The EUMETSAT Network of Satellite Application Facilities

NWP SAF

Numerical Weather Prediction

Document NWPSAF_MF_VS_005

Version 1.0

20/2/15

Preliminary studies towards the use of IASI PC products in the Météo-France global data assimilation system

Stephanie Guedj, Vincent Guidard, and Jean-Francois Mahfouf

Met Office

ECMWF

Royal Netherlands Meteorological Institute
Ministry of Transport, Public Works and Water Management

METEO FRANCE
Toujours un temps d'avance

Preliminary studies towards the use of IASI PC products in the Météo-France global data assimilation system

Stephanie Guedj, Vincent Guidard, and Jean-Francois Mahfouf

CNRM-GAME/METEO-FRANCE and CNRS
42 av. Gaspard Coriolis, 31057 Toulouse, FRANCE

This documentation was developed within the context of the EUMETSAT Satellite Application Facility on Numerical Weather Prediction (NWP SAF), under the Cooperation Agreement dated 29 June 2011, between EUMETSAT and the Met Office, UK, by one or more partners within the NWP SAF. The partners in the NWP SAF are the Met Office, ECMWF, KNMI and Météo France.

Copyright 2015, EUMETSAT, All Rights Reserved.

Change record			
Version	Date	Author / changed by	Remarks
1.0	20/2/15	Stephanie Guedj, Vincent Guidard, and Jean-Francois Mahfouf	

Preliminary studies towards the use of IASI PC products in the Météo-France global data assimilation system

Report : NWP-AS14-P01

Stephanie Guedj, Vincent Guidard*,
and Jean-Francois Mahfouf

CNRM-GAME/METEO-FRANCE and CNRS
42 av. Gaspard Coriolis, 31057 Toulouse, FRANCE
*correspondig author: vincent.guidard@meteo.fr

February 20, 2015

Abstract

A pre-processing chain has been set up in the Météo-France research environment to get Eumetsat BUFR of IASI PC scores and reconstruct radiance spectra from these PC scores. Reconstructed data have been compared with original data and globally showed a good agreement. Nevertheless, some discrepancies occuring either over cloudy areas or over land indicating that some numerical errors may occur either in the coding of the BUFR files or in the reconstruction of the spectra in Météo-France pre-processing.

When ingested in the assimilation process of the global model ARPEGE, reconstructed BTs have the same impact on analyses and forecasts as the original BTs. No impact is seen on the bias correction, which means that assimilating reconstruted BTs introduces no bias in the system. The standard deviations of observations minus simulations from model fields are smaller for reconstructed BTs than for the original BTs, which is consistent with the a posteriori Desroziers diagnostic. Especially in band 1, reconstructed BTs are less "noisy" and could be assimilated with smaller σ_o (which means with a larger weight). But it appears that the PCA technique introduces inter-channel correlation in the observation errors.

Results are consistent with those obtained at the Met-Office.

Contents

Acknowledgements

This research is funded by the NWP-SAF of Eumetsat. The authors would like to thank Pascal Brunel from Météo-France for his help with RTTOV. The authors warmly thank Eumetsat teams, in particular Tim Hultberg for his help on PC scores. Many thanks to Jean Maziejewski for his help to revise the manuscript.

Acronyms

AROME	Applications of Research to Operations at MEsoscale
ARPEGE	Action de Recherche Petite Echelle Grande Echelle
BT	Brightness Temperature
CrIS	Cross-track Infrared Sounder
ECMWF	European Centre for Medium Range Forecasts
GNSS	Global Navigation Satellite System
IASI	Infrared Atmosperic Sounding Interferometer
IRS	InfraRed Sounder
ITCZ	InterTropical Convergence Zone
MTG	Meteosat Third Generation
NWP	Numerical Weather Prediction
PC	Principal Component
PCA	Principal Componet Analysis
RTTOV	Radiative Transfer for TOVS
SAF	Satellite Application Facility
STD	Standard Deviation
TIROS	Television InfraRed Observation Satellite
TOVS	TIROS Operational Vertical Sounder

1 Work plan

The activity will be to examine the use of the real-time Eumetsat IASI PC scores in order to reconstruct IASI radiances and to assimilate them in the Météo-France ARPEGE global model using a 4D-Var data assimilation system. Data monitoring of reconstructed radiances will be implemented and assimilation experiments performed over a two month period. These two aspects will be compared to the current operational practice at Météo-France where 314 original IASI radiances are monitored and about 123 assimilated. The main components of the work are as follows.

1. Install and test the various tools and data (available from the NWP SAF) needed to reconstruct the radiances on local computers. The computational efficiency of the reconstruction of the original spectrum will be evaluated for operational purposes. This could lead to the selection of a reduced number of PC scores to reconstruct the IASI radiances.

2. Monitor the original set of selected IASI radiances and the reconstructed ones through the 4D-Var assimilation system "screening". Special attention will be paid to low peaking channels and to channels sensitive to water vapour since they have rather large variability. The variational bias correction scheme will be applied in the 4D-Var to the reconstructed radiances and resulting biases will be compared to those from the original radiances. The comparison will be restricted to clear sky pixels.

3. Examine the specification of observation errors. Reconstructed radiances should have lower error values (noise reduction) but these errors should be correlated between channels (and this could require an artificial increase of diagonal errors). Diagnostics will be performed on background and analysis departures to estimate observation error variances and correlations using the a-posteriori Desroziers method.

4. Prepare report for publication on the NWP SAF web-site. The conclusions from this study will provide guidance on the strength and weaknesses of the reconstructed IASI radiances in various aspects of a global NWP data assimilation system (e.g. bias correction, cloud detection, channel selection, error specification). Recommendations will be given on the use of PC scores for the monitoring and the assimilation of reconstructed radiances both from a data provider perspective and from a data user perspective. The study will be undertaken using the same version of the RTTOV radiative transfer model to simulate real and reconstructed radiances, with an implicit assumption that it is not a dominant source of differences.

