

Potential for the use of reconstructed IASI radiances in the detection of atmospheric trace gases

N. C. Atkinson¹, F. I. Hilton^{1,2}, S. M. Illingworth², J. R. Eyre¹, and T. Hultberg³

¹Met Office, Fitzroy Road, Exeter, EX1 3PB, UK

²Department of Physics and Astronomy, University of Leicester, LE1 7RH, UK

³EUMETSAT, Eumetsat-Allee 1, 64295 Darmstadt, Germany

Received: 8 February 2010 – Published in Atmos. Meas. Tech. Discuss.: 11 February 2010

Revised: 15 July 2010 – Accepted: 16 July 2010 – Published: 30 July 2010

Abstract. Principal component (PC) analysis has received considerable attention as a technique for the extraction of meteorological signals from hyperspectral infra-red sounders such as the Infrared Atmospheric Sounding Interferometer (IASI) and the Atmospheric Infrared Sounder (AIRS). In addition to achieving substantial bit-volume reductions for dissemination purposes, the technique can also be used to generate reconstructed radiances in which random instrument noise has been reduced. Studies on PC analysis of hyperspectral infrared sounder data have been undertaken in the context of numerical weather prediction, instrument monitoring and geophysical variable retrieval, as well as data compression. This study examines the potential of PC analysis for chemistry applications.

A major concern in the use of PC analysis for chemistry is that the spectral features associated with trace gases may not be well represented in the reconstructed spectra, either due to deficiencies in the training set or due to the limited number of PC scores used in the radiance reconstruction. In this paper we show examples of reconstructed IASI radiances for several trace gases: ammonia, sulphur dioxide, methane and carbon monoxide. It is shown that care must be taken in the selection of spectra for the initial training set: an iterative technique, in which outlier spectra are added to a base training set, gives the best results. For the four trace gases examined, key features of the chemical signatures are retained in the reconstructed radiances, whilst achieving a substantial reduction in instrument noise.

A new regional re-transmission service for IASI is scheduled to start in 2010, as part of the EUMETSAT Advanced Retransmission Service (EARS). For this

EARS-IASI service it is intended to include PC scores as part of the data stream. The paper describes the generation of the reference eigenvectors for this new service.

1 Introduction

The IASI instrument (Infrared Atmospheric Sounding Interferometer; Siméoni et al., 1997) on the MetOp-A satellite is a unique instrument for Earth observation (Phulpin et al., 2007). Since its launch in 2007, IASI has been used in a wide range of applications, including numerical weather prediction (NWP) (e.g. Hilton et al., 2009a), air quality monitoring and atmospheric chemistry (e.g. Clerbaux, 2009).

The instrument has the following characteristics:

- 8461 spectral channels, sampled every 0.25 cm^{-1} , covering the region $645\text{--}2760\text{ cm}^{-1}$
- 3 overlapping spectral bands (described in more detail in Sect. 2)
- Cross-track scan geometry
- 30 scan positions, separated by 3.3° ($\sim 50\text{ km}$ at nadir); 1 scan per 8 s
- 4 fields of view per scan position, each being sampled by a different detector
- Footprint 12 km at nadir for a single field of view

For NWP the focus is typically on those spectral channels that give information on atmospheric temperature and humidity profiles, in the carbon dioxide and water vapour absorption bands. To assimilate the radiances from all 8461 IASI channels would be prohibitively expensive computationally, and in any case the full information content of the



Correspondence to: N. C. Atkinson
(nigel.atkinson@metoffice.gov.uk)

spectrum is much lower than the number of channels, so a subset of a few hundred channels is often used (Collard, 2007). Timeliness is critical for NWP applications – with global data generally being made available by EUMETSAT within 2 h of the sensing time, and distributed to NWP users via the Global Telecommunications System (GTS) or via the EUMETCast re-transmission service. The GTS data stream comprises a subset of 300 IASI channels whereas the EUMETCast service currently carries all channels and it is up to the user to select those that are required for the application.

For chemistry applications, on the other hand, the requirements tend to be different. Full resolution spectra are needed covering specific spectral regions where trace gas absorption features are found. Timeliness is usually not critical for climatological studies: for example, Clarisse et al. (2009) have recently reported the first IASI-derived measurements of ammonia total column, and have mapped the concentrations over the course of 2008. However, spectra are increasingly processed in near-real time (e.g. George et al., 2009), and in fact near-real time atmospheric chemistry products are becoming more important in the context of the international efforts under the Global Monitoring for Environment and Security (GMES) initiative, in particular through its atmospheric service designed in the framework of the Monitoring Atmospheric Composition and Climate (MACC) project (<http://www.gmes-atmosphere.eu/>).

One of the tools being investigated by the NWP community in the context of IASI is the use of Principal Component (PC) analysis for the extraction of meteorological signals from the spectra and suppression of random noise (e.g. Antonelli et al., 2004). By representing a spectrum in terms of up to a few hundred PC scores (against a set of reference eigenvectors) instead of 8461 channels, a significant reduction in data volume can be achieved. Experiments are being undertaken both in the direct assimilation of PC scores (M. Matricardi, personal communication, 2009) and in the assimilation of radiances reconstructed from the PC scores (Hilton and Collard, 2009b).

PC compression can potentially lead to cost savings in near-real-time data transmission. Motivated by the desire for increased timeliness in IASI data distribution for regional NWP, EUMETSAT are planning a pilot IASI re-transmission service as part of the EUMETSAT Advanced Retransmission Service (EARS) program. For this service it is planned to distribute PC scores, together with raw radiances for selected channels. The same technique may in the future also be used for distribution of the global IASI data – especially after the launch of MetOp-B in 2012, as present communications channels would not be able to accommodate full-resolution data from two IASI instruments without some form of compression. EUMETSAT delegate bodies have recommended the PC approach for near-real-time data transmission, since a significant increase in bandwidth would represent a substantial cost to member states.

To date there have been no published studies in the use of IASIPC scores in chemistry applications. A major advantage of the PC technique is its potential to suppress noise, which may provide benefit to the retrieval of chemical species characterised by weak signals. However, there is a concern that the signatures of trace gases may not be retained in the reconstructed spectra. This could occur if the trace gas signal is very weak (significantly below instrument noise level) or if the trace gas signature is not captured in the training set of spectra from which the reference eigenvectors were derived. It is therefore important to understand the implications of PC compression of IASI data for atmospheric chemistry where the use of near-real time data is important.

In this paper we look for the characteristic signatures of some trace gases in IASI reconstructed radiances and compare with the signatures found in the raw radiances. The gases and spectral bands chosen are as follows:

- Ammonia (NH₃) lines in the 850–1000 cm⁻¹ region. Ammonia is a short-lived gas emitted primarily from livestock, agriculture and biomass burning (Beer et al., 2008; Clarisse et al., 2009). Ammonia is very variable spatially.
- Sulphur dioxide (SO₂) in the 1300–1400 cm⁻¹ region. A strong feature but only found in rare events such as volcanic eruptions (Clarisse et al., 2008a)
- Methane (CH₄) in the ν_3 band at 2550–2760 cm⁻¹. This band can be used to complement the ν_4 band at 1220–1380 cm⁻¹ which is usually used to retrieve CH₄ (Crevoisier et al., 2009b), as the ν_3 band can, under certain conditions, improve the information in the boundary layer (Razavi et al., 2009). This band is an example of a spectral region where IASI instrument noise is very high, and could benefit from PC noise reduction.
- Carbon Monoxide (CO) in the 2143–2181 cm⁻¹ region. CO is formed during biomass burning (Turquety et al., 2009), by incomplete combustion of fossil fuels and by oxidation of volatile organic compounds. CO has a longer lifetime than NH₃ of a few weeks to months and can be used as a transport tracer (e.g. George et al., 2009).

For each of these trace gases, the use of different sets of reference eigenvectors was examined. The three sets used initially are described in Sect. 2, whilst a fourth is introduced in Sect. 4.

