Satellite Data Assimilation in Numerical Weather Prediction: an Overview

Jean-Noël Thépaut

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire, RG2 9AX, U. K.

1. Introduction

Over the last few years, satellite data have progressively become the predominant source of information assimilated in Numerical Weather Prediction (NWP) models. This has been made possible thanks to a substantial enhancement of the remote sensing instruments measuring various atmospheric quantities but also largely to the improvements in data assimilation techniques to better exploit the information contained in such data. The advantage of satellite data is that they provide a uniform spatial and temporal coverage of the atmosphere. This advantage is however balanced by a general poor vertical resolution of the instruments currently used, and the difficulty to handle clouds, precipitations and surface contributions to the information content of the data. The future improvements of NWP models and a better handling of new observing techniques (radio-occultation, passive limb soundings, active sensors) in data assimilation schemes may overcome some of these limitations.

This paper briefly presents the Satellite Observing System (section 2). Different ways of observing the Earth from space are described as well as how geophysical information is inferred from these measurements into NWP models (section 3). Section 4 stresses how satellite data have become an essential part of current data assimilation schemes. Section 5 finally describes the current issues and the foreseen evolution of the observing system from space and the opportunities as well as challenges this will entail for operational NWP data assimilation systems.

2. The Satellite Observing System

The Space-based Global Observing System complements the in-situ Observing System to provide what is commonly known as the World Weather Watch’s Global Observing System. The main providers of Earth-observation satellite systems and space-based observations of the atmosphere for NWP centres are the American (NASA and NOAA), European (ESA and EUMETSAT) and Japanese (JAXA and JMA) space agencies. Other Earth-observation satellite systems are operated by the Russian Federation, People’s Republic of China, etc…

Research and Development (R&D) Space Agencies (NASA, ESA, JAXA) usually promote demonstration missions, with innovative technologies, thus paving the way for future long-term operational missions. The primary goal of R&D Space Agencies is in principle not the delivery of Near Real Time products to the community. However, R&D satellite missions prove to be crucial to better characterize diverse features of the model (UARS, POLDER), develop new methodologies in view of assimilation of future operational instruments (AIRS on AQUA, TRMM), or sometimes improve simply the quality of the operational NWP assimilation system when data are of good quality and available on time (ERS-2, QuikScat, AIRS, ENVISAT). A special care has to be taken when assimilating R&D satellite products operationally. In close collaboration with the different space agencies involved, ECMWF has benefited from an excellent access to

R&D data to enable a quick feedback between the data producers and the user community, these CAL/VAL exercises being obviously beneficial to both sides.

Operational Spaces Agencies (NOAA, EUMETSAT, JMA) operate instruments that inherit from demonstration missions. Operational systems ensure a stabilised long-life mission technology (e.g. the HIRS instrument onboard NOAA satellites will have lasted some 30 years), which eases the investment decisions at the NWP community end. Operational missions moreover ensure robustness in the processing chain and time delivery to the end-users in agreement with their requirements. Today, they constitute the backbone of the Global Observing System and provide the major part of the data currently assimilated in NWP.

Operational agencies ensure the long-term continuity of the operational systems in polar as well as geostationary orbits. Both ways of observing the Earth/Atmosphere are very complementary. Geostationary platforms (GEOS), located at 36 000 km in an equatorial plan, orbit the Earth with the same angular velocity as the Earth and therefore provide an almost continuous view (repetition time of down to a few minutes) of the same part of the Earth. The high temporal resolution of the GEOS makes them essentially suitable for nowcasting applications, but also for NWP four-dimensional data assimilation systems through the provision of Atmospheric Motion Winds derived from cloud tracking or sequences of radiance data. The orbit geometry of GEOS makes them unable to observe polar regions. Figure 1 (top) displays the current constellation of Geostationary satellites currently assimilated at ECMWF. Low Earth Orbiting satellites

