INTERNATIONAL ATOVS WORKING GROUP

Towards a synergetic approach for the retrieval of atmospheric parameters from passive optical, infrared and microwave measurements

O. AZNAY (1), F. AIRES (1,2), M. PAUL (1) and C. PRIGENT (3)

(1) LMD/CNRS/France - (2) Estellus, France - (3) LERMA/CNRS/France

Introduction

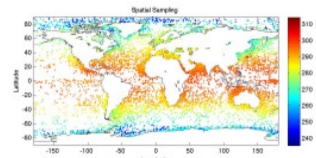
A wealth of Earth satellite observations is now available, covering the entire globe, and providing a large diversity of information over a broad frequency range (UV, visible, infrared and microwave), in order to obtain a global and continuous monitoring of the state of the atmosphere. Space agencies have designed satellite platforms that include instruments from the different regions of the electromagnetic spectrum. In parallel, accurate Radiative Transfer Models (RTM) have been developed to simulate the responses of these multi-spectral observations to atmospheric changes in composition or temperature. However, the retrieval accuracy of key variables such as temperature, water vapour or ozone profiles is still not always satisfying.

In order to develop new approaches to perform satellite data fusion, it is necessary first to understand the basic concepts behind synergy. To illustrate the various types of synergy, it is a good strategy to use simple schematic model. Then, a methodology needs to be put in place to measure the synergy. Since assimilation is the most widely used technique to fusion data, the traditional information content analysis is the first candidate. To test the potential of this method, we apply it to selected atmospheric parameters and wavelength bands under specific instrument geometry for the MetOp-A satellite. This platform provides coincident observations in the visible, GOME-2 (Global Ozone Monitoring Experiment) instrument, in the infrared, IASI (Improved Atmospheric Sounding in the Infrared) instrument, and in the microwaves, AMSU-A (Atmospheric Microwave Sounding Unit - A) and MHS (Microwave Humidy Sounder), with nadir geometries. We concentrate on the major atmospheric parameters, namely temperature, water vapour and ozone profiles for which the selected MetOp-A instruments are particularly sensitive.

The geophysical database

ECMWF operational analyses:

The atmospheric profiles and surface properties from the 6-hourly operational global analyses from the Integrated Forecasting System (IFS) of the European Center for Medium Range Forecasting (ECMWF) are at the origin of the datasets used in our study. In order to run accurate radiative transfer simulations, the following information is kept: the temperature, water vapour and ozone profiles on 43 pressure levels ranging from 1000 to 1 hPa (these levels have been interpolated for the initial 21 levels in order to be used with the RTTOV code) and surface properties such as the temperature. Selecting cloud-free oceanic cases, few millions atmospheric and surface situations are extracted over one-year of data.
Retrieval algorithms cannot handle this huge amount of data. To reduce the size of this geophysical dataset while keeping its spatial and temporal variability a sampling procedure has been used [Aires & Prigent, 2007]

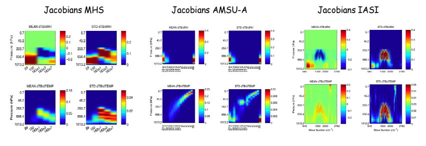


Spatial location of the 10 000 atmospheric situations over ocean used to train and test the retrieval methods. The colour shades represent the temperature (K) at the surface layer for illustration purpose.

The satellite observation databases

The satellite instruments:

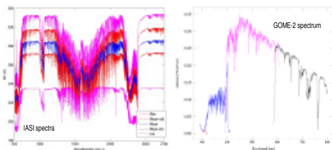
- **AMSU-A** (Advanced Microwave Sounding Unit-A) measures the oxygen band between 50 and 60 GHz, for the retrieval of atmospheric temperature profiles (15 channels).
- **MHS** (Microwave Humidity Sounder) is designed to measure the atmospheric water vapour profile (5 channels).
- **IASI** (Infrared Atmospheric Sounding Interferometer) has been designed to retrieve temperature and water vapour profiles in the troposphere and the lower stratosphere, as well as measure concentrations of ozone, carbon monoxide, methane and other compounds between the wavelengths of 3.2 and 15.5 microns (8461 channels).
- **GOME-2** (Global Ozone Monitoring Experiment) provides vertical profiles or total column amounts for each of the gases (ozone, nitrogen dioxide, sulphur dioxide and other trace gases) and ultraviolet radiation. These profiles are representative of the lowermost 50 kilometres of the Earth's atmosphere (4096 channels).