2 Introduction

This report focuses on the assimilation of reconstructed radiances from PC compression in the framework of NWP applications. In the last few years, with each generation of satellite instruments, the data volumes have increased dramatically. The Infrared Atmospheric Sounding Interferometer (IASI; Cayla, 2001) is an infrared Fourier transform spectrometer which was first launched on the MetOp-A satellite in October 2006. It is the first such instrument which is truly operational, and has 8461 channels covering the spectral interval from 645 to 2760 cm^{-1} at a resolution of 0.5 cm^{-1} (apodised). IASI radiances are currently being assimilated operationally at a number of NWP centres (e.g. Hilton et al., 2011; Guidard et al.,2011). A significant challenge in the coming decade will be to accommodate MTG-IRS, the hyperspectral sounder on Meteosat Third Generation (MTG), scheduled to launch in 2021.

Rationale

The ability to assimilate hyperspectral sounding data in the form of reconstructed radiances is likely to be very useful for making improved use of IASI and CrIS data in the NWP context and for preparing for MTS-IRS data. Indeed, the recent growth in satellite radiances from infrared hyperspectral instruments available to the NWP community raises issues due to the huge increase in data volumes both for assimilation and storage. The dissemination cost is also becoming an issue for satellite agencies. An efficient representation of these satellite radiances is therefore highly desirable. The Principal Component Analysis (PCA) is an appealing method for compressing information without significant losses. The PCA has also filtering properties that could be of interest by reducing the noise from the original data. Since February 2011, Eumetsat has been operationally disseminating IASI PC scores through EumetCast. Even though the direct assimilation of PC scores could be possible, as experienced recently by ECMWF, it would necessitate, for many NWP centres, a significant investment to adapt many aspects of their data assimilation systems. An interesting alternative is the direct assimilation of reconstructed radiances. Such activity has started at the Met Office but also needs to be undertaken by other NWP centres. Since the generation of PC scores relies on various hypotheses regarding the size of the training database (that could lead to representativeness issues) and the noise normalisation (that requires the specification of an error covariance matrix), real and reconstructed radiances could have different properties. It is therefore important to investigate whether the choices made at Eumetsat in order to establish the current real time IASI PC scores are compatible with the requirements of NWP centres that would like to assimilate reconstructed radiances.

3 Monitoring of the original and reconstructed set of IASI radiances

3.1 Pre-processing of IASI Radiances

The operational Eumetsat stream of PC scores from Metop-A and Metop-B has been taken from the EumetCast distribution. Auxiliary data, namely the Eigenvectors, Noise and Mean for the 3 bands, have been downloaded from the Eumetsat website in HDF5 format. The routines to reconstruct radiances have been extracted from the NWP-SAF PCA package and they have been adapted and included in Météo-France NWP pre-processing tool. In order to save some computational time, only the subset of 314 IASI channels, which are routinely monitored in operations, are reconstructed. Then, the reconstructed radiances undergo the same treatment as the original ones. Figure 1 shows an example of reconstructed versus original raw radiances. Comparable structures and brightness temperature (BT) amplitude are found for both datasets after this pre-processing step.

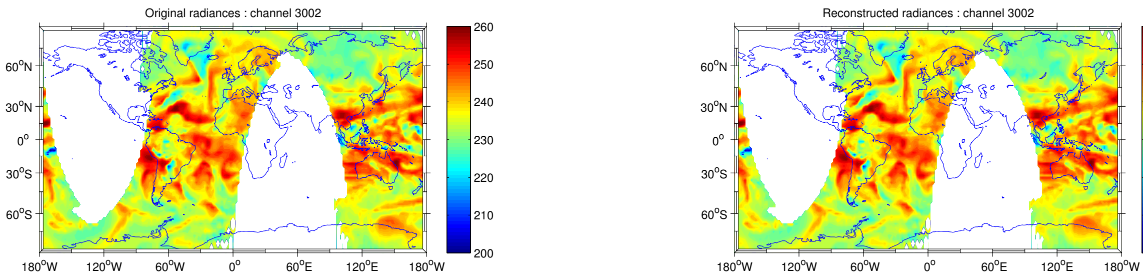


Figure 1: Maps of original raw BT (left) and reconstructed BT (right) for the 20/11/2014 - 09UTC to 15UTC.

3.2 Objective comparison between original and reconstructed radiances

The operational monitoring of raw radiances is actually performed in BT space. The monitoring of reconstructed radiances is also done in BT space. Statistic are computed for different conditions: land and sea surfaces using the ARPEGE land-sea mask, clear and cloudy conditions using the AVHRR cloud cover product. Statistics are also applied seperately on the 3 bands:

- band 1 : from 648,75 to 1142.5 cm^{-1} (temperature, window and ozone)
- band 2 : from 1142.75 to 1927.25 cm^{-1} (humidity)
- band 3 : from 1927.5 to 2646.5 cm^{-1}

Among the pre-selected 314 channels, 165 are located in Band 1, 104 in Band 2 and 45 in Band 3. Figure 2 shows spectrum of BT mean and standard deviation (STD) for the original and reconstructed radiances within the 6-hour time window centered on the

20/11/2014 at 00UTC, which means 80720 pixels × 314 channels. Very good agreement is found between original and reconstructed IASI BT mean et standard deviation show similar values. As mentioned in Table 1, larger errors affect reconstructed radiances over land surface and in cloudy conditions, especially for band 3. Among the pre-selected 314 channels, 165 are located in band 1, 104 in band 2 and 45 in band 3. As shown in Figure 3, relatively to band 1 which exhibits best scores, errors in band 3 are increased by a factor close to 6 for global, 4 over land, 3 over sea, 5 in cloudy pixels and 4 in clear sky conditions.