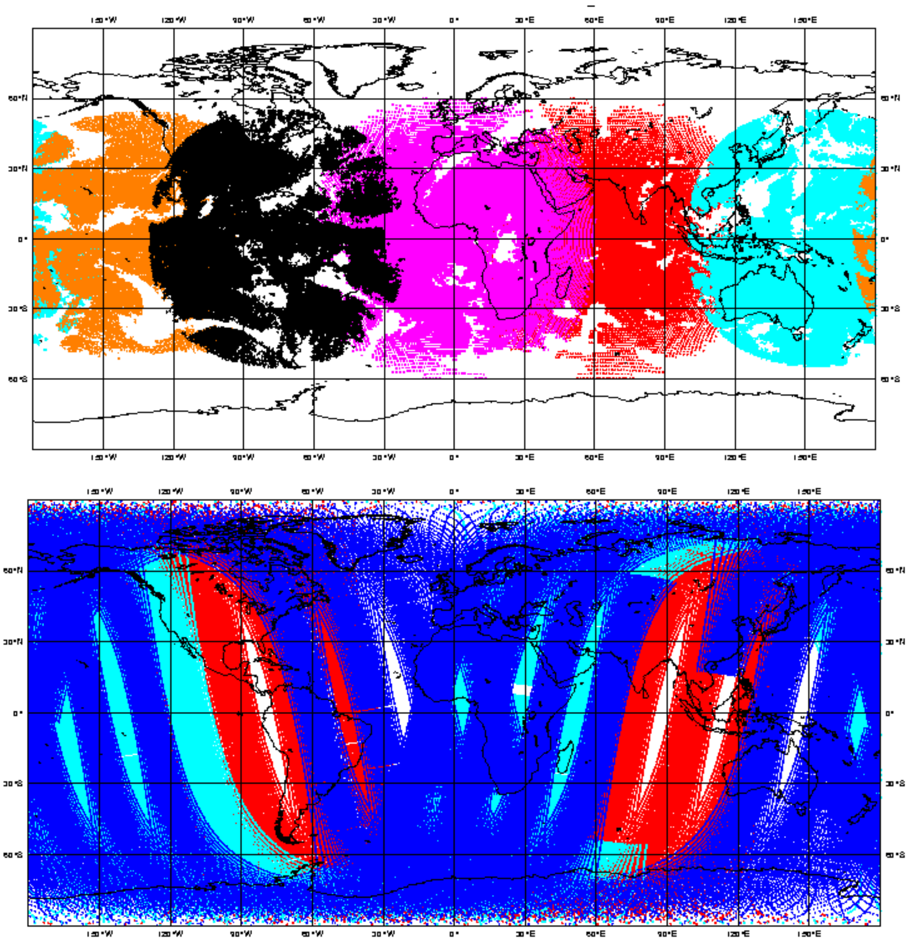

Figure 1: Typical data coverage provided by the Geostationary constellation (top): GOES-W/E (orange/black), Meteosat-7/5 (pink/red) and GMS-5/GOES-9 (cyan). Bottom plot displays the LEO constellation from the NOAA satellites (NOAA-15 in red, NOAA-16 in cyan, NOAA-17 in blue).

(LEOs), at least the operational ones, orbit the Earth at around 800 km with a repetition time over the pole of about 100 mn. Being closer to the atmosphere than the GEOS, these satellites are more suitable to sound the atmosphere (specially in the microwave spectrum). Figure 1 (bottom) also displays the constellation of the operational AMSU-A instruments on board NOAA satellites. In both cases (GEOs and LEOs), a constellation of satellites is required to provide an adequate global coverage. This is currently achieved by the current respective operational satellite constellations.

3. Earth Observing techniques from Space

What is specific to satellite observations? what are they actually measuring?

It is important to realize that the measured quantities by satellite instruments do not relate directly to geophysical quantities. The conversion of measured quantities into geophysical quantities is an inverse problem that data assimilation schemes try to solve in an optimal way. As such, satellite instruments do not measure temperature, do not measure humidity, do not measure wind. Satellite instruments (active and passive) measure essentially the radiance L that reaches the top of the atmosphere at a given frequency **ν**. The radiance is related to geophysical parameters through the radiative transfer equation. In short, the radiative transfer equation can be summarized as follows:

$$L(\nu) = \int_0^\infty B(\nu, T(z)) \left[\frac{d\tau(\nu)}{dz} \right] dz$$

+ (surface emission)

+(surface reflection)

+(surface scattering)

+(cloud/rain contribution) (1)

where B(**ν**, T(Z)) is the Planck radiance for a scene temperature T at altitude Z, and τ(**ν**) the altitude Z to space transmittance. Note that Eq. 1 is a particular case of the generalized direct problem mapping the atmospheric model X_b to a given observation Y, $Y = H(X_b)$, the observation operator H being in this specific case the radiative transfer equation.

3.1. Different Ways of Sensing the Earth/Atmosphere

Eyre (2000) provides an excellent overview of the different instrument technologies commonly used to observe the atmosphere from space, and a brief summary is given here. By selecting radiation at different frequencies (or channels), a satellite instrument can provide information on a range of geophysical variables (upper air temperature, surface parameters, clouds,…). A distinction has to be made between passive and active instruments. Passive instruments sense radiation emitted by the surface and/or atmosphere (or the solar radiation reflected by it). Active instruments emit radiation and measure how much of it is reflected or back-scattered by the surface and/or atmosphere.

In general the channels currently used for NWP applications maybe considered as one of 3 different types.

3.1.1. Atmospheric sounding channels from passive instruments

These channels are located in parts of the infrared and microwave spectrum for which the main contribution to the measured radiance is described by the first term of the right hand side of Eq 1:

$$L(\nu) = \int_0^\infty B(\nu, T(z)) \left[\frac{d\tau(\nu)}{dz} \right] dz \tag{2}$$

These channels avoid frequencies for which surface radiation or cloud contribution are important. These channels are primarily used to obtain information about atmospheric temperature and humidity (Atmospheric sounding channels from the HIRS (High resolution Infrared Sounder) and AMSU (Advanced Microwave Sounding Unit) on board NOAA satellites fall into this category).