For NH₃, SO₂ and CH₄ we do not attempt to perform any trace gas retrievals from reconstructed radiances, but by identifying the characteristics of the spectra we show the potential for use of the reconstructed radiances in further studies. In the case of CO, we test the effects of PC compression and radiance reconstruction on total column retrievals for an area of elevated CO concentration over central Africa.

2 PC methodology

For a radiance spectrum \mathbf{x}_i (column-vector with m channels at location i), the Principal Component scores, \mathbf{p}_i (column-vector rank r where $r \leq m$), are computed from a set of pre-computed eigenvectors, \mathbf{E} (rank $m \times r$), instrument noise covariance \mathbf{N} ($m \times m$) and a (reference) mean radiance $\bar{\mathbf{x}}$:

$$\mathbf{p}_i = \mathbf{E}^T \mathbf{N}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \quad (1)$$

The reconstructed radiances, \mathbf{x}'_i , are given by:

$$\mathbf{x}'_i = \mathbf{N} \mathbf{E} \mathbf{p}_i + \bar{\mathbf{x}} \quad (2)$$

To simplify the computation, the noise normalisation matrix is usually assumed to be diagonal (and is assumed diagonal in this paper).

The eigenvectors are generated via the covariance matrix, \mathbf{C} , of the noise-normalised training-set spectra, \mathbf{y}_i ($= \mathbf{N}^{-1} \mathbf{x}_i$), as described by Collard (2008):

$$\mathbf{C} = \frac{1}{n} \sum_{i=1}^n \mathbf{y}_i \mathbf{y}_i^T - \bar{\mathbf{y}} \bar{\mathbf{y}}^T = \mathbf{E} \mathbf{\Lambda} \mathbf{E}^T \quad (3)$$

where n is the number of spectra in the training set, \mathbf{E} is the matrix of eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix containing eigenvalues. $\bar{\mathbf{y}}$ denotes the mean of the n noise-normalised spectra. A software package called the ‘‘IASI PCA-based Compression package’’ is available to allow users to perform this computation, see <http://www.nwpsaf.org/>.

Equation (3) can in principle yield eigenvectors of rank $m \times m$ – in which case \mathbf{x}'_i is identical to \mathbf{x}_i . However most of the atmospheric signal is contained in a relatively small number of leading eigenvectors, with the higher-rank eigenvectors containing mainly instrument noise. Therefore the higher-rank eigenvectors are usually discarded.

Three sets of reference eigenvectors were used initially in this study:

1. A set of 150 eigenvectors generated by the Met Office and distributed with version 6.6 of the ATOVS and AVHRR Pre-processing Package (AAPP – Atkinson et al., 2008), derived from 6 months of thinned IASI data (15 736 de-apodized spectra) from July to December 2007. The noise normalisation was the RMS reconstruction error after an initial set of PCs had been generated using a CNES pre-launch level 1B noise profile (from the March 2001 meeting of the IASI Sounding Science Working Group). Referred to as set 1.
2. A ‘‘base’’ set eigenvectors generated by EUMETSAT from 74 719 apodized level 1C spectra chosen randomly (one per IASI scan line) for 7 different days (7 December 2007, 12 March 2008, 9 May 2008, 26 July 2008, 26 August 2008, 25 September 2008 and 7 January 2009). Each of the three IASI spectral bands is treated separately. There are 90 eigenvectors selected in band 1 (see below for band definitions), 120 in band

2 and 80 in band 3. The noise normalisation was the square root of the diagonal of the CNES level 1C covariance dated 28 April 2008, which is derived from internal calibration view measurements. Referred to as set 2.

3. As for set 2., but with 6664 ‘‘outliers’’ added to the training set. The outliers are taken from the period 8 to 12 August 2008, which includes the eruption of the Kasatochi volcano in the Aleutian Islands, off south western Alaska. Referred to as set 3.

Since single band failures are more frequent for band 3 than band 2 (which again are more frequent than for band 1), in this study we choose the band limits at the lower end of the band overlap regions. This avoids a failure of band 3 affecting any of the other bands. Thus band 1 comprises channels 1 to 1997 (645–1144 cm^{-1}), band 2 comprises channels 1998 to 5116 (1144.25–1923.75 cm^{-1}) and band 3 comprises channels 5117 to 8461 (1924–2760 cm^{-1}).

For sets 2 and 3, the optimal number of PCs was selected for each band by plotting the PC score spatial correlation as a function of eigenvector rank, as described in Atkinson et al. (2009). Low-rank eigenvectors show high spatial correlation as they mainly represent atmospheric structure, whereas high-rank eigenvectors show low spatial correlation as they are dominated by random noise. In the case of set 1 a simpler method was used: a plot of the eigenvalues was examined and PCs were selected up to the point where the slope of the curve stabilised.

In each case, a random selection of scanning angles was used, and data from all four detectors were included in the training set. Some experiments were performed in which separate training sets were used for each detector (Hultberg, 2010), but whilst some differences were observed between the detectors, the reconstruction scores and spatial correlations were not substantially affected. Therefore this issue is not discussed further.

Experiments were also performed to investigate whether any improvement could be made by including off-diagonal components of \mathbf{N} , when working with apodized spectra. This might be expected to reduce the amount of signal contained in the higher-rank PC scores. But in practice no clear improvement was found and therefore the decision was made to use a diagonal noise matrix, which is computationally simpler.

To assess the fit of the reconstructed radiances to the raw radiances, it is useful to compute Reconstruction Scores, S_j , for each spectrum, defined as

$$S_j = \sqrt{\left(\mathbf{y}'_j - \mathbf{y}_j\right)^T \left(\mathbf{y}'_j - \mathbf{y}_j\right) / m} \quad (4)$$

where \mathbf{y}'_j is the noise-normalised reconstructed radiance vector for IASI band j , with m channels. The reconstruction scores are typically computed for each of the three IASI spectral bands (corresponding to three physical detectors). Values

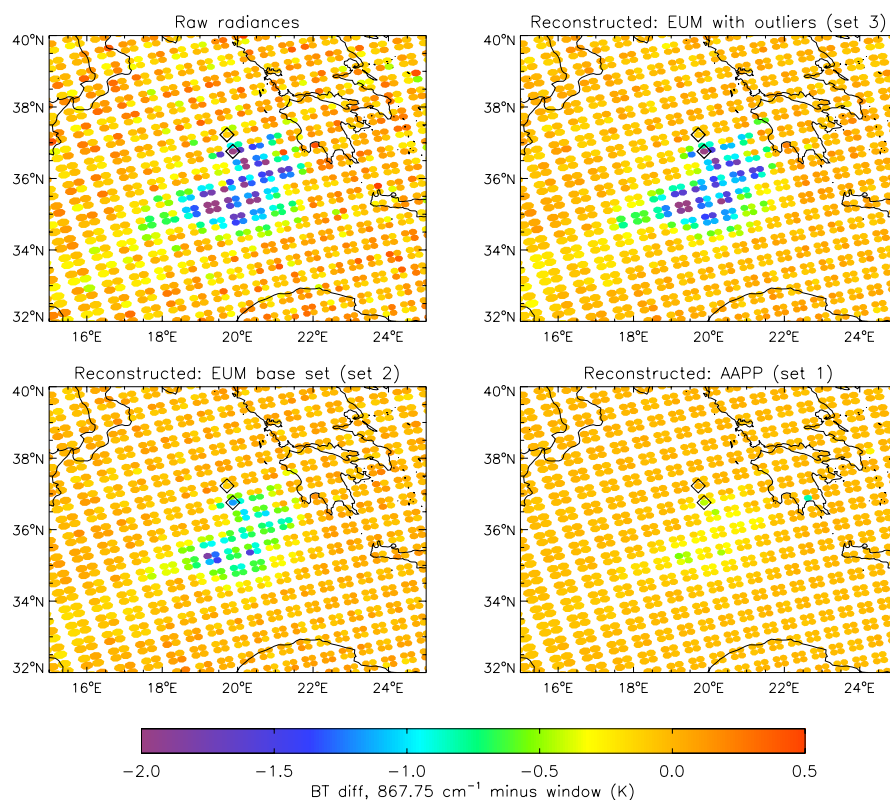


Fig. 1. Ammonia signal associated with biomass burning near Greece on 25 August 2007, 19:38 MetOp overpass. The brightness temperature at 867.75 cm^{-1} has been subtracted from the brightness temperature in adjacent window channels. Shown for raw radiances and three varieties of reconstructed radiances. The BT scale has been fixed at -2.0 to $+0.5\text{ K}$ in order to show background noise, but it should be noted that the strongest signal was -3.7 K . The size of the IASI spots has been increased from 12 km at nadir to 20 km in order to show colours clearly. The spots selected for display in Fig. 2 are marked by black diamonds.

of S_j are expected to be close to 1.0 if the reconstructed radiance accurately represents the raw radiance and if the noise profile \mathbf{N} is realistic (Goldberg et al., 2003). In qualitative terms, a value of S_j much greater than 1.0 implies that the reconstruction is not fitting the raw spectrum to the expected level, whilst a value consistently less than 1.0 implies that the assumed noise is greater than the actual random noise.