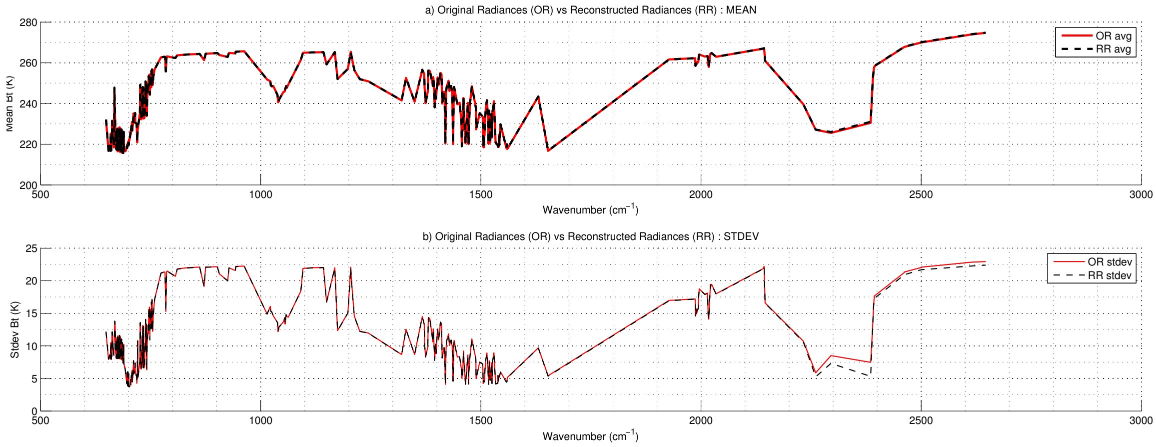


Figure 2: a) Mean and b) standard-deviation of original BT (red line) and reconstructed (black line) as a function of the wavenumber, for data within the 6-hour time window centered on the 20/11/2014 at 00UTC. The channel plotted are those used operationally at Météo-France.

All bands (N=314)	correlation	bias	STD
global	1.0000	0.0182	0.0447
	1.0000	*0.0268*	*0.0360*
land	1.0000	**0.0423**	0.0735
	1.0000	*0.0271*	*0.0529*
sea	1.0000	0.0143	0.0354
	1.0000	*0.0278*	*0.0321*
cloudy	1.0000	**0.0271**	0.0417
	1.0000	*0.0432*	*0.0385*
clear	1.0000	0.0141	0.0476
	1.0000	*0.0140*	*0.0371*

Table 1: Correlation, bias and standard-deviation of original versus reconstruced BTs for all bands for a night-time period (6-hour window centered on 20/11/2014 at 00UTC) and a day-time period in italic (6-hour window centered on the 20/11/2014 at 12UTC). The selected channels are those used operationally at Météo-France (N=314).

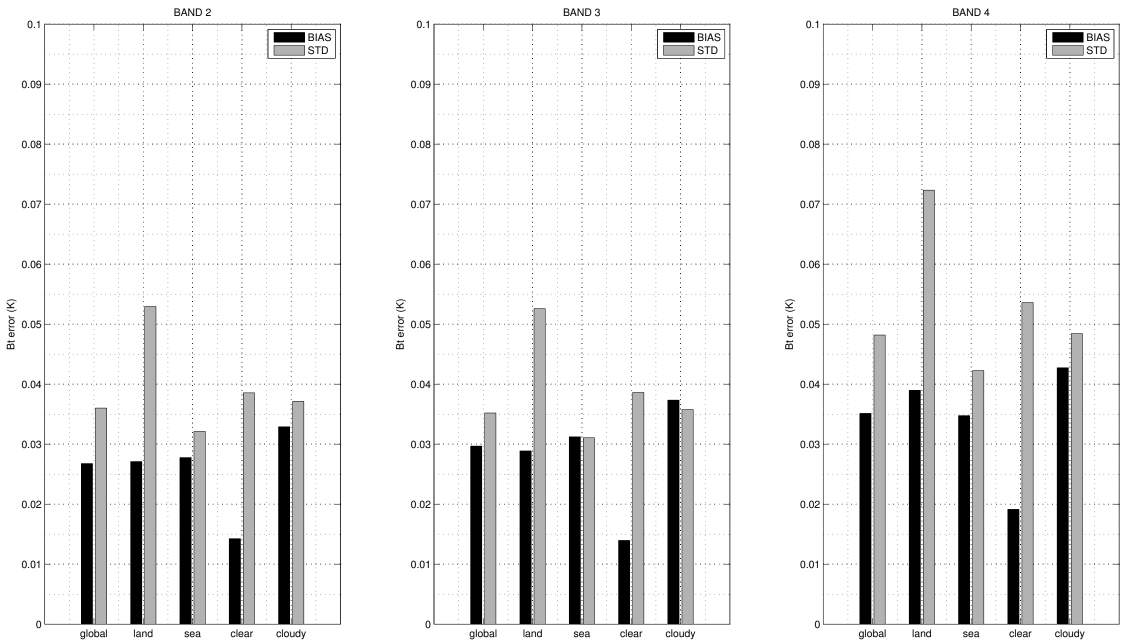


Figure 3: Bias and standard-deviation of original versus reconstructed BTs on the day-time period (6-hour window centered on 20/11/2014 at 12UTC) for band 1, band 2 and band 3. Statistics are applied over a subset of observations for various conditions (all, land, sea, clear cloudy).

Special attention is paid to low peaking channels and to channels sensitive to water vapour since they have rather large variability. 3 channels were selected: channel 226 (701.25 cm^{-1}, Temperature), 1191 (942.5 cm^{-1}, Window) and 3002 (1395.25 cm^{-1}, Humidity) may give good and representative examples. Their weighting functions are plotted in Figure 4.