3.1.2. *Surface sensing channels from passive instruments*

These channels, called "imaging" channels, are located in atmospheric "window" regions of the infra-red and microwave spectrum at frequencies where there is very little interaction with the atmosphere and the main contribution to the measured radiance in this case is:

$$L(\nu) = \textbf{Surface emission } [T_{surf}, \varepsilon(\text{wind},\ldots)] \tag{3}$$

where **T**surf is the surface temperature and **ε** the surface emissivity. These channels are primarily used to obtain information on surface temperature and quantities that influence the surface emissivity such as wind (through the roughness over sea) and vegetation (land). They can also be used to obtain information on cloud top (in the Infrared) and rain (in the microwave). In addition, sequences of Infrared images from geostationary satellites can be used to track the cloud movements and indirectly derive wind information.

3.1.3. *Surface sensing channels from active instruments*

These instruments (e.g. scatterometer) emit microwave radiation towards the surface in the atmospheric window parts of the spectrum such that radiance scattered back from the surface is:

$$L(\nu) = \textbf{Surface scattering } [\varepsilon(\text{wind})] \tag{4}$$

These instruments provide information on ocean winds. Some similar-class active instruments such as altimeters and SARS (Synthetic Aperture Radars) provide information on wave height and spectra.

3.1.4. *Other technologies*

Active instruments operating in the visible (Lidars) or the microwave (radars) can also analyse the signal backscattered from atmospheric targets such as molecules, aerosols, water droplets or ice particles. Their penetration capability allow the derivation of information on cloud base, cloud top, wind profiles (Lidars) or cloud and rain profiles (radars).

Radio-occulation technique using GPS (Global Positioning System) is another novel way of extracting atmospheric information. These techniques exploit an opportunity that the GPS constellation (originally designed for other applications) already exists. GPS receivers (such as the GRAS instrument on board METOP) measure the Doppler shift of a GPS signal refracted along the atmospheric limb path. This refraction is proportional to (among other parameters) the density of the atmosphere, and therefore indirectly to temperature and humidity profiles. Provided a sufficient number of receivers are installed on LEOs, this technique could offer high vertical resolution (balanced by a somewhat coarse horizontal resolution of ~200-300 km), self-calibrated and “all weather” observations of atmospheric temperature (and possibly humidity).

3.2. Weighting Functions

Let us consider the simple case of a channel for which the primary absorber is a well mixed gas (oxygen or carbon dioxide). It can be shown that the measured radiance is in this case a weighted average of the atmospheric temperature profile, that is Eq. 1 reduces to:

$$L(\nu) = \int_{0}^{\infty} B(\nu, T(z))K(z)dz \tag{5}$$

where . $K(z) = d\tau / dz$

The function K(Z) that defines this vertical average is known as a weighting function. Figure 2 a) and b) represent two idealized weighting functions. If the weighting function of a channel is a delta-function, the measured radiance is uniquely sensitive to the temperature at a single level in the atmosphere. Conversely, if the weighting function is constant with Z throughout the atmosphere, the measured radiance is sensitive to the mean temperature of the atmospheric profile.

In reality, the shape of the real atmospheric weighting functions are somewhere in between these two idealized cases (see Figure 3 illustrating the AMSU-A temperature weighting functions). As we can see, they are fairly broad, i. e. the associated radiances sense very broad atmospheric layers. A reasonable vertical sampling of the atmosphere by satellite radiances therefore comes from an appropriate selection of channels, with varying absorption strengths.

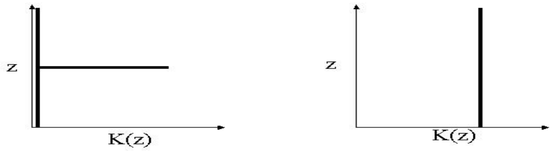

Figure 2. Two idealized weighting functions: The delta function corresponds to a sensitivity of the radiance to a single atmospheric level (a). The vertically constant weighting function corresponds to a sensitivity of the radiance to the mean temperature between the surface and the top of the atmosphere (b).

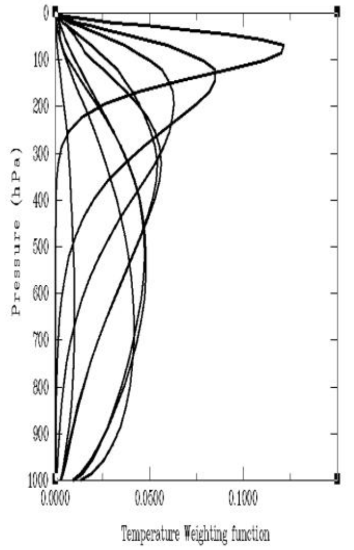

Figure 3: temperature weighting functions of the AMSU-A instrument on board NOAA satellites