The noise covariance of the reconstructed radiances is substantially different from that of the raw radiances. Random (uncorrelated) noise is expected to be spread evenly across all PCs, and will therefore be reduced by virtue of the fact that only the low-rank PCs are retained. The remaining noise shows increased inter-channel correlations and may be calculated as $\mathbf{NEE}^T \mathbf{N}^{-1} \mathbf{RN}^{-1} \mathbf{EE}^T \mathbf{N}$ (Hultberg, 2009), where \mathbf{R} is the error covariance matrix of the IASI level 1C radiances. For further discussion of correlated noise, see Hilton and Collard (2009).

3 Case studies

3.1 Detection of ammonia

Ammonia has a series of spectral lines in the region 800 to 1200 cm^{-1} . Following Clarisse et al. (2009), we can obtain an indication of the ammonia abundance from the brightness temperature difference between the IASI channel at 867.75 cm^{-1} and two neighbouring window channels, at 861.25 and 873.5 cm^{-1} . We use the case study that is described in detail in Clarisse et al. (2008b) (also in Coheur et al., 2009), in which wildfires close to Greece were observed by IASI in the evening of 25 August 2007.

Figure 1 shows a map of the brightness temperature difference, obtained from raw radiances and from the three varieties of reconstructed radiances described in Sect. 2. The raw radiance map and the reconstructed radiance map obtained from eigenvectors set 3 match well. The reconstructed radiance signal shows evidence of a lower background noise level, as would be expected. However, the plume feature is significantly attenuated in the reconstructed radiance map obtained from set 2, while for set 1 the feature is almost

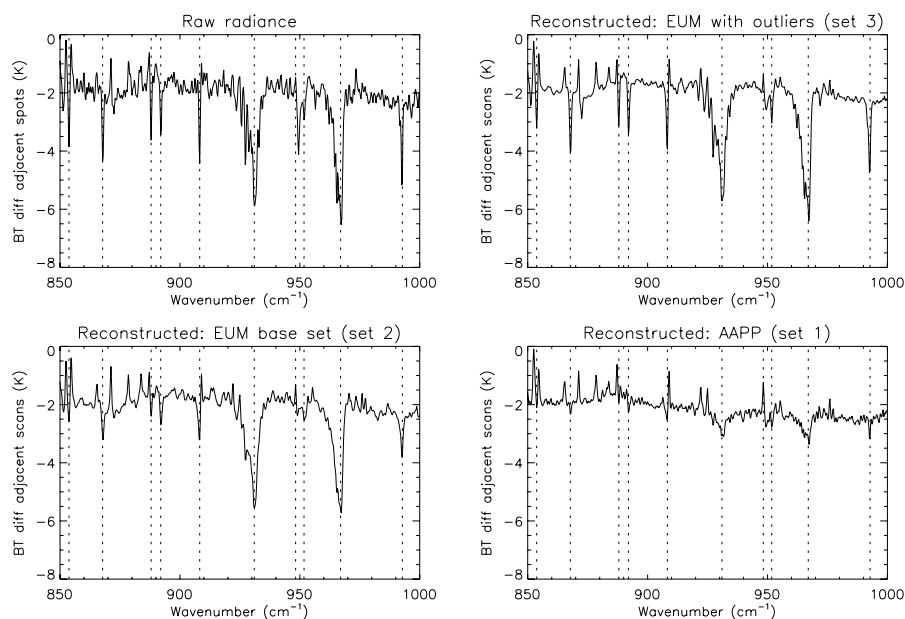


Fig. 2. Difference spectra between a spot on the northern edge of the plume and a spot on the following IASI scan line which is outside the plume. Vertical dotted lines mark the positions of ammonia absorption lines, from Clarisse et al. (2008b).

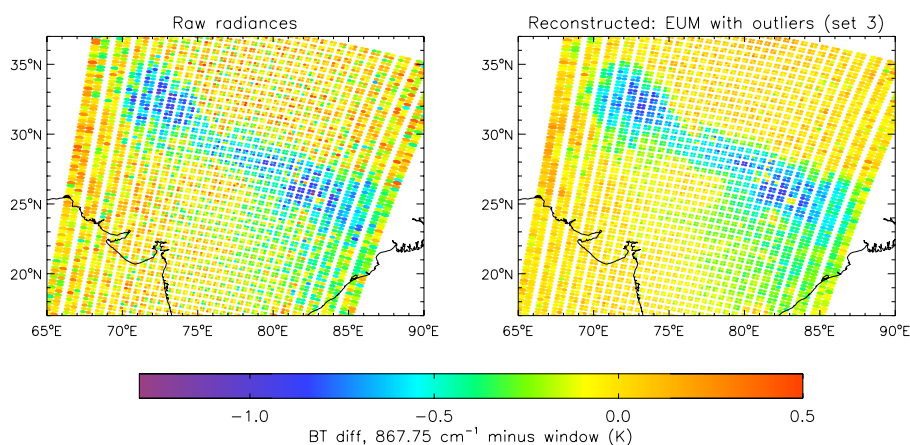


Fig. 3. Ammonia signal associated with agriculture over northern India. 13 May 2007, morning overpass.

absent. It must be concluded that the set 1 and set 2 training sets do not include sufficient representation of ammonia.

Figure 2 shows difference spectra between a spot inside the plume and a spot on the next IASI scan line which was outside the plume. The positions of the ammonia lines are marked with dashed vertical lines. The set 3 reconstructed radiances show close agreement with the raw radiances, but the set 2 and set 1 plots show reduced ammonia signal. Set 2 does show the strong lines at 931 and 967 cm^{-1} but even these are greatly attenuated in set 1. In order to ascertain whether the poor performance of set 1 was due to the use of only 150 PCs an experiment was carried out in which the number of PCs was increased to 290. The change made little

difference; therefore we conclude that the inability of set 1 to detect the ammonia signal is due primarily to a lack of significant ammonia episodes within the relatively small set 1 training set (15 736 spectra).

As a second example, Fig. 3 shows the ammonia signal over northern India on 13 May 2008. This region has been identified by Clarisse et al. (2009) as a persistent source of ammonia, especially during the months May to August. As in the case over Greece, it can be seen that the raw radiances and the reconstructed radiances give a very similar geographical distribution for the ammonia signal, but the reconstructed radiances (set 3) contain a lower level of random noise.