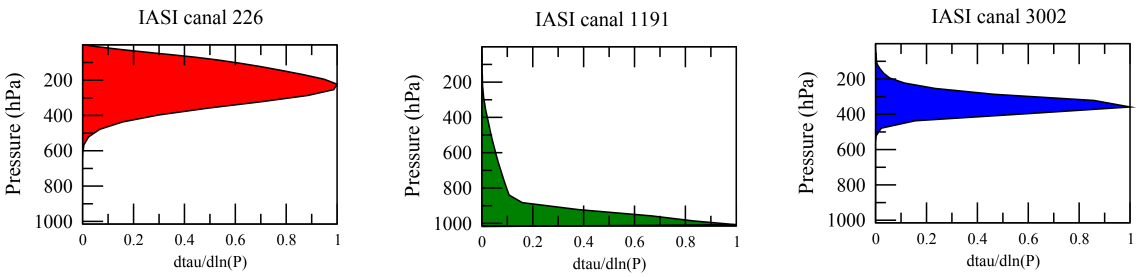


Figure 4: Weighting functions for 3 representative channels : 226 (701.25 cm^{-1}, Temperature), 1191 (942.5 cm^{-1}, Window) and 3002 (1395.25 cm^{-1}, Humidity)

Figure 5 shows maps of original and reconstructed BTs averaged over 10 days (20-30/11/2014) for the 3 selected channels. The averaged difference is also presented. Maps show similar structures and amplitudes. Positive bias is drawn close to the ITCZ indicating that the use of reconstructed radiances produces higher value of BTs than the original radiances in this region. This effect may be related to the presence of clouds (see section 4.2).

Time series of mean and standard deviation differences (reconstructed minus original) are plotted in Figure 6 for the 3 selected channels. Statistics were applied over data filtered for different conditions: global, land, sea, clear and cloudy. In all conditions (global) data, reconstructed BTs exhibit a warm bias and less variability with respect to orignal BTs. The high-tropospheric temperature channel provides smaller differences than both other channels (between -0.1 and 0.1). The highest differences concern the window channel, over land surface. Large biases between original and reconstructed BTs can reach up to 0.5 K. Water Vapour channel statistics show less differences for the clear sky class.

3.3 Radiance simulation skills

Both datasets were simulated using the RTTOV radiative transfer model (v10) over 10 days (from 20/11/2014 to 30/11/2014). Atmospheric profiles (temperature, humidity, ozone) as well as surface parameters (emissivity, skin temperature) were extracted from the operational ARPEGE system. Input atmospheric and surface fields are the same for both simulation experiments. Simulations are calculated with a clear sky assumtion. Figure 7 shows scatterplots of reconstructed BTS minus simulated BTs versus original

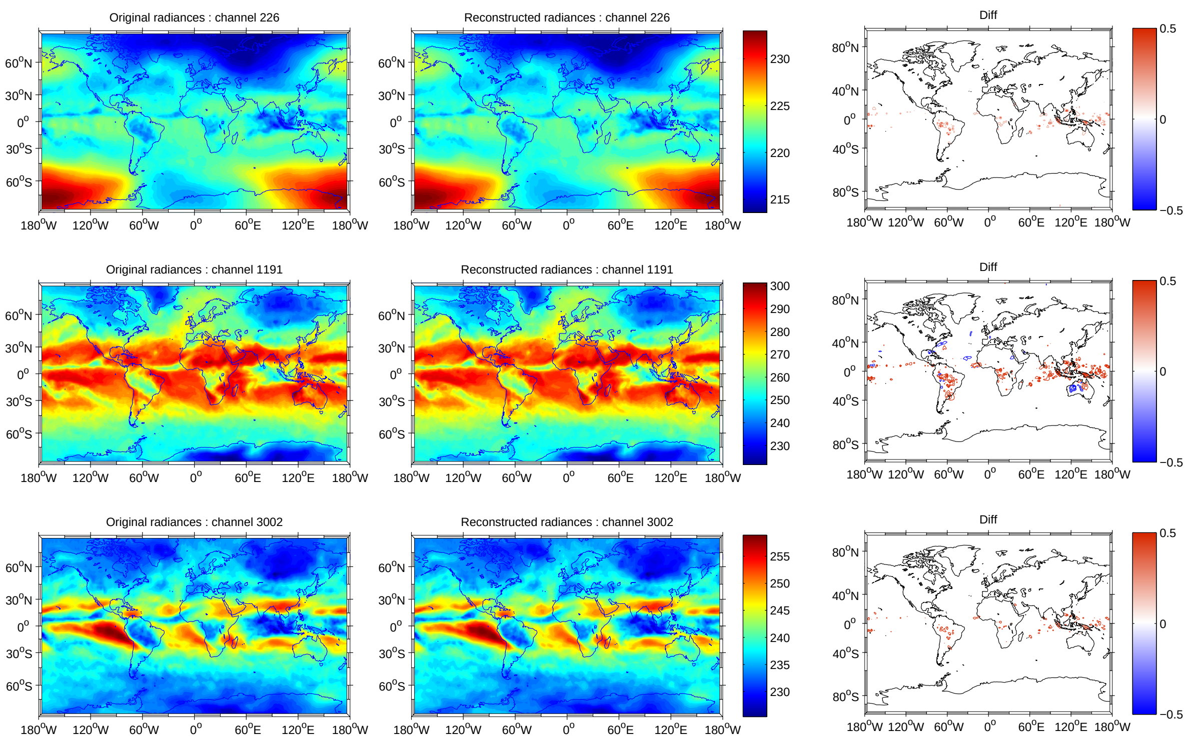


Figure 5: Mean original BTs (left column), reconstructed BTs (middle column) and difference [reconstructed - original] (right column) over 10 days (20-30/11/2014).