3.3. The Inverse Problem

The problem of extracting the atmospheric temperature profile from a set of measured radiances is called the retrieval or inverse problem. Unfortunately, with a finite number of channels and with weighting functions that are generally quite broad, the inverse problem is generally ill-posed (i. e. an infinite number of different temperature profiles could give the same measured radiance). If one wants to utilize satellite radiances to determine the initial conditions of a NWP model, the role of data assimilation is to solve this ill-posed problem. Any technique requires the use of a prior information, the quality of which will drive the accuracy of the final retrieved product. Let us write the (1D) inversion equation simply as:

$$X_a = X_b + \frac{\mathbf{BH}^t}{\mathbf{HBH}^t + \mathbf{R}}\left[Y - H(X_b)\right] \tag{6}$$

where,

X_b represents the atmospheric background, or prior information,

Y represents the radiance observation,

B and **R** represent the associated error covariance matrices

X_a represents the final analysis.

It is clear from Eq. 6 that it is through the convolution of **B** and **H** (jacobian of *H*, proportional to its weighting function) that a given measurement information will be distributed in space and among different geophysical quantities defining the atmospheric state. In particular, since satellite radiances sense very broad layers, the vertical propagation of an observed increment is left to **B**. The design of **B** is therefore crucial for a proper assimilation of satellite radiances. On can easily see that the complexity of the problem increases when radiance information has to be further distributed through temperature, moisture and other atmospheric quantities (ozone, CO2, clouds, rain) for which the direct modeling is difficult and the error statistics (e.g. the associated **B**) poorly known. Additional complexity occurs when *H* incorporates the horizontal/vertical (3D-Var) and time (4D-Var) dimension.

Different options have been commonly used to use satellite data in NWP data assimilation and solve Eq. 6.

- *Assimilation of retrieved products from space agencies or research institutes*

These retrieved products have the advantage of being of the same nature as meteorological variables. The principle of the retrieval can vary from centre to centre but it is generally based on a statistical relationship that predicts atmospheric temperature/humidity from measured radiances. One limitation of this approach for NWP is that it is limited by the statistical characteristics of the training dataset and may miss some extreme but important atmospheric features that are statistically rare in the training sample. Another risk is that the retrieved product will suffer from a poorer prior estimate of the atmosphere than that from an NWP model. Last but not least, the error characteristics of the retrieved products are not easy to assign when the inputs and the chain of processing the product (cloud clearing, limb adjustment, surface corrections,…) are hardly known by the NWP centre.

- *Locally produced or "1D-Var" retrievals*

The retrievals are produced by the NWP centre prior to the main assimilation suite, using background information from a short-range forecast (typically 6 hour). The retrieval is the outcome of an optimal estimation (minimizing for example a cost function or solving the standard optimum interpolation equation) adjusting atmospheric profiles to background atmospheric profiles and measured radiances. In that case, prior information is generally very accurate and contains information about important atmospheric phenomena (such as fronts, tropopause folding,…). In principle, the error characteristics (covariances) of the prior information and resulting retrieval are better known. This ingredient is vital for the subsequent assimilation process. However, the error characteristics of the retrieval may remain complicated due to its correlation with the forecast background that is in fact used twice in the subsequent assimilation.

- *Direct assimilation of radiances*

Variational techniques such as 3D-Var or 4D-Var (Rabier et al. 1998) allow the direct assimilation of radiance observations and therefore avoid the need for an explicit retrieval step. The retrieval step is essentially incorporated within the main analysis by finding the model variables that minimize a cost function measuring the departure between the analysed state and both the background and available observations. In this case, the forecast background still provides the prior information to supplement the radiances. However, it is not used twice (as it is in a 1D-Var preprocessor context) and this avoids the problem of assimilating retrievals with complicated error structures. Furthermore, the inversion is further constrained by the simultaneous assimilation of other observations. Note in particular that in 4D-Var, the adjustments forced by radiances at different times of the assimilation window will be consistent with the forecast model physics and dynamics. Last, the characterization of observational errors in radiance space is much easier. In practice, the approach adopted by NWP centres is pragmatic and observations are used in the space where errors are easier to characterize. As it happens, the “model-to-satellite” approach tends to become the rule, but exceptions exist. For example, atmospheric motion winds derived from cloud tracking are assimilated in all NWP systems because the direct assimilation of cloud information is not mature enough. Stoffelen (2000) also explains the advantages of assimilating ambiguous wind products versus raw signals from scatterometers, claiming that observational errors are better characterized in the wind space (gaussian distribution) than in the sigma-naught space (skewed distribution). Last, the assimilation of ozone information from satellites is preferred in “retrieval” space due to currently poor fast radiative transfer modeling performance in the UV domain.