Simulation of the synthetic database

The radiative transfer simulations:

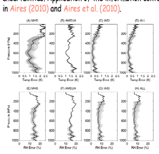
- **RTTOV-8.7** radiative transfer model originally developed at ECMWF [Eyre, 1991] and now supported by EUMETSAT provides rapid simulations of radiances for satellite infrared and microwave radiometers for a given atmospheric state vector. Over ocean, the emissivities are computed by the FASTEM-3 [Deblonde & English, 2001] surface emissivity model.
- **4A** (Automatized Atmospheric Absorption Atlas) is a fast and accurate Radiative Transfer Model for the infrared co-developed by LMD and Noveltis with the support of CNES [Scott & Chédin, 1981]. 4A computates transmittance and radiance, using a comprehensive database (atlases) of monochromatic. optical thicknesses for up to 43 atmospheric molecular species.
- **LBLDOM** code [Dubuisson et al., 1996] allows solving the radiative transfer equation (RTE) from the Discrete Ordinates Method [DOM; Stamnes et al.,1988], at each step of a line-by-line model (LBL). Radiances can be calculated at any atmospheric level in the solar spectral range, with a spectral resolution of about 0.01 cm^{-1}. The atmosphere is assumed to be vertically inhomogeneous and stratified into plane-parallel layers.



Information content: Retrieval and comparison

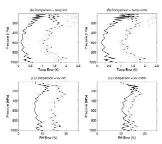
Theoretical aspects:

The most widely used technique exploiting synergy among Earth observation instruments is, without any doubt, the assimilation [Kalnay, 2002]: a wide spectrum of visible, infrared and microwave satellite observations are combined with model forecasts and *in situ* measurements to better characterize and predict the state of the atmosphere, continental surfaces or oceans. In association to the assimilation technique, various tools have been designed to estimate the theoretical quality of retrievals. Since we are interested in this study by the synergy, it is essential to test this approach first. Assimilation and information content analysis share the same theoretical hypothesis: *Gaussian* character of the stochastic variables, linearization around the First Guess (FG), same observation, radiative transfer and *a priori* uncertainties. Application of the information content technique on IR/MW measurements is fully described in Aires (2010) and Aires et al. (2010).



•Strong dependency to the Jacobians => Need to perform statistics on the dataset.

•Estimation too optimistic for IASI retrievals => We cannot measure the synergy.

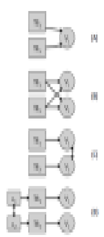


•Retrievals of temperature and water vapour profiles individually are better than simultaneous retrieval of both parameters.

Retrieval algorithms (1)

Theoretical considerations:

Synergy (from the Greek syn-ergo, working together) refers to the phenomenon in which two or more discrete influences or agents acting together create an effect greater than that predicted by knowing only the separate effects of the individual agents.



Four synthetic types of synergy:
additive synergy (**A**), un-mixing synergy (**B**), indirect synergy (**C**) and de-noising synergy (**D**). TB_1 and TB_2 are the satellite observations, ε_1 and ε_2 are the corresponding instrument noise, and V_1 and V_2 are the geophysical variables to retrieve.

Retrieval algorithms (2)

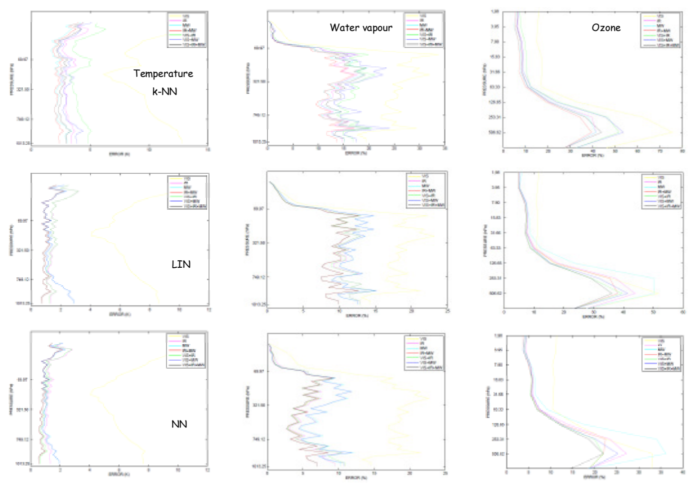
Retrieval approaches:

-In the ***k*-NN** (Nearest Neighbours) retrieval approach, a "reference" dataset, R, is built that includes a number E of situations described by a set of geophysical variables (i.e., the variables to be retrieved) and by a set of the associated satellite observations: $R = \{(GEO^i, TB^i); i = 1, \ldots, E\}$. The Brightness Temperatures (TBs) can be real observations or radiative transfer simulations (the later is chosen in this paper). The k closer situations in R are used to define the inversion. Few approaches can be used: if only the closest situation is taken (k = 1), this scheme is a pattern recognition algorithm. The *k*-NN algorithm is a nonlinear model. It is also a truly multivariate method. However, it should be noted that the method is entirely based on a distance in the satellite observation space. This distance gives the same weight to each of the TB inputs and no information content analysis on the GEO is used. If some non-pertinent channels are included in the TB-space, they will not add any useful information and even worse, they can contaminate the actual pertinent information.

-Multiple **LIN**ear regression (LIN) is a very simple and classical technique. For more details, details to present this method, see any elementary statistical textbook. We just mention here that the LIN model is truly multivariate. Furthermore, contrarily to the *k*-NN approach, only the pertinent information is used from the inputs for the retrieval of a particular output. This means that meaningless information will not pollute the retrieval. By definition, and contrarily to *k*-NN, the LIN approach is not nonlinear and can suffer from outliers, saturations, interactions between inputs, and any nonlinear behaviour. This technique is often used as a reference test for the NN method presented next.

- **NN** techniques have proved very successful in developing computationally efficient algorithms for remote sensing applications. The Multi-Layered Perceptron (MLP) model [Rumelhart et al., 1986] is selected here. It is a nonlinear mapping model: given an input TB, it provides an output *f* in a nonlinear way. The MLP model is defined by the number of input neurons (i.e., the size of the inputs, number of channels), the number of outputs (i.e., the size of the geophysical variables to retrieve = 3 x 43 = 129) and the number of neurons in the hidden layers that control the complexity of the model.

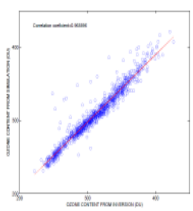
Retrieval results (1): the different instrument combinations with the different methods



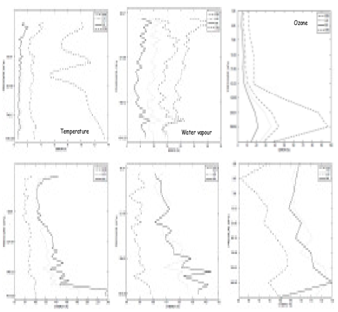
Retrieval results (2): Restitution of the ozone content

Correlation coefficients between the ozone integrated content obtained from simulations and the ozone integrated content obtained from inversions.

RMS error statistics (in %) for the retrieval of the ozone integrated content obtained from simulations and the ozone integrated content obtained from inversions.



Retrieval results (3): A measure of the synergy



Conclusion (1)

***k*-NN method:**

The synergy cannot be said to be an improvement for retrieval of temperature, water vapour or ozone profile.

LIN method:

The input information being additive in a linear model, the more information available, the best the retrieval. As a consequence, the combination of the four instruments provides a synergy effect, improving the retrieval of the **temperature** by up to 0.5 K near the surface, and never degrading the retrieval of the best instrument.
For **water vapour** profile retrieval, the synergy impact is very positive, the RMS error is lower at every atmospheric layer than any individual-instrument retrieval. The benefit can be large, especially near the surface where the RMS error can decrease from ~13% to ~10% (i.e., a ~25% decrease of the error).
Finally for the **ozone** profile retrieval, the benefit of the synergy can be large, especially at low altitude where the RMS error can decrease from ~50% to ~35% (i.e., a ~30% decrease of the error). The retrieval of the **ozone integrated content** is consistent with ozone profile retrieval.

NN method:

Combining the four instruments improves considerably the retrieval of the **temperature**, especially near the surface where the RMS error decreases from 1.5 K (IASI retrieval) to about 0.5 K for the combined configuration. Like for the LIN method, the more information available in the inputs, the best the retrieval.
The synergy impact is very positive for the retrieval of the **water vapour** profile, the RMS error is lower at every atmospheric layer than any individual instrument retrieval. The synergy from the four instruments is always positive with an important decrease of the RMS error (from 10 to 7% RMS error in the lower layers).
It can be seen that the merging of all the information is the best retrieval of the ozone profile for most of the atmospheric layers. Again, the benefit of the synergy can be large, especially at low altitude where the RMS error can decrease from ~37% to ~22% (i.e., a ~40% decrease of the error).