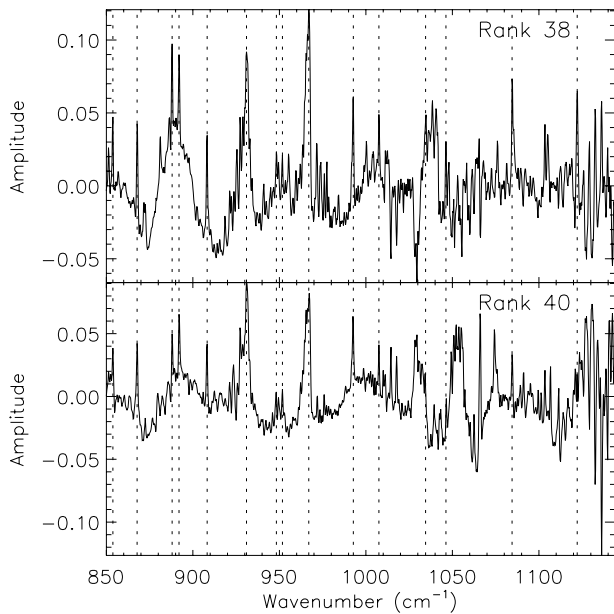


Fig. 4. Eigenvectors of rank 38 and 40 in IASI band 1. Both contain a strong signature of ammonia. Vertical lines mark the ammonia absorption lines, from Clarisse et al. (2008b).

In order to identify the reason why eigenvector set 3 performs better than set 2, the individual spectra in the “outliers” training set were examined for ammonia signals. It was found that there were a small number of spectra in eastern Kazakhstan with very high band 1 reconstruction scores (up to 3.2). This is one of the ammonia hot-spot regions reported by Clarisse et al. (2009). When analyzed using set 3, the peak reconstruction score dropped to 1.2, confirming that the inclusion of these spectra has added an ammonia signal to the eigenvectors (though the residual is still somewhat elevated compared with neighbouring spots). For reference, the peak reconstruction score for set 1 was 4.6.

Looking more closely at the eigenvectors themselves, the ammonia signal was found to be most obvious in the band 1 eigenvectors of rank 38 and 40, as shown in Fig. 4. Note that there are 90 eigenvectors in band 1, so the ammonia signal is by no means one of the weakest signals to be included – implying that on the rare occasions that the signal is present, it is well above instrument noise.

3.2 Detection of volcanic SO₂

The characteristics of SO₂ spectra are described in detail by Clarisse et al. (2008a). There are two bands that can be detected by IASI, the ν_1 band centred on 1150 cm⁻¹ and the ν_3 band centred on 1360 cm⁻¹. According to Clarisse et al. the ν_1 band is difficult to use because of an ice signature and emissivity features on arid land, therefore we concentrate on the ν_3 band. This is within IASI band 2. Following Clarisse et al., we identify major concentrations of SO₂ by measuring

the brightness temperature difference between two channels in the absorption band (1371.5 and 1371.75 cm⁻¹) and two “baseline” channels (1407.25 and 1408.75 cm⁻¹).

Figure 5 shows a map of the SO₂ signal on 8 August 2007 in the vicinity of the Kasatochi eruption, for raw radiances and the three varieties of reconstructed radiances. As in the ammonia cases, the reconstructed radiances from the set 3 eigenvectors give good agreement with the raw radiances. The set 2 eigenvectors significantly over-estimate the signal, whereas the set 1 eigenvectors significantly under-estimate the signal. Difference spectra (inside the plume minus outside) are shown in Fig. 6. It is not clear why set 2 over-estimates the signal, but it seems likely that there is a low-level (noisy) SO₂ signal in the training set that is not representative of the situation where there is a high concentration of the gas.

The band 2 reconstruction scores using the two sets of EU-METSAT eigenvectors are shown in Fig. 7. The reduction in reconstruction error for the set 3 eigenvectors is clear, though there is still some residual structure. The removal of this residual structure is discussed in Sect. 4.

Because some of the Kasatochi outliers were included in the training set for set 3, the analysis was repeated for an independent case – the eruption on Jebel at Tair (30 September 2007), described in Clarisse et al. (2008a). The results (not shown) were very similar to the Kasatochi case, with good agreement between the set 3 reconstructed spectra and the raw spectra. As before, the set 2 eigenvectors significantly over-estimated the signal and the set 1 eigenvectors under-estimated the signal.

3.3 Detection of methane in the short-wave band

As mentioned in Sect. 1, the methane short-wave ν_3 band can be used to improve methane retrievals in the boundary layer – complementing the long-wave ν_4 band. During the day the solar radiation reflected by the Earth’s surface increases the IASI signal substantially in the short-wave band. However, at night, or if the reflected solar radiation is low, the signal to noise ratio (SNR) can be very poor. Razavi et al. (2009) found that it was necessary to improve the SNR in order to perform satisfactory retrievals of CH₄, even during the day: to do this they averaged four contiguous IASI measurements (50×50 km).

Reconstructed radiances could be used to improve the SNR, and hence reduce the need to perform spatial averaging. This is illustrated in Fig. 8, which shows raw and reconstructed brightness temperature spectra for a night-time case, using the set 3 eigenvectors. The improved SNR for the reconstructed spectrum is striking, with the methane (and HDO) signatures clearly visible.

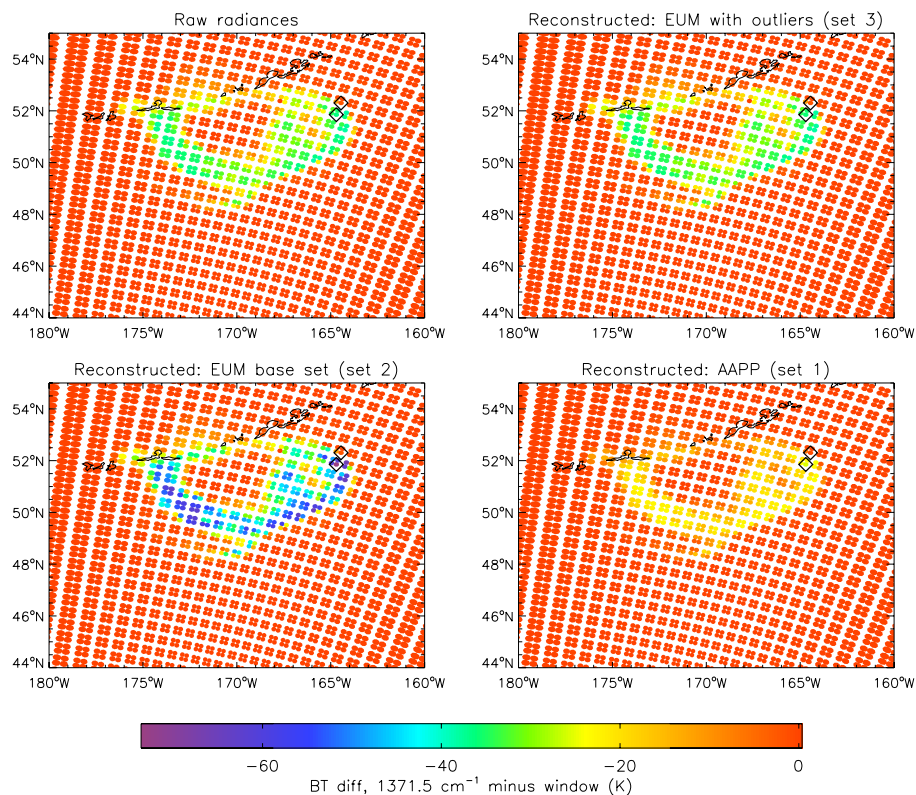


Fig. 5. SO₂ signal associated with the Kasatochi volcano eruption on 8 August 2008, 21:23 MetOp overpass. The brightness temperature at 1371.5 and 1371.75 cm⁻¹ has been subtracted from the brightness temperature in adjacent window channels. Note the anomalously large signal in the reconstructed radiances from the set 2 eigenvectors. The volcano is located at 52.18° N, 175.51° W. The spots selected for display in Fig. 6 are marked by black diamonds.