4. The importance of satellite data in Numerical Weather Prediction

4.1. Current Usage of Satellite Data

In terms of number of observations, satellite data dominate by far the volume currently ingested by NWP data assimilation systems. So far the main provider of temperature and humidity sounding data has been the NOAA-series (TOVS and ATOVS) of polar-orbiting satellites. Currently data from three NOAA satellites (15, 16 and 17), each with three main instruments (HIRS in the infrared, AMSU-A and AMSU-B in the microwave) are used in ECMWFs operational system (see Table 1). The humidity-sensitive microwave instruments (SSMI) on the US Navy’s DMSP-series of polar-orbiting satellites also provide an important source of data for assimilation. Data from three DMSP satellites (13, 14 and 15) is currently used operationally. The European METEOSAT and American GOES geostationary satellites provide frequent full-disc radiance data within about 50º North and South of the Equator. These data are predominantly sensitive to the upper tropospheric humidity. Data is currently used from METEOSAT-7 (at 0ºW), METEOSAT-5 (65ºE) except during eclipse season, and from GOES-9, GOES-12 and 10 (at 155ºE, 75ºW and 135ºW, respectively). Accurate near-surface wind data over the oceans is provided by the scatterometer instruments (active microwave) on the European ERS-satellite (with a very limited coverage since summer 2003), and by the American QuikSCAT. Wind data from ERS were used operationally from 1996 to 2002. Currently (and in the past 12 months), QuikSCAT data are assimilated. Last, ozone data are assimilated from the SBUV instrument onboard NOAA-16 and from the MIPAS instrument onboard the European ENVISAT instrument. Data counts of the number of processed and actively used data from each of the above-mentioned satellite systems are given in Table 1, for 20031028-00 UTC. It is worth noticing that the total number of actively used data from these space-based observing systems (3,160,000), now far exceeds the number of used data from the terrestrial observing systems (about 300,000).

Table 1 List of satellite data (excluding ozone data and cloud-drift winds) used in ECMWF operational system, for the date 20031028-00 UTC (12 hour window).

Space craft	Instrument	Total number processed	Number of used data	Measurement
METEOSAT-5	Imager	182,519		water-vapour radiances, sensitive to upper tropospheric humidity
METEOSAT-7		273,498	26,542	
GOES-9		905,376	37,319	
GOES-10		546,256	25,333	
GOES-12		538,132	18,359	
NOAA-16	HIRS	1,683,020	78,804	infrared, temperature and humidity sounding
NOAA-17		1,433,436	66,949	
NOAA-16	AMSU-B	402,450	29,217	microwave humidity sounding
NOAA-17		402,790	29,799	
NOAA-15	AMSU-A	2,046,460	173,216	microwave radiances, temperature sounding in stratosphere and upper/mid troposphere
NOAA-16		2,380,110	188,935	
NOAA-17		2,365,200	213,424	
AQUA		2,427,030	153,380	
DMSP-13	SSMI	89,579	40,901	microwave, tropospheric humidity
DMSP-14		86,394	33,166	
DMSP-15		88,403	39,452	
QuikSCAT	Seawinds	203,272	111,800	Near-surface winds over ocean
AQUA	AIRS	52,019,901	1,893,169	infrared temperature and humidity sounding
Total		68,073,826	3,159,765	

The rapid increase in the number of data used at ECMWF, per assimilation cycle is shown in Figure 4. The increase has been a factor of ten, compared to 1997, when a 6-hourly 3D-Var assimilation system was operational. The increase towards the end of year 2000 was due to the change from 6-hourly to 12-hourly cycling of 4D-Var. The increases in 2002 mainly reflect enhanced use of ATOVS radiances, the introduction of QuikSCAT winds and water-vapour radiances from METEOSAT. The increase (red bar) in January 2003 is due to the introduction of a third DMSP satellite, a third NOAA satellite, and the use of water-vapour radiances from the two GOES satellites. Last, the most recent increase (by a factor of 3) is due to the introduction of the advanced infrared sounder AIRS from which hundreds of channels are now assimilated.

Obviously, the increased data count due to the more aggressive use of satellite observations is only a very partial indication of the importance of these data for NWP applications (although they clearly drive the computer requirements in terms of observation handling). The next two subsections will illustrate the meteorological impact of satellite observations in modern NWP data assimilation systems.

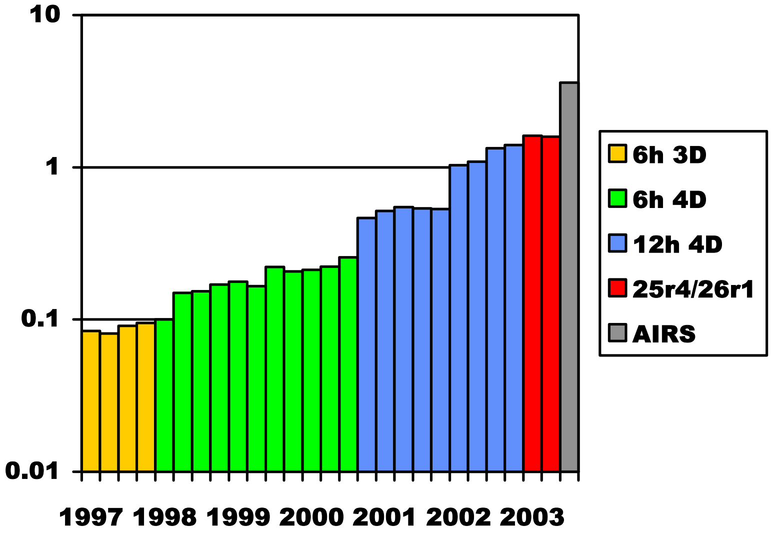

Figure 4: Evolution in the number of observations actively used at ECMWF per assimilation cycle, from 1997 until October 2003.