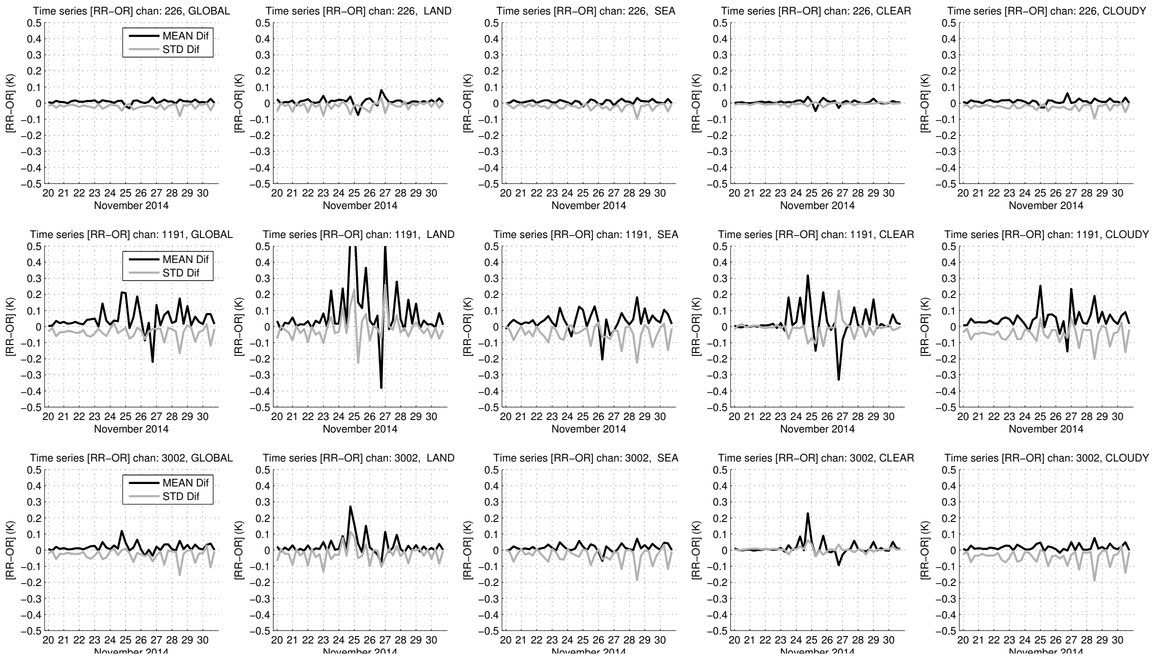


Figure 6: Time series of mean and standard deviation differences (reconstructed minus original) for several conditions (global, sea, land, clear and cloudy) over 10 days (20-30/11/2014), for the 3 selected channels.

BTs minus simulated BTs, for various classes. A better agreement is found in clear sky conditions.

3.4 Conclusion

Reconstructed and original BTs show a good agreement. One has to point out differences occuring in cloudy areas, which could be due to problems in the computation of PC scores in those areas. This could be due to a numerical accuracy in conding in the BUFR format or in the training database. The numerical representation problem could also occur during the reconstruction process included in Météo-France pre-processing.

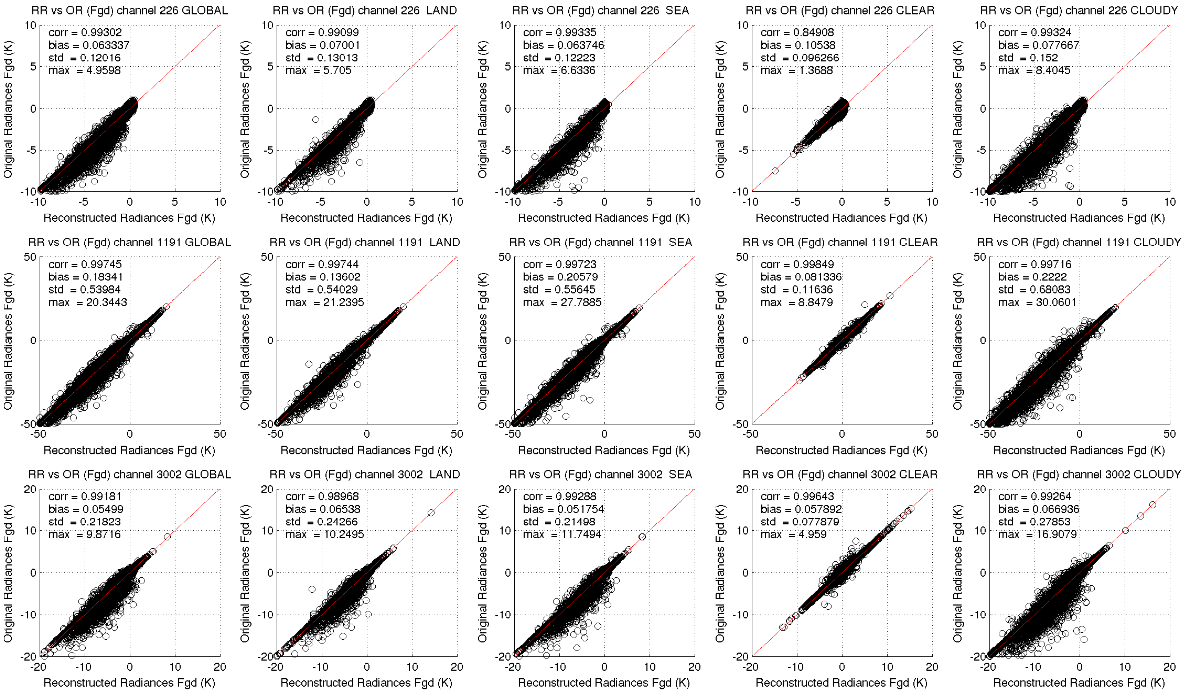


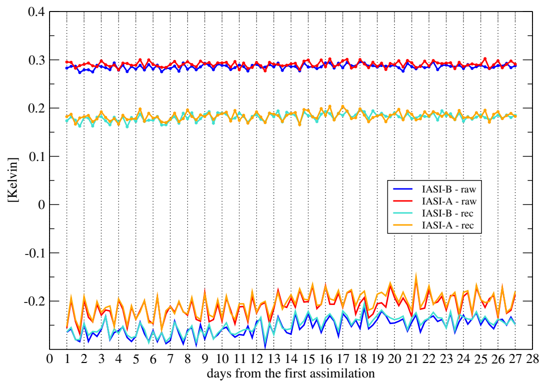
Figure 7: Scatterplots of reconstructed BTs minus simulated BTs versus original BTs minus simulated BTs, columns represent various conditions, one line per channel. Statistics over 10 days (20-30/11/2014).

