9**: 9313-9366.

Smith LW, Wu X. 1997. Emissivity of rough sea surface for 8-13 um: modelling and verification. *Appl. Op.* **36**: 2609-2619.

Strow LL, Hannon SE, De Souza-Machado S, Motteler SH, Tobin D. 2003: An Overview of the AIRS Radiative Transfer Model, *IEEE T. Geosci. Remote Sens.*, **41**, 303-313.

Watts PW, Allen MR, Nightingale TJ. 1996. Wind speed effects on sea surface emission and reflection for the along track scanning radiometer. *J. Atmos. Ocean. Technol.* **13**: 126-141.