4.2. Information Content of Satellite Observations

Different ways exist to assess the information content of a single observation or a given Observing System. A popular method consists in evaluating a measure called the Degrees of Freedom for Signal (*DFS*) that can be defined as:

$$DFS = tr(\mathbf{I} - \mathbf{AB}^{-1})$$

or

$$DFS = n - \sum_{\lambda \in \sigma(\mathbf{AB}^{-1})} \lambda$$

where **I** represents the identity matrix, **A** and **B** the background and analysis error covariance matrices, ***λs*** the eigen values of the product $\mathbf{AB}^{-1}$ and n the number of degrees of freedom of the system. *DFS*, the difference between the total number of degrees of freedom and the sum of the eigenvalues, is a measure of the number of degrees of freedom for which the error variance is significantly reduced. Small values for ***λ*** correspond to well observed directions ("degree of freedom for signal"). Eigen values close to one correspond to "unobserved" directions where the error variance is not significantly reduced (i. e. degrees of freedom for noise). Fisher (2003) proposed a method for estimating the degrees of freedom for signal resulting from the assimilation of various observations in the current ECMWF operational 4D-Var system analysis system. The main outcome of the study is displayed in Figure 5, which presents the DFS of the baseline system in which all the observations are assimilated, as well as when individual observation types are removed from the baseline. The most striking feature is the reduction of DFS when ATOVS data (AMSU-A and HIRS) are withdrawn from the assimilation. As pointed out by Fisher (2003), part of the signal maybe a consequence of the presence of the skin temperature in the control variable of 4D-Var for which background errors are artificially large (5K) compared to the accuracy of the AMSU-A instruments. Nevertheless, even if one removes this effect, the contribution of ATOVS data in terms of information content is roughly twice that from TEMP and PILOT data (not shown). It is also interesting to note that TEMP/PILOT observation types are the second contributors, well above Atmospheric Motion Vectors (SATOBs) or Radiances (GEOS) from geostationary satellites.

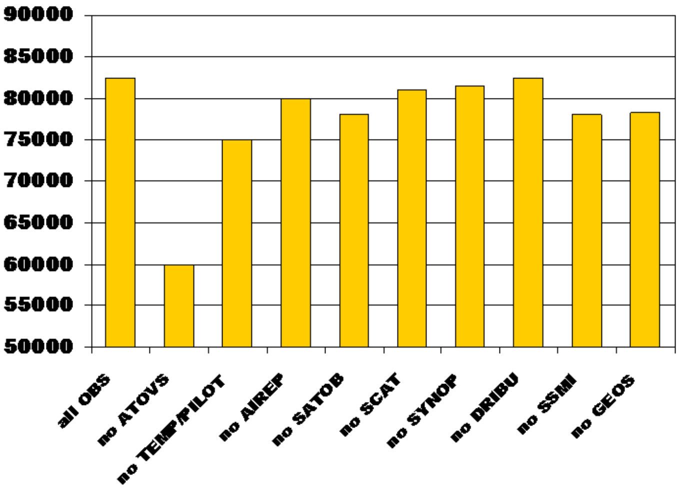

Figure 5: Information content of each individual observation type in the ECMWF assimilation system as represented by the DFS estimation for each individual denial experiment.

