The assimilation of the IASI long-wave and mid-wave full spectrum using reconstructed radiances

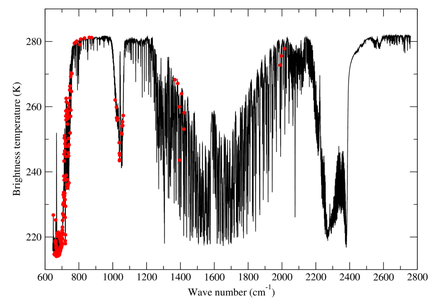
Marco Matricardi

Research Department, ECMWF, Reading, United Kingdom

1. Principal Component Analysis

- The operational use of Infrared Atmospheric Sounder Interferometer (IASI) radiances at the European Centre for Medium Range Weather Forecasts (ECMWF) is currently restricted to a selection of temperature, ozone and water vapour sounding channels in the long-wave and mid-wave regions of the spectrum. In principle, to exploit the full information content of IASI, the number of channels used in the assimilation could be increased to cover the full spectrum. Numerical Weather Prediction (NWP) users are limited to assimilating less than the full IASI spectrum by the prohibitive computational cost.
- Principal Component Analysis (PCA) is a classical statistical method for the efficient encapsulation of information from voluminous data [1]. As such, it has been proposed as a solution to the above problem.
- ECMWF [2] has been at the forefront of the exploitation of PC-based data demonstrating the direct assimilation of IASI Principal Component (PC) scores in a global 4D-Var assimilation system. More recently, the ECMWF focus has shifted from the direct assimilation of PC scores to the assimilation of radiances reconstructed from truncated principal components (hereafter RR). Although at a theoretical level the assimilation of PCs or RRs can be considered equivalent, the latter have the advantage of being able to exploit existing science and infrastructure developed for raw radiance assimilation, in particular cloud and aerosol screening. To this end, the ECMWF 4D-Var data assimilation system has been modified to allow testing the assimilation of a selected number of 400 IASI reconstructed radiances which effectively encapsulate the information content of the full number of IASI channels in the long-wave band-1 (2221 temperature, surface, and ozone sounding channels) and the mid-wave band-2 (3201 water vapour sounding channels). Note that we have excluded short-wave channels for a number of reasons, which include day/night variations in data usability due to non-local thermodynamic equilibrium effects and high instrument noise. Indeed the latter was one of the incentives of initial PCA investigation.

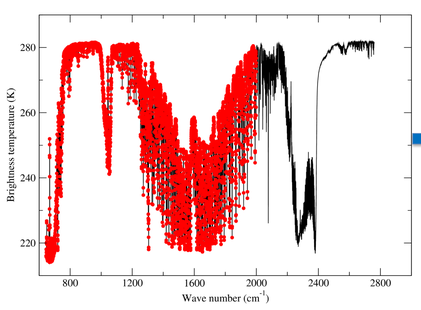
We currently use only ~ 2% of the available IASI channels (the red dots indicate the IASI channels assimilated in the ECMWF operational system (cycle 43r3))



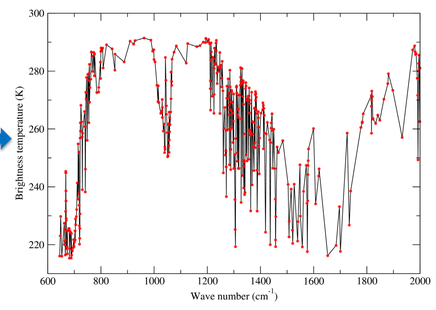
The *m* leading eigenvectors of a covariance matrix (***A***) describing variations in a dataset of synthetic IASI spectra (***r***) are used to calculate *m* truncated PC scores (***p***) and *m* Reconstructed Radiances (***rr)***

The PCA technique has been used to encapsulate most of the information content of 5421 IASI raw radiances into 400 reconstructed radiances.