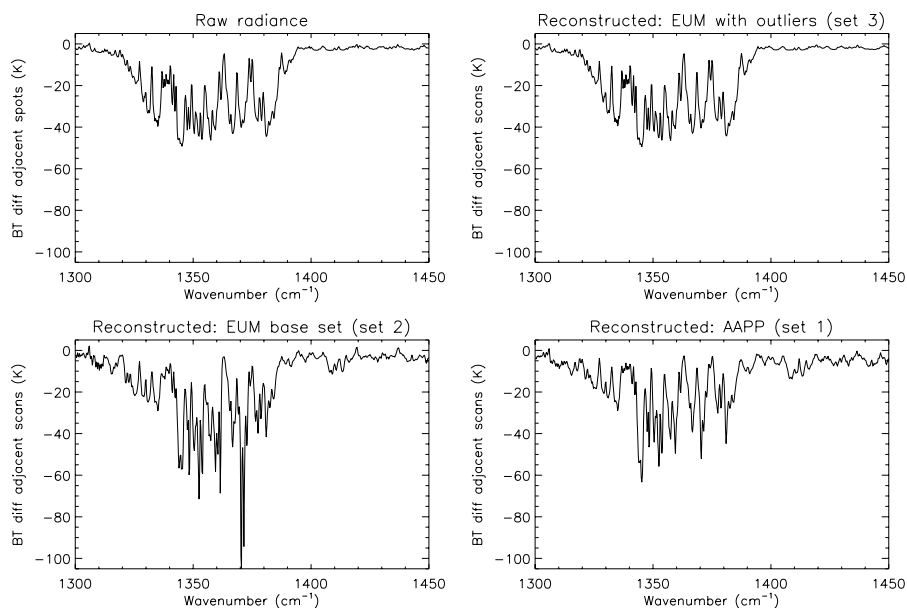


Fig. 6. Difference spectra between a spot on the north-eastern edge of the Kasatochi plume and a spot on the previous IASI scan line which is outside the plume. The ν_3 band of SO₂ covers the range 1320 to 1400 cm⁻¹.

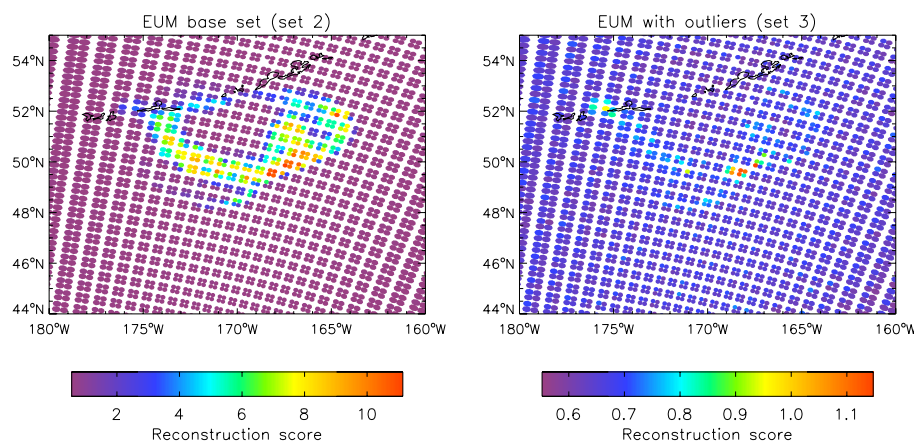


Fig. 7. Band 2 reconstruction scores for the Kasatochi eruption, showing the improvement in scores when outliers are included in the PC training set.

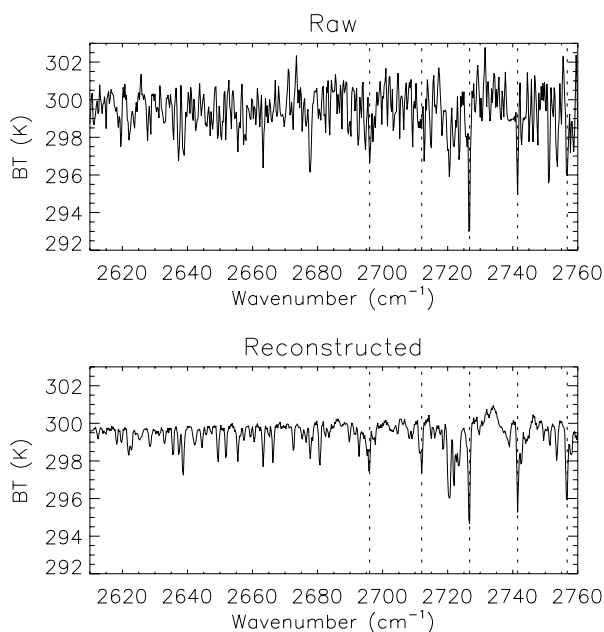


Fig. 8. Raw and reconstructed brightness temperature spectra showing methane lines in IASI band 3, for a night-time overpass. Vertical dotted lines mark the positions of methane absorption lines, from Razavi et al., 2009. The other spectral lines in the 2620 to 2690 cm^{-1} range (and also at 2720 cm^{-1}) are HDO. Data taken from 25 August 2007, 19:38 MetOp overpass, over the Mediterranean (same case as Fig. 1, but well away from the biomass burning jet plume).

3.4 Retrieval of carbon monoxide

The impact of PC compression on the retrieval of carbon monoxide from IASI observations has been investigated using the University of Leicester IASI Retrieval Scheme (ULIRS). This scheme has been developed at the University of Leicester to specifically retrieve CO from IASI measured

Top of Atmosphere (TOA) radiances, utilising an Optimal Estimation Method (OEM) (Rodgers, 2000). As with any OEM it makes use of an a priori profile and covariance matrix to restrain the ill-conditioned nature of the retrieval problem. The ULIRS utilises the Oxford RFM as a forward model, which itself is based on the line-by-line model GENLN2 (Edwards, 1992). The CO a priori profile and covariance matrix have been constructed using over 8000 profiles from the TOMCAT CTM (Chipperfield, 2006), which incorporate a wide variety of CO scenarios, and the retrieval scheme makes use of the spectral interval 2143–2181 cm^{-1} , for reasons outlined in Barret et al. (2005). This spectral interval contains 10 CO lines. The ULIRS utilises a spatially precise surface elevation and emissivity (Seeman et al., 2008), as well as incorporating a quantified back-scattered solar radiation component into the retrieval process. More details of the ULIRS can be found in Illingworth et al. (2010), which is currently in preparation.

Unlike the other gases that are under investigation, the CO retrieval does not use a simple differencing technique, so we have tested the impact of the use of reconstructed radiances on the retrievals themselves, using PC scores calculated from set 1 and set 3 eigenvectors. We have tested the direct substitution of reconstructed radiances for raw radiances in the retrieval scheme, and therefore the same error covariances were used in each case.

Figure 9 shows raw and reconstructed spectra using set 3 eigenvectors and the difference between them, for one observation with relatively high retrieved CO. The reconstructed spectrum is extremely close to the raw data in this spectral region. There is no marked correspondence between the differences and the location of the CO lines in general, although the spectra do differ at 2162 cm^{-1} , close to the centre of a CO absorption line. Furthermore, the differences seen in this spectral band for this observation are much smaller than some of those seen in the CH_4 case (Fig. 8).

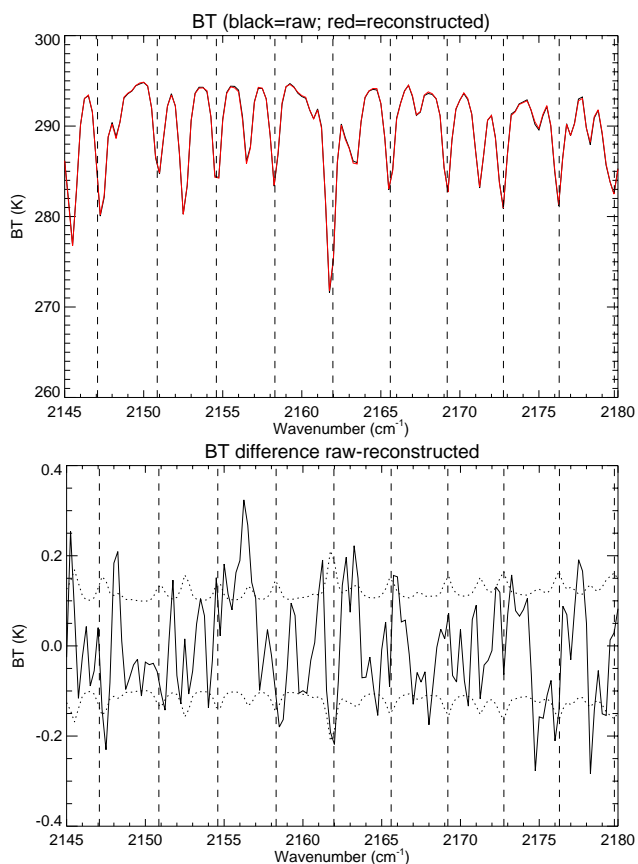


Fig. 9. Raw and reconstructed brightness temperature spectra (left) and the difference between them (right) for an enhanced CO observation. Set 3 eigenvectors were used. CO absorption lines are marked with vertical dashed lines. The set 3 noise level is plotted as a dotted line in the difference plot.