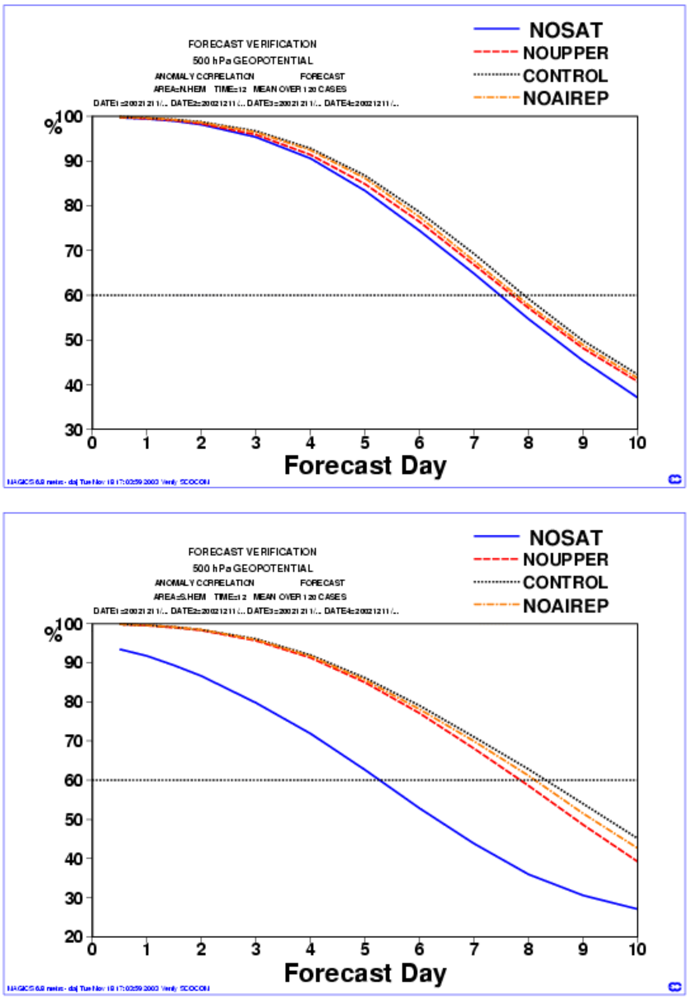

Figure 6. 500 hPa height anomaly correlations for the extratropical northern (top) and southern (bottom) hemispheres averaged over 120-day forecasts from various assimilation experiments (see text for details)

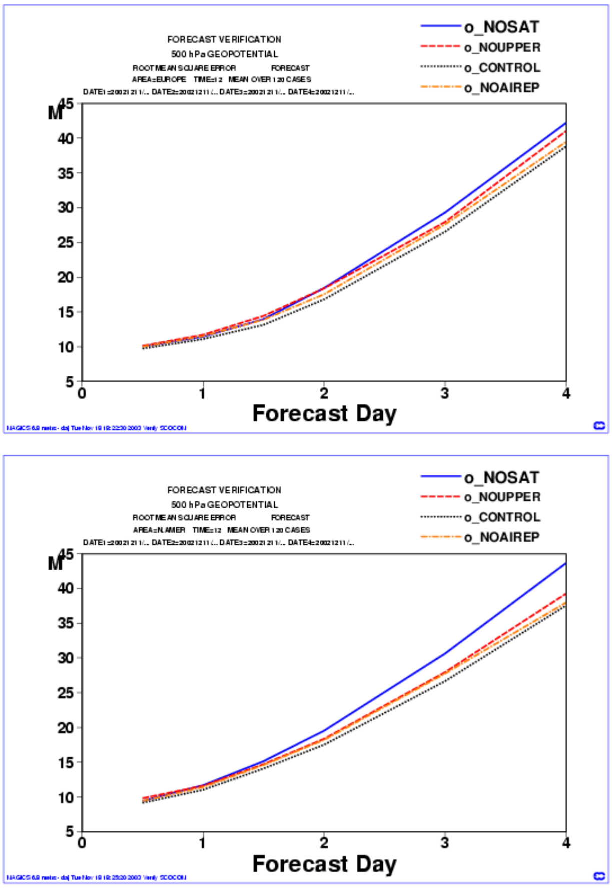

Figure 7. 500 hPa height RMS forecast errors as verified against radiosondes, over Europe (top) and North America (bottom). Colour code is identical as in Figure 6.

4.3. Observing System Experiments

Observing System Experiments (OSEs) constitute other means to study the impact of a given Observing System (see Dumelow, this volume). The impact of satellite data on the improvement of numerical forecasts has been confirmed by Kelly and Thépaut (2003) in the context of a large set of OSEs covering two periods of three months. Figure 6 represents the 500 hPa height anomaly correlation for the extratropical northern (top) and southern (bottom) hemispheres (averaged over 120 days) forecasts for various assimilation experiments. The dotted black curve corresponds to the control experiment (all data in = CONTROL), the orange curve to the experiment when no aircraft data are used (NOAIREP), the red curve to the experiment when no TEMP/PILOT data are used (NOUPPER), and the blue curve when all satellite data are withdrawn (NOSAT). It is not surprising to see the large impact of satellite data on the forecast quality over the southern hemisphere, although the amplitude (3 days of skill at day 5) has substantially increased with respect to previous similar OSEs (Bouttier and Kelly, 2001). What is particularly striking is that the impact of the satellite data is now noticeably larger than the impact of conventional sondes **in the northern hemisphere**. As illustrated in Figure 7 where Z500 RMS forecast error scores verified against radiosondes are displayed for Europe (top) and North America (bottom) up to day 4, these results hold also for short ranges. Note that the relatively small impact of sondes over North America (smaller than the impact reported in Bouttier and Kelly, 2001) can probably be explained by an active use of numerous aircraft data during descent/ascent phases in the ECMWF 4D-Var system. The difference of forecast error slopes between the NOSAT and NOUPPER/NOAIREP remains very impressive and demonstrates that the investment in methodologies for exploiting remote sensed observations in NWP is paying off.