The red dots indicate the 5421 IASI channels in band-1 and band-2

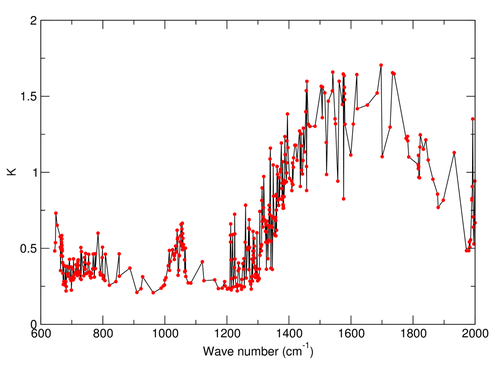


The red dots indicate the 400 reconstructed radiance channels used in the reconstructed radiances assimilation trials

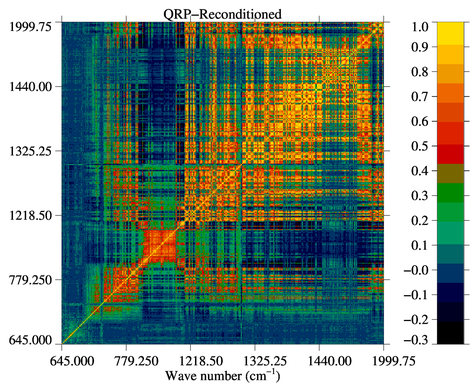


2. Reconstructed radiance assimilation methodology

- Observed IASI spectra are converted to PC scores by projecting the radiance vector on the fixed basis of synthetic eigenvectors utilised in the PC-based fast radiative transfer model used in the assimilation. Note that the synthetic eigenvectors used in this study correspond to cloudy and clear sky signals. Each vector of observed PC scores has length *n*=5421, but crucially we retain only the first *m*=400 of these. In truncating the vector of observed PC scores we make the assimilation highly efficient, while preferentially retaining highest rank PC scores that convey most information about the atmospheric state.
- The truncated PC scores are used to reconstruct a new radiance vector. In principle, the *m* truncated PC scores can be used to reconstruct a radiance vector containing all *n* channels. However, the *n* RRs can only contain as many pieces of information as the *m* PC scores (that is, *n* reconstructed radiances are not equivalent to *n* raw radiances - PCA is not a lossless technique). The radiance reconstruction process reduces the noise in the observations (random instrument noise is spectrally uncorrelated and is thus represented by low rank eigenvectors), yet at the same time introduces significant inter-channel correlations because the matrix used to map PC scores into reconstructed radiances is rank deficient - the rank cannot be larger than *m*, therefore only a maximum of *m* reconstructed radiances may be independent. This means that for a choice of more than *m* RRs the resulting error covariance matrix is rank-deficient and may not be invertible. To use as much as possible of the information contained in the *m* PC scores, we have carried out a selection of the most independent *m* reconstructed radiance channels ensuring that the resulting observation error covariance matrix is not only invertible but also well conditioned.
- The RR system uses a dedicated observation error covariance diagnosed from innovation statistics [3] and the reconstructed radiances are simulated in the 4D-Var by PC-RTTOV [4] - the principal component based version of the RTTOV fast radiative transfer model. In all other respects, the RR assimilation system is identical to the conventional radiance assimilation system.



The observation errors assigned to the RRs



The observation error covariances used for the assimilation of the RRs

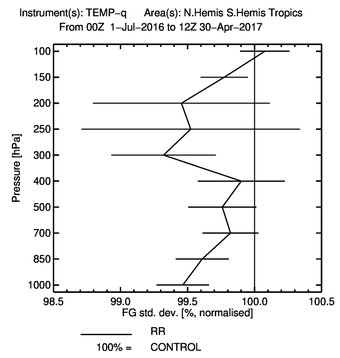
3. Assimilation experiments

- To quantify the performance of the RR assimilation system we have used two assimilation and forecast experiments. The first experiment (hereafter referred to as the CONTROL) uses the ECMWF operational configuration. The second experiment (hereafter referred to as the RR) is identical to CONTROL but assimilates 400 RRs instead of the 191 IASI operational raw radiances. Note that ECMWF assimilates IASI observations from both MetOp-A and MetOp-B platforms.
- Experiments use ECMWF cycle 43r3 which was operational from July 2017 to July 2018 and has 137 vertical levels. All experiments have been run at a reduced horizontal resolution of ~40 km (T511). Because chaotic variability in the quality of individual forecasts can be very large, to achieve statistical significance [5] we have used a very large sample size that consists of 594 forecasts covering 1 July 2016 to 27 April 2017.