Figure 10 shows retrievals for clear sky observations from 1 April 2008 over the Congo river basin and southern Cameroon from raw IASI radiances, set 1 and set 3 reconstructed radiances. Higher CO levels are observed in the north of the area of study. The retrievals from the reconstructed radiances are similar to those from the raw radiances. Set 1 retrievals show some small differences in the tropospheric column amount, but the overall pattern of high and low retrieved values corresponds well. The set 3 retrievals seem to show a slightly more coherent pattern of high retrieved CO for the northernmost points, but are otherwise very similar. Figure 11 shows the difference in retrieved CO for set 3 eigenvectors, and retrieved CO for raw radiances, plotted as a fraction of the estimated total column retrieval error. The differences between raw and set 3 radiances produce differences in retrieved values which are generally less than 50% of the retrieval error. The mean retrieval error was 24.4% of the total column, with a standard deviation of 3.4%.

One way in which the performance of a retrieval scheme can be assessed is in terms of the fit of the retrieved quantities to the data used in the retrieval. Figure 12 shows the average

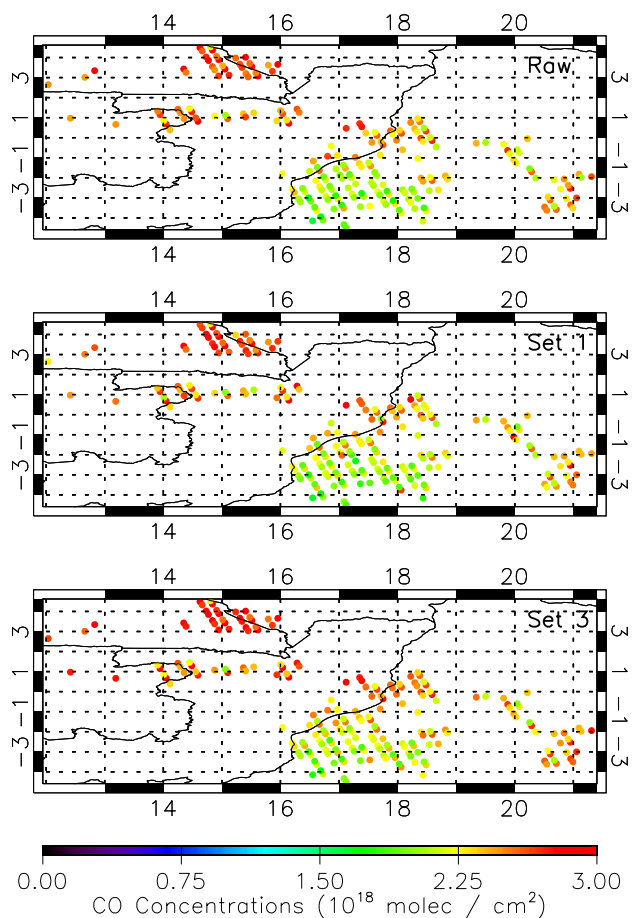


Fig. 10. Total column CO retrievals for 1 April 2008 over West Central Africa using (a) IASI raw radiances, (b) set 1 eigenvectors, (c) set 3 eigenvectors.

residual fit to IASI of the forward-modelled CO retrievals from Fig. 10. In both cases, the residual fit is within the instrument noise across the spectral region, indicating that the principal component compression has had no major effect on the behaviour of the retrieval scheme. The set 1 residuals (not shown) were slightly worse, going outside instrument noise for two channels, which is indicative of a poor retrieval for some footprints. This is in keeping with the results for the other trace gases which indicated that the set 1 eigenvectors performed less well than set 3 for trace gases associated with biomass burning.

It is emphasised that this is a one-off case study, and more detailed studies would be needed to assess whether or not there are any statistically significant systematic biases in the retrieved fields for raw versus reconstructed radiances. Also, further studies could investigate the effect of modifying the observation covariance matrix for the reconstructed radiances, to account for the increase in correlated noise and reduction in random noise, discussed in Sect. 2.

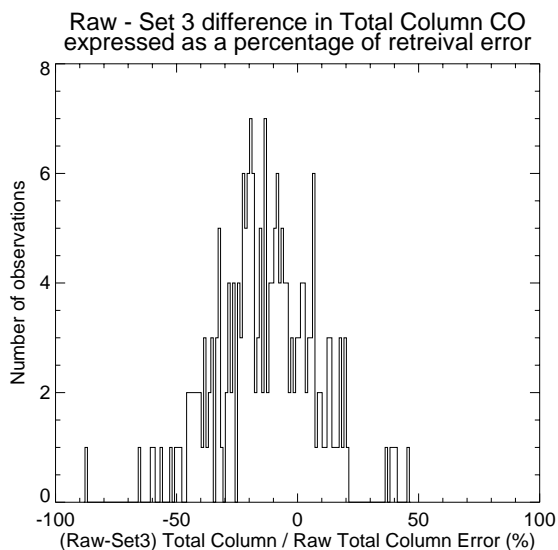


Fig. 11. Difference in total column CO amount retrieved from set 3 eigenvectors and raw radiances, expressed as a fraction of retrieval error.

4 Further refinement of the eigenvectors

The previous section has shown that by adding outliers to the initial training set, the resulting reconstructed radiances retain qualitatively many of the trace gas features observed in the raw radiances, for these case studies. However, close examination of the reconstruction scores showed that there was still scope for improvement. Therefore to generate a set of eigenvectors for operational use, further iterations were carried out. The process involved the addition of further outlier spectra to the training set – identified through the reconstruction scores, and also through more detailed examination of reconstruction errors in specific parts of the IASI spectrum. The aim of this process is to minimize the number of spectra with anomalously high reconstruction scores.

At the same time the noise normalisation matrix, \mathbf{N} , was refined. The refined noise covariance estimate was computed as a sum of the covariance of the reconstruction residual and the noise covariance of the reconstructed radiances. The latter can be computed if an estimate of the original IASI noise covariance, \mathbf{R} , is known (see Sect. 2). The IASI noise covariance from CNES (referenced in Sect. 2) was used for \mathbf{R} in this computation. In a second iteration the previously obtained overall noise estimate was used for \mathbf{R} . It is to be expected that the refined noise normalisation, \mathbf{N} , differs somewhat from the CNES covariance; the latter is derived from measurements of the on-board black body, which is at a fixed temperature, whereas \mathbf{N} is a measure of the radiance noise at typical atmospheric scene temperatures. For band 1 the radiance noise is nearly independent of scene temperature but for band 3 in particular there is a significant increase in noise with scene temperature.