4 Assimilation experiments of reconstructed radiances

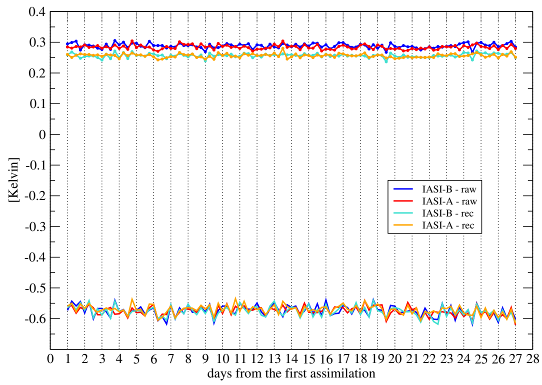
Twin assimilation cycles were carried out: the first one assimilating IASI raw BTs (hereafter RAW) and the other one assimilating reconstructed BTs (hereafter REC). The cycles begin at the analysis date 7 November 2014 at 18UTC and end at the analysis date 7 December 2014 at 18UTC. The standard evaluation procedure has been used to compare the two cycles with each other.

4.1 Impact on the innovations

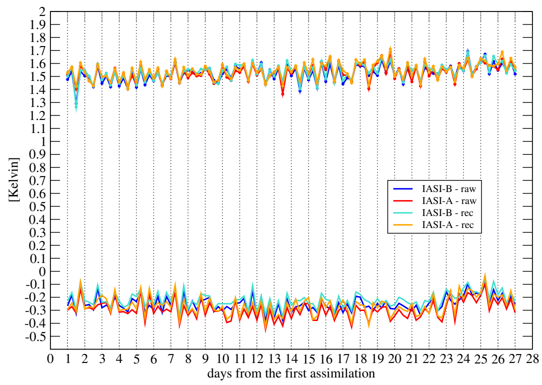
Innovations are the differences between the observed values and the values simulated from model fields. Time series of averages (also called biases) and standard deviations of the innovations are drawn for the 3 reference channels (Figure 8).



(a) channel #0226



(b) channel #1191



(c) channel #3002

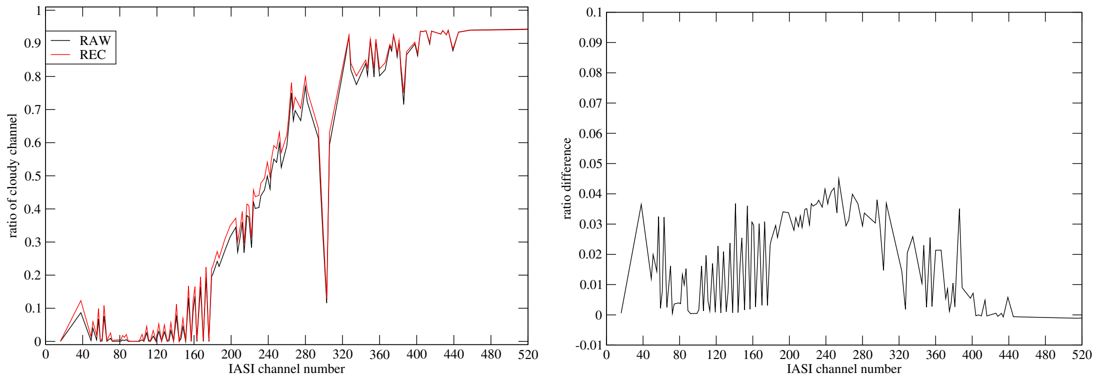
Figure 8: Time series of averages (solid lines - lower curves) and standard deviations (solid lines with bullets - upper curves) for the innovations (observations minus model) before bias correction. Original BTs are in blue (resp. red) for IASI-B (resp. IASI-A) and reconstructed BTs are in cyan (resp. orange) for IASI-B (resp. IASI-A).

RAW and REC exhibit similar biases for all channels, both for IASI-A and IASI-B. The impact of assimilating reconstructed radiances on the variational bias correction

(VarBC) implemented within the assimilation system is neutral. REC has smaller standard deviations than RAW for temperature channels (both atmospheric and window channels), whereas standard deviations are similar for water vapour channels. These findings are consistent with the Principal Componant Analysis technique which is supposed to reduce the noise, leaving biases unchanged.

4.2 Effects on cloud detection

Two algorithms are used in the assimilation process: the McNally and Watts technique -hereafter cloud detect- (McNally and Watts, 2003), which flags each channel of each pixel clear or cloudy, and the CO2-slicing technique (Chahine, 1974; Guidard *et al.*, 2011), which diagnoses the top pressure and fraction of a single-layer cloud in the pixel. The ratio of cloudy data obtained with the cloud detect is higher for REC than for RAW (cf. Figure 9). The ratio increase is about 0.02-0.04 for channels in the lower stratosphere - upper troposphere.