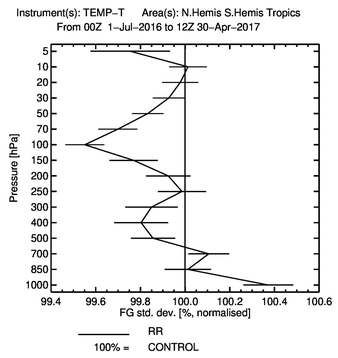
4. Impact on the assimilation: the fit to radiosonde data and the impact on forecasts

- The statistics of the fit to radiosonde temperatures and humidities for the background forecasts are shown in terms of standard deviations with respect to the CONTROL system (values are explicitly normalised to the CONTROL to improve visualisation). Thus, reduced values for the RR system indicate the extent to which the assimilation of reconstructed radiances improves the fit to radiosonde data compared to the operational assimilation.

Humidity

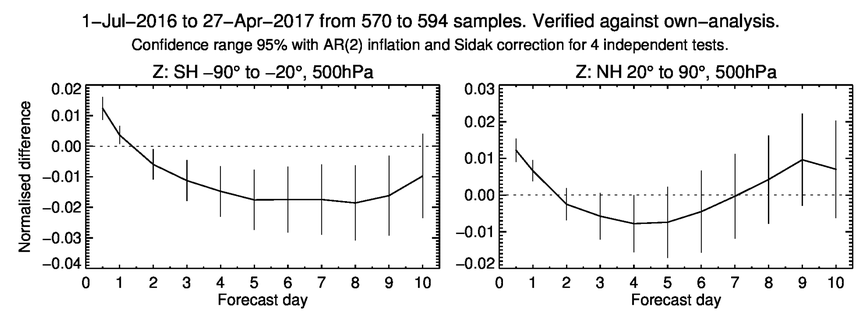


Temperature



The global fit to radiosonde temperature (left panel) and humidity (right panel) data for the background forecasts

- Forecasts have been run from analyses generated by the RR and CONTROL experiments and verified against the experiment's own analysis. Forecasts scores over the 10 month testing period have been computed as the change in the root-mean-square error compared to the CONTROL and the differences normalised by the forecast errors of the CONTROL experiment. Results are presented in terms of 500 hPa geopotential height for the extratropical Northern and Southern Hemispheres. A negative value of the forecast score means that the use of IASI reconstructed radiances improved forecast accuracy compared to the operational system.



Normalised root-mean-square error difference for 500 hPa geopotential

5. Conclusions

- Nine months of assimilation trials show that, overall, background forecasts produced by the RR system fit radiosonde and satellite observations better than background forecasts produced by the operational system, especially for water vapour.
- Medium range forecasts launched from the analyses produced by the assimilation of reconstructed radiances demonstrate a performance advantage over the operational system, especially in the Southern Hemisphere, where 500hPa geopotential scores are improved up to ~ 2 % in the medium-range during the testing period. Note that this improvement is comparable to the typical annual improvement in scores from all developments in data assimilation, modelling and the observing system.
- The short term plan is to migrate the RRs libraries to the latest ECMWF cycle. If assimilation trials are successful, then the medium term plan is to introduce the assimilation of IASI RRs into the ECMWF operational system.

References

[1] Joliffe IT. 2002. Principal Component Analysis. Springer: New York.

[2] Matricardi, M. and McNally, A. P. 2014. The direct assimilation of principal components of IASI spectra in the ECMWF 4D-Var. Q.J.R. Meteorol. Soc., 140: 573-582. doi:10.1002/qj.2156

[3] Desroziers G., Berre L., Chapnik B., Poli, P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. Quarterly Journal of the Royal Meteorological Society, **131**, 3385-3386.

[4] Matricardi, M. 2010. A principal component based version of the RTTOV fast radiative transfer model. Q.J.R. Meteorol. Soc., 136: 1823-1835. doi:10.1002/qj.680

[5] Geer, A. J. 2016. Significance of changes in medium-range forecast scores,Tellus A: Dynamic Meteorology and Oceanography, 68:1, 30229, DOI: 10.3402/tellusa.v68.30229