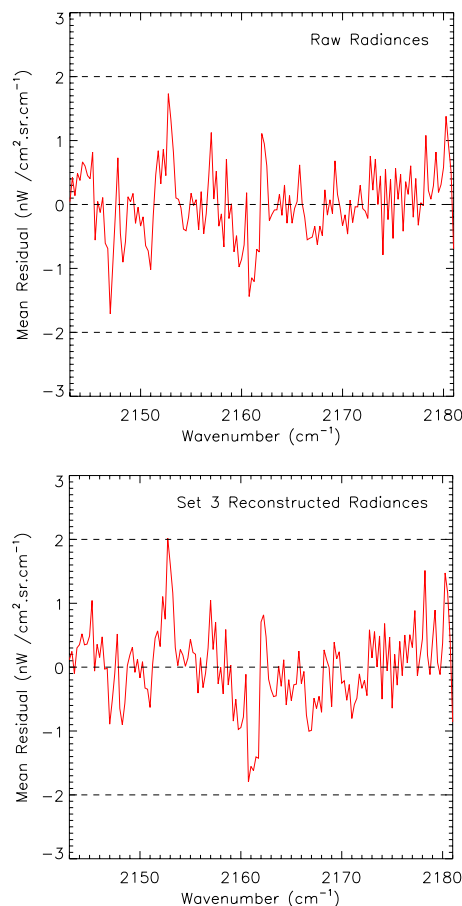


Fig. 12. Mean residual fit to IASI observations from all footprints in Fig. 10 for the raw radiances (left) for the set 3 reconstructed radiances (right). The dashed lines show the instrument noise level.

The use of eigenvectors based on the new noise normalisation led to reduced spatial correlation of the residuals: spatial correlations would imply the presence of scene structure that is not being represented in the reconstructed radiances.

The training set from which the final set of eigenvectors (set 4) was generated comprises the following data:

- The base set of 74 719 spectra, selected from 7 days of IASI data (see Sect. 2). Having selected the base set, the noise normalisation matrix was computed, as described above.
- 26 150 outlier spectra selected from 12 months of IASI data. An iterative procedure was used in which the eigenvectors were re-computed as further sets of outliers were identified; 7 iterations were used.
- A further 960 outlier spectra selected from 15 months of IASI data, using the thresholds 1.2, 1.25 and 1.45 in the reconstruction scores for the three bands. Care was taken to exclude anomalous spectra resulting from instrument mode transitions, and also to exclude anomalous spectra associated with the South Atlantic Anomaly.

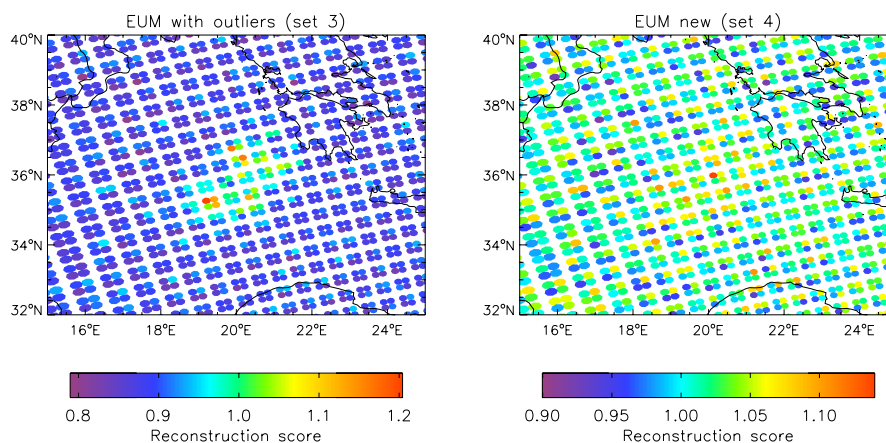


Fig. 13. Band 1 reconstruction scores for the case presented in Fig. 1, showing the improvement in homogeneity when the new set of EUMETSAT eigenvectors is used (see Sect. 4). Note the expanded scale in the right hand plot, allowing the background variability outside the biomass burning plume to be seen.

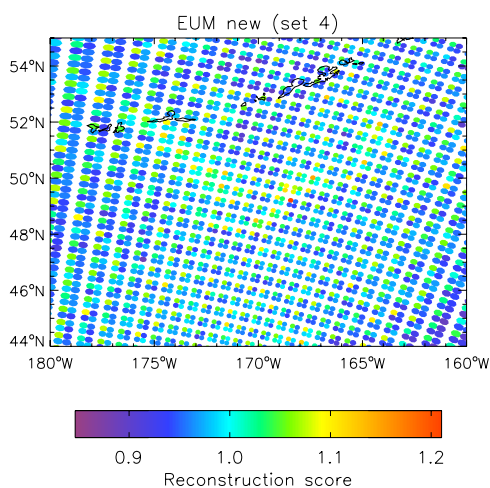


Fig. 14. Band 2 reconstruction scores for the Kasatochi eruption, using the set 4 eigenvectors. To be compared with Fig. 7.

Figure 13 shows the band 1 reconstruction scores for the biomass burning plume presented in Fig. 1 – comparing the scores from the new set of eigenvectors (set 4) with those from set 3. The new set is much more homogeneous than the old set (noting the expanded scale), and the mean reconstruction score is very close to 1.0, confirming that the new estimate for the noise covariance is more accurate than the old.

Similarly, Fig. 14 shows the band 2 reconstruction scores for the SO₂ case using the set 4 eigenvectors - compare this with Fig. 7. Again, the new set is more homogeneous than the old set, though the SO₂ plume is just visible in a few spots.

It is proposed to use this new set of eigenvectors in the pilot phase of the EARS-IASI system and in a trial

parallel dissemination of PC-compressed global IASI data. It is planned that the datasets for these services will include 290 PCs, together with 366 raw channel radiances per spectrum. Details are given in the EARS Operational Service Specification, available from <http://www.eumetsat.int/>. For a definitive specification of the eigenvectors associated with a particular EUMETSAT product, and their derivation, the user is advised to consult the EUMETSAT operational documentation. An updated version of the AAPP package allows users to read the eigenvectors and readily generate reconstructed radiances (Atkinson et al., 2009).

The performance of the scheme will need to be fully tested by the IASI user community, preferably over an extended period of at least one year in order to take seasonal cycles into account. Updates to the eigenvectors may be needed in the future if deficiencies are found in the training set, or if there are changes in the instrument characteristics, ground processing or atmospheric state. If updates are provided then there may be a need for further periods of testing. Note, however, that the method of selecting the training set does not depend on the prior identification of chemical species of interest – this is one of the strengths of using real spectra in the training set, as opposed to spectra simulated using a forward model.

5 Conclusions

PC compression allows IASI spectra to be disseminated and archived efficiently and additionally reduces considerably the amount of instrument noise in the reconstructed spectra. Random noise is much reduced, though there is an increase in correlated noise. The case studies in this paper have shown that (i) trace gas signals can be retained in the reconstructed spectra, although the amount of information that might be

lost is still uncertain; and (ii) their signal-to-noise ratios can be enhanced. It is very important that the training set used in the generation of the eigenvectors contains adequate representation of rare events.

Random selection of the training set spectra (as was used in the eigenvectors distributed in 2008 with AAPP) does not adequately represent rare events such as biomass burning or volcanic eruptions. Moreover it does not capture the more persistent but geographically limited ammonia emissions associated with agriculture. However, an iterative procedure, involving refinement of a base training set by the addition of outlier spectra, is successful in improving the residuals, such that there is good qualitative agreement between raw and reconstructed radiances. The final set of eigenvectors, described in Sect. 4, will form the basis of the PC compression in the pilot phase of the EARS-IASI system. There will be a need for thorough testing of the scheme by the IASI user community.

Whilst the qualitative results shown in this paper are very encouraging for PC reconstructions, it is recommended that more detailed and quantitative studies are carried out into trace gas retrievals using reconstructed radiances. For example, the use of PCA in the analysis of weak signals, such as NO₂ or CO₂ variability retrieved from IASI (Ricaud et al. 2009; Crevoisier et al., 2009), particularly in the context of GMES-MACC, is an area that would merit further study.