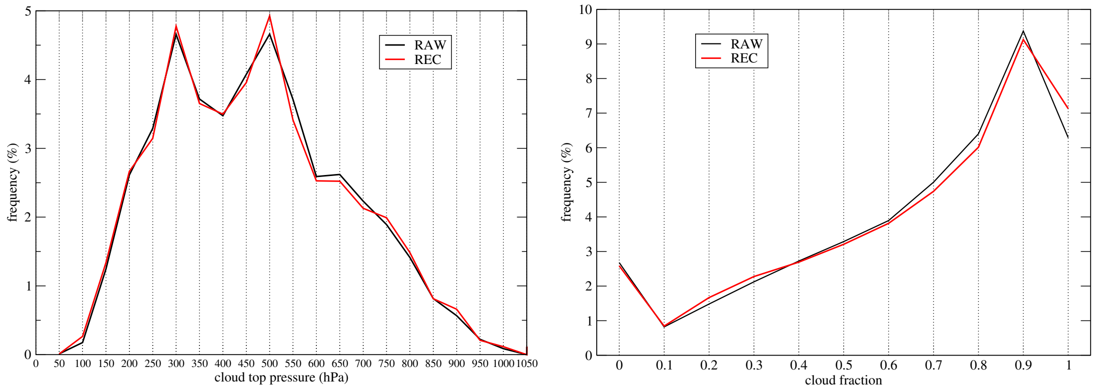
(a) comparison between RAW and REC (b) difference REC minus RAW

Figure 9: Ratio of data flagged cloudy with the McNally and Watts technique for each channel in the long-wave temperature region.

The CO2-slicing technique exhibits neutral impact on retrieved cloud top pressures (cf. Figure 10). Retrieved cloud fractions are more often equal to 1 with REC than with RAW. Which is consistent with the results on the McNally and Watts technique. Both techniques use bias-corrected innovations as inputs. Even though the innovations and the bias correction applied to them are similar between RAW and REC (as seen in previous paragraph), the small differences seem sufficient to modify the cloud characterisation techniques which are very sensitive.

4.3 Analysis and forecast skills

Statistics (averages and standard deviations) on innovations (observations minus model) and residuals (observations minus analysis) have been calculated for the whole set of observations assimilated in ARPEGE (radiosoundings, aircraft data, winds derived from satellite imagery, scatterometer, GNSS radio-occultation, radiances from polar-orbitting



(a) cloud top pressure (b) cloud fraction

Figure 10: Comparison of cloud parameters retrieved using the CO2-slicing technique for RAW (black lines) and REC (red lines).

and geostationary satellites, etc.). Figure 11 gives some statistics for temperature observations from the radiosoundings. Differences between statistics for RAW and REC are hardly visible (same conclusion for all other observations). Impact of assimilating reconstructed radiances instead of raw radiances appears to be negligible on the analyses and 6-hour forecasts. Thus, no scores have been calculated on longer term forecasts.

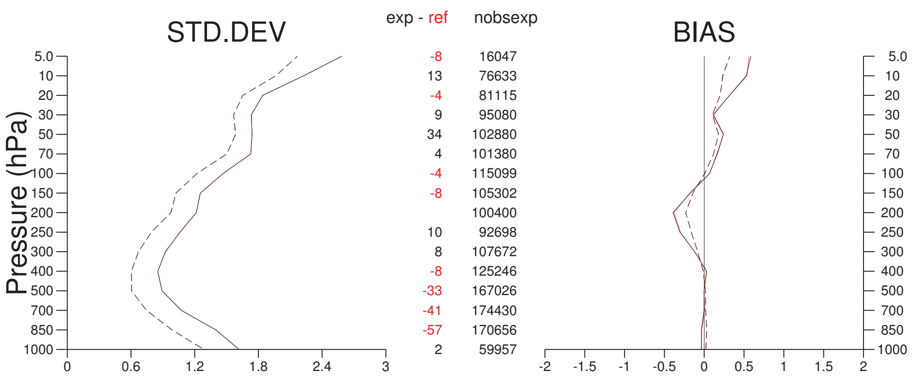


Figure 11: Bias (right) and standard deviation(left) statistics for radiosounding temperature over the whole month. RAW is in red, REC in black. Solid lines represent statistics for observations minus 6-hour forecast and dashed lines represent statistics for observations minus analysis. Column nobsexp is the number of radiosounding observations assimilated for each vertical range, column exp-ref is the difference between REC and RAW.

4.4 Potential changes in observation error correlations

A posteriori diagnostics have been calculated to evaluate the observation errors. The so-called Desroziers method (Desroziers *et al.*, 2005) has been applied to both raw and reconstructed BTs. This technique can provides the full observation error covariance matrix (the R matrix), but only for channels that have been assimilated. The R matrix can be splitted into observation error standard deviations (σ_o) and correlation matrix.