Conclusion (2)

Comparing methods:

The **NN** and **LIN** outperform the ***k*-NN** methods, with very interesting levels of accuracy especially when compared to GOME-2 errors budget. This study clearly shows that the **NN** and **LIN** make the difference with the other method when the relationship from the satellite observations to the geophysical parameter to retrieve (i.e. the water vapour or the ozone) is complex and nonlinear; the impact is less important but still exist for simpler problems such as the retrieval of the temperature.

Measure of the synergy:

For the **temperature** retrieval, the *k*-NN algorithm is close to 100% with some gain in the lower atmosphere, up to 800 hPa, but with a negative effect on higher levels. The LIN method benefits from the synergy for all the atmospheric layers, in particular close to the surface. The impact can be important with an improvement of the retrieval statistics by more than 80% next to the surface and close to 20% in the middle troposphere. The NN inversion benefits significantly from the synergy: the retrieval errors can be reduced by a factor 2 at the surface.
For the **water vapour** profile retrieval, the *k*-NN method is not optimal to merge information from the various captors: The impact of using simultaneously the four instruments (GOME-2, MHS, AMSUA and IASI) is always negative, with an increase of errors up to 20%. The LIN method has a synergy factor close to 100% (meaning that no synergy is observed for the retrieval of water vapour) for higher atmospheric layers but the impact is quite positive for layers lower than 300 hPa. The NN method benefits from the synergy: The synergy factor reaches 140%, meaning that the errors are decreased by 40% when the four instruments are used together.
For the **ozone** retrieval, the LIN and NN methods appear to be quite optimal for some layers to merge information from the various captors while for the *k*-NN method the impact of using simultaneously the four instruments is mostly negative, with an increase of errors up to 20%. The LIN and NN methods have a synergy factor close to 105% meaning that a small synergy is observed for stratospheric layers (pressure ≤ 120 hPa) but the impact is quite positive for higher pressures. Both methods benefit from the synergy: The synergy factor reaches 125% for NN method, meaning that the errors, for these layers, are decreased by 25% when the four instruments are used together.

References

Aires, F., and C. Prigent (**2007**), Sampling Techniques in High-Dimensional Spaces for Satellite Remote Sensing Databases Generation, *J. Geophys. Res.*, 112, D20301, doi:10.1029/2007JD008391.

Aires, F. (**2010**), Measure and exploitation of multi-sensor and multi-wavelength synergy for remote sensing - Part I: Theoretical considerations, *J. Geophys. Res.*, submitted.

Aires, F., M. Paul, C. Prigent, O. Aznay, B. Rommen, and M. Bouvet (**2010**), Measure and exploitation of multi-sensor and multi-wavelength synergy for remote sensing - Part II: Application to the retrieval of atmospheric temperature and water vapour from MetOp, *J. Geophys. Res.*, submitted.

Deblonde, G. and S. J. English (**2001**), Evaluation of the FASTEM-2 fast microwave oceanic surface emissivity model, *Tech. Proc. ITSCXI Budapest*, 2026 Sept 2000, pp. 67-78.

Dubuisson P., J.C. Buriez and Y. Fouquart (**1996**), High Spectral Resolution Solar Radiative Transfer in Absorbing and Scattering media, application to the satellite simulation, *J. Quant. Spectros. Radiat. Transfer*, 55, pp. 103-126.

Eyre, J.R. (**1991**), A fast radiative transfer model for satellite sounding systems, *ECMWF Research Dept. Tech. Memo. 176* (available from the librarian at ECMWF).

Kalnay, E. (**2002**), Atmospheric Modelling, Data Assimilation and Predictability, *Cambridge University Press*, 364 pp.

Rumelhart, D.E., G.E. Hinton, and R.J. Williams (**1986**), Learning Internal Representations by Error Propagation, in Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. I, Foundations, edited by D.E. Rumelhart, J.L. McClelland, and the PDP Research Group, pp. 318–362, *MIT Press, Cambridge*.

Scott, N.A. and A. Chédin (**1981**), A fast line-by-line method for atmospheric absorption computations: The Automatized Atmospheric Absorption Atlas, *J. Appl. Meteorol.*, 20, 556564.

Stamnes, K., S. Tsay, W. Wiscombe, and K. Jayaweera (**1988**), Numerically stable algorithm for discrete ordinate method radiative transfer in multiple scattering and emitting layered media, *Appl. Opt.*, 27, 25022509.

ITSC-XVII Conference: Monterey, CA, USA, 14-20 April 2010