Edited by: H. Worden

References

- Antonelli, P., Revercomb, H. E., Sromovsky, L., Smith, W. L., Knuteson, R. O., Tobin, D. C., Garcia, R. K., Howell, H. B., Huang, H.-L., and Best, F. A.: A principal component noise filter for high spectral resolution infrared measurements, *J. Geophys. Res.*, 109, D23102, doi:10.1029/2004JD004862, 2004.
- Atkinson, N. C., Brunel, P., Marguinaud, P., and Labrot, T.: AAPP developments and experiences with processing METOP data, *Tech. Proc. 16th Int. TOVS Study Conf.*, Angra dos Reis, Brazil, 6–13 May, 2008.
- Atkinson, N. C., Ponsard, C., and Hultberg, T.: AAPP enhancements for the EARS-IASI service, *Proc. EUMETSAT Meteorological Satellite Conf.*, Bath, UK, 21–25 September, 2009.
- Barret, B., Turquety, S., Hurtmans, D., Clerbaux, C., Hadji-Lazaro, J., Bey, I., Auvray, M., and Coheur, P.-F.: Global carbon monoxide vertical distributions from spaceborne high-resolution FTIR nadir measurements, *Atmos. Chem. Phys.*, 5, 2901–2914, doi:10.5194/acp-5-2901-2005, 2005.
- Beer, R., Shephard, M. W., Kulawik, S. S., Clough, S. A., Eldering, A., Bowman, K. W., Sander, S. P., Fisher, B. M., Payne, V. H., Luo, M., Osterman, G. B., and Worden, J. R.: First satellite observations of lower tropospheric ammonia and methanol, *Geophys. Res. Lett.*, 35, L09801, doi:10.1029/2008GL033642, 2008.
- Chipperfield, M. P.: New version of the TOMCAT/SLIMCAT offline chemical transport model: Intercomparison of stratospheric tracer experiments, *Q. J. Roy. Meteor. Soc.*, 132, 1179–1203, 2006.
- Clarisse, L., Coheur, P. F., Prata, A. J., Hurtmans, D., Razavi, A., Phulpin, T., Hadji-Lazaro, J., and Clerbaux, C.: Tracking and quantifying volcanic SO₂ with IASI, the September 2007 eruption at Jebel at Tair, *Atmos. Chem. Phys.*, 8, 7723–7734, doi:10.5194/acp-8-7723-2008, 2008a.
- Clarisse, L., Coheur, P.-F., Hurtmans, D., Clerbaux, C., Turquety, S., and Hadji-Lazaro, J.: IASI measurements of Trace Species and Particles in Volcanic and Fire Plumes, *ACCENT-TROPOSAT-2 Activities 2007–2008 and Final Report*, 82–86, 2008b.
- Clarisse, L., Clerbaux, C., Dentener, F., Hurtmans, D., and Coheur, P.-F.: Global ammonia distribution derived from infrared satellite observations, *Nat. Geosci.*, 2, 479–483, 2009.
- Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, *Atmos. Chem. Phys.*, 9, 6041–6054, doi:10.5194/acp-9-6041-2009, 2009.
- Coheur, P.-F., Clarisse, L., Turquety, S., Hurtmans, D., and Clerbaux, C.: IASI measurements of reactive trace species in biomass burning plumes, *Atmos. Chem. Phys.*, 9, 5655–5667, doi:10.5194/acp-9-5655-2009, 2009.
- Collard, A. D.: Selection of IASI channels for use in numerical weather prediction, *ECMWF Tech. Memorandum*, 532, 2007.
- Collard, A. D.: NWPSAF ECMWF IASI PCA-Based Compression Package Manual, NWP SAF document NWPSAF-EC-UD-011, available online at: <http://www.nwpsaf.org/>, 2008.
- Crevoisier, C., Chédin, A., Matsueda, H., Machida, T., Armante, R., and Scott, N. A.: First year of upper tropospheric integrated content of CO₂ from IASI hyperspectral infrared observations, *Atmos. Chem. Phys.*, 9, 4797–4810, doi:10.5194/acp-9-4797-2009, 2009.
- Crevoisier, C., Nobileau, D., Fiore, A. M., Armante, R., Chédin, A., and Scott, N. A.: Tropospheric methane in the tropics – first year from IASI hyperspectral infrared observations, *Atmos. Chem. Phys.*, 9, 6337–6350, doi:10.5194/acp-9-6337-2009, 2009.
- Edwards, D. P.: GENLN2: A general line-by-line atmospheric transmittance and radiance model, *NCAR Tech. Note*, NCAR/TN-367+STR, 1992.
- George, M., Clerbaux, C., Hurtmans, D., Turquety, S., Coheur, P.-F., Pommier, M., Hadji-Lazaro, J., Edwards, D. P., Worden, H., Luo, M., Rinsland, C., and McMillan, W.: Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors, *Atmos. Chem. Phys.*, 9, 8317–8330, doi:10.5194/acp-9-8317-2009, 2009.
- Goldberg, M. D., Qu, Y., McMillin, L. M., Wolf, W., Zhou, L., and Divakarla, M.: AIRS near-real-time products and algorithms in support of operational numerical weather prediction, *IEEE Trans. Geosci. Rem. Sens.*, 41, 379–389, 2003.
- Hilton, F., Atkinson, N. C., English, S. J., and Eyre, J. R.: Assimilation of IASI at the Met Office and assessment of its impact through observing system experiments, *Q. J. Roy. Meteor. Soc.*, 135, 495–505, 2009a.
- Hilton, F. and Collard, A. D.: Recommendations for the use of principal component-compressed observations from infrared hyperspectral sounders, *Met Office Forecasting Re-*

- search and Development Technical Report 536, available online at: http://research.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/, access: 3 June 2010, 2009b.
- Hultberg, T.: IASI Principal Component Compression (IASI PCC) FAQ, Technical Note, available online at: <http://www.eumetsat.int/Home/Main/DataProducts/Resources/index.htm>, access: 27 July 2010, 2009.
- Hultberg, T.: IASI Principal Component Compression (PCC) – first experiences, 2nd IASI International Conference, Annecy, 25–29 January, 2010.
- Illingworth, S. M., Remedios, J. J., Boesch, H., Moore, D. P., Sembhi, H., Dudia, A., and Walker, J. C.: A New Optimal Estimation Retrieval Scheme for Carbon Monoxide using IASI spectral radiances: Part I Sensitivity Analysis, Error Budget and Simulations, in preparation, 2010.
- Phulpin, T., Blumstein, D., Prel, F., Tournier, B., Prunet, P., and Schlusser, P.: Applications of IASI on MetOp-A: first results and illustration of potential use for meteorology, climate monitoring, and atmospheric chemistry, *Proc. SPIE*, 6684, 66840F, 2007.
- Razavi, A., Clerbaux, C., Wespes, C., Clarisse, L., Hurtmans, D., Payan, S., Camy-Peyret, C., and Coheur, P. F.: Characterization of methane retrievals from the IASI space-borne sounder, *Atmos. Chem. Phys.*, 9, 7889–7899, doi:10.5194/acp-9-7889-2009, 2009.
- Ricaud, P., Attié, J.-L., Teyssède, H., El Amraoui, L., Peuch, V.-H., Matricardi, M., and Schluessel, P.: Equatorial total column of nitrous oxide as measured by IASI on MetOp-A: implications for transport processes, *Atmos. Chem. Phys.*, 9, 3947–3956, doi:10.5194/acp-9-3947-2009, 2009.
- Rodgers, C.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific, 2000.
- Seemann, S. W., Borbas, E. E., Knuteson, R. O., Stephenson, G. R., and Huang, H.-L.: Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals, *J. Appl. Meteorol. Climatol.*, 47, 108–123, 2008.
- Siméoni, D., Singer, C., and Chalon, G.: Infrared atmospheric sounding interferometer, *Acta Astronautica*, 40, 113–118, 1997.
- Turquety, S., Hurtmans, D., Hadji-Lazaro, J., Coheur, P.-F., Clerbaux, C., Josset, D., and Tsamalis, C.: Tracking the emission and transport of pollution from wildfires using the IASI CO retrievals: analysis of the summer 2007 Greek fires, *Atmos. Chem. Phys.*, 9, 4897–4913, doi:10.5194/acp-9-4897-2009, 2009.