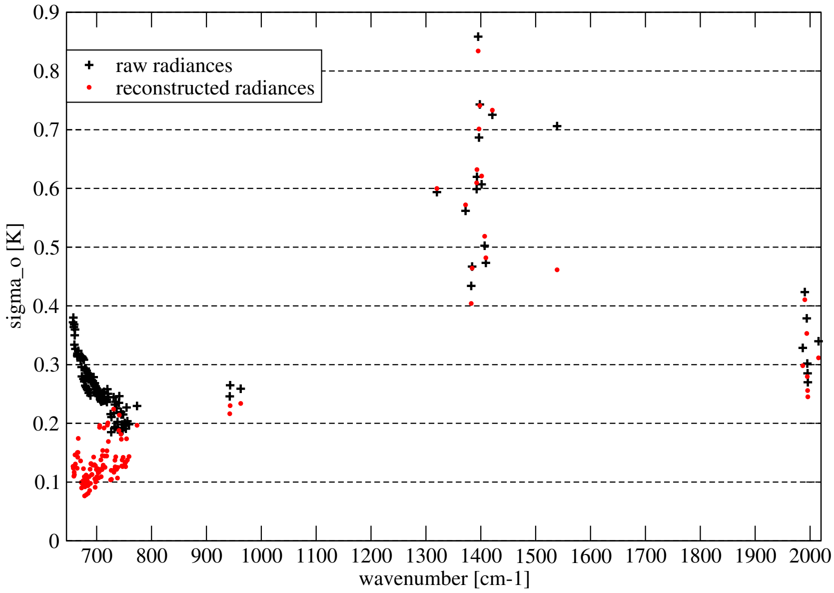
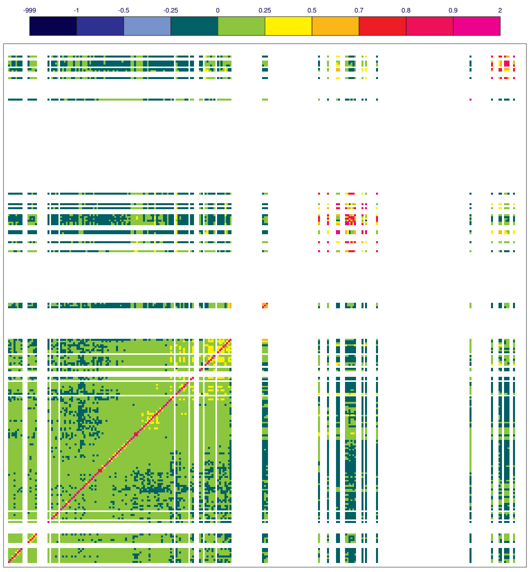
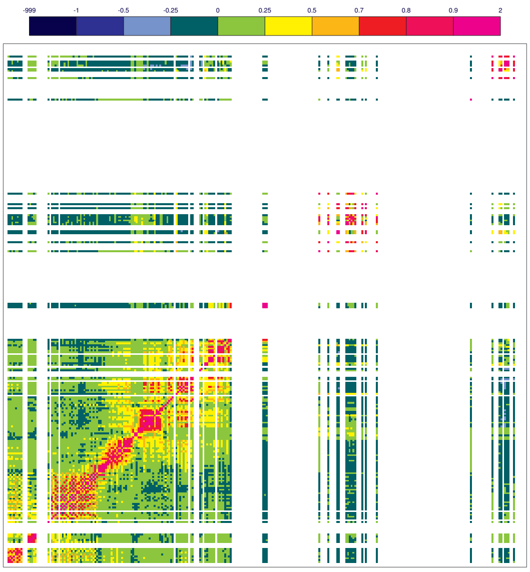


Figure 12: Observation error standard deviations obtained from the a posteriori Desroziers diagnostic

Standard deviations are shown in Figure 12. For those channels in band 2 (water vapour channels), σ_o are rather similar between REC and RAW. σ_o in band 1 are dramatically lower for the reconstructed BTs than for the original BTs. When comparing the correlation part of the R matrices (Figure 13), it turns out that this decrease in standard deviations is linked to an increase of the inter-channel correlations. This is consistent with what is known of the PC analysis technique.



(a) raw radiances



(b) reconstructed radiances

Figure 13: Correlation part of the R matrices obtained from the a posteriori Desroziers diagnostic. On x-axis, wavenumbers increase from left to right; on y-axis, wavenumbers increase from bottom to top. The temperature channels (band 1) are in the left lower corner and the water vapour channels (band 2) are in the top right corner.

5 Conclusion and future plans

A pre-processing chain has been set up in the Météo-France research environment to get Eumetsat BUFR of IASI PC scores and reconstruct radiance spectra from these PC scores. Reconstructed data have been compared with original data and globally showed a good agreement. Nevertheless, some discrepancies occuring either over cloudy areas or over land indicating that some numerical errors may occur either in the coding of the BUFR files or in the reconstruction of the spectra in Météo-France pre-processing. As only one pixel out of 8 is monitored in the Météo-France global model, further studies in the frame of the convective-scale model AROME could be of interest (all pixels are monitored in AROME).

When ingested in the assimilation process of the global model ARPEGE, reconstructed BTs have the same impact on analyses and forecasts as the original BTs. No impact is seen on the bias correction, which means that assimilating reconstruted BTs introduces no bias in the system. The standard deviations of observations minus simulations from model fields are smaller for reconstructed BTs than for the original BTs, which is consistent with the a posteriori Desroziers diagnostic. Especially in band 1, reconstructed BTs are less "noisy" and could be assimilated with smaller σ_o (which means with a larger weight). But it appears that the PCA technique introduces inter-channel correlation in the observation errors. As no inter-channel correlation is used in ARPEGE assimilation for the time being, some non-neutral impact of assimilating reconstructed BTs with the proper R matrix may be found. This study should be performed again when the possibility of using a non-diagonal R matrix is available in ARPEGE.

Results are consistent with those obtained at the Met-Office (F. Smith, pers.comm.).

References

Cayla, F. (2001). L'interféromètre IASI, un nouveau sondeur satellitaire à haute résolution. *La Météorologie 8ème série 32*, 23–39.

Chahine, M. (1974). Remote sounding of cloudy atmospheres. i: The single cloud layer. *J. Atmos. Sci.*.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli (2005). Diagnosis of observation, background and analysis error statistics in observation space. *Quart. J. Roy. Meteor. Soc. 131*.

Guidard, V., N. Fourrié, P. Brousseau, and F. Rabier (2011). Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes. *Quart. J. Roy. Meteor. Soc. 137*, 1975–1987.

Hilton, F., R. Armante, T. August, C. Barnet, A. Bouchard, C. Camy-Peyret, V. Capelle, L. Clarisse, C. Clerbaux, P.-F. Coheur, A. Collard, C. Crevoisier, G. Dufour, D. Edwards, F. Faijan, N. Fourrié, A. Gambacorta, M. Goldberg, V. Guidard, D. Hurtmans, S. Illingworth, N. Jacquinet-Husson, T. Kerzenmacher, D. Klaes, L. Lavanant, G. Masiello, M. Matricardi, A. McNally, S. Newman, E. Pavelin, S. Payan, E. Péquignot, S. Peyridieu, T. Phulpin, J. Remedios, P. Schlüssel, C. Serio, L. Strow, C. Stubenrauch, J. Taylor, D. Tobin, W. Wolf, and D. Zhou (2011). Hyperspectral earth observation from IASI: four years of accomplishments. *BAMS 93*, 347–370.

McNally, A. and P. Wats (2003). A cloud detection algorithm for high-spectral-resolution infrared sounder. *Quart. J. Roy. Meteor. Soc. 129*, 3411–3423.