5. Current issues and future challenges

5.1. Important Issues for the Assimilation of Satellite Radiances

If the direct use of radiances in variational techniques has become the preferred choice to incorporate measured satellite sounder observations in most of the operational NWP centres, important issues still need to be addressed to handle properly the treatment of remote sensed observations in data assimilation.

5.1.1. Monitoring and Bias correction

Systematic errors must be removed before the assimilation is performed to avoid propagation of various biases into the analysis. Biases in radiance data assimilation can originate from various sources: instrument error (calibration), radiative transfer error, cloud/rain screening error, background model error,…In practice, NWP centres address the problem of **bias correction** by **monitoring** departures between forecast background (or analysis) and radiance observations, thus identifying possible systematic differences. The use of several independent instruments is then recommended to disentangle the possible source of these differences. An illustration of this (borrowed from Thépaut, 2003) is given in Figure 8 showing the standard monitoring statistics daily produced at ECMWF (http://www.ecmwf.int). This plot shows on a daily basis during April 2001, the global mean departure between the ECMWF model background and channel 5 (temperature channel peaking at 600 hPa) from the HIRS instrument on board NOAA-14 (left panel) and NOAA-16 (right panel). The grey line represents the raw mean departure while the dark thick line represents the bias corrected departure. In this particular comparison, one clearly sees that NOAA-14 satellite has a +2K radiance bias against the model, while NOAA-16 has no radiance bias against the model (the same in both comparisons). This illustrates how powerful and important monitoring procedures and cross calibration techniques are to quantify biases in satellite instruments, although problems are not always as simple as the one illustrated in Figure 8. Indeed, although bias correction is mainly dedicated to remove radiative transfer and observational biases, it is very difficult to make absolutely sure that none of the model biases are removed at the same time. Monitoring is also essential to identify sudden deterioration in instrument performance, as a trigger for blacklisting actions.

Figure 8. Example of monitoring of channel 5 from the HIRS instrument on board NOAA-14 (left panel) and NOAA-16 (right panel) for April 2001. The grey line represents the non-corrected mean departure between observations and model first-guess. The thick dark line represents the bias corrected mean departure. Thin black lines show one standard deviation around the mean departure. From Thépaut (2003)

5.1.2. Quality Control and thinning

Quality control is an essential step in any data assimilation scheme to reject data of "bad" quality but also data that cannot be properly simulated by the NWP model. This is particularly true for satellite radiances that can be difficult to use because of the contribution of clouds, rain, surface emission and other parameters which are not represented correctly in the forecast model. For example, most of the atmospheric sounding instruments are currently not used over land due to the difficulty to handle the surface contribution to the radiative transfer equation. As pointed out by Derber (2000), a second difficulty arises from that the spacing of the satellite observations is small (a few kilometers). However, the horizontal analysis grid of global NWP models is fairly coarse (120km in the case of the ECMWF 4D-Var analysis) and furthermore, background error covariances exhibit currently fairly large horizontal scales and therefore forbid the details observed by satellites to be properly analysed. In addition, Liu and Rabier (2002) have shown that when observation errors are correlated, increasing the observation density beyond a certain threshold adds very little information even within an optimal data assimilation scheme. The problem gets worse in a suboptimal context that neglects observational error correlations. Therefore and also due to the additional computational burden of processing high resolution radiance data, satellite observations are thinned in boxes to a resolution that is scientifically "digestable" by the error statistics of the assimilation system, and technically compliant with computational constraints. The challenge is of course to design a "clever" thinning procedure to select the most informative spot in each thinning box (warmest-clearest radiance, smallest first-guess departure,…).

5.1.3. Observational error characterization

As stated above, the characterization of observational errors is in principle much easier for raw radiances than for retrievals because of far less processing steps involved. However, the observational error covariance matrix $\boldsymbol{R}$ should represent accurately the instrument (calibration), radiative transfer and representativeness errors (errors due to the inability of the analysis system to represent scales at a resolution compatible with the observations). Although it is likely that $\boldsymbol{R}$ is not diagonal (merely because errors in the radiative transfer model lead to inter channel correlations), most NWP centres ignore it in practice. Specifying properly $\boldsymbol{R}$ in radiance data assimilation remains a difficult issue that is becoming more important with the new generation of multi-channel advanced sounders.

6. Future challenges in satellite data assimilation

If satellite observations are already playing a major role in improving numerical forecasts today, the combination of increased model resolution, improved physics and treatment of the interfaces (ocean and land surface), enhanced data assimilation techniques, together with advances in satellite instruments over the next decade offer new potential observations available for NWP data assimilation centres. The horizontal and vertical resolution of numerical forecast models has been steadily increasing for the last twenty years in most NWP centres (following this trend and as an example, the horizontal resolution of the ECMWF model is expected to typically reach 10-15km in 2015, with a vertical resolution of around 80m in the boundary layer, 200-300m in the troposphere and 500-600m in the stratosphere). In parallel to model and data assimilation improvements, the evolution of the current and future satellite observing systems clearly indicates a massively increased spectral resolution in the infrared instruments onboard polar orbiters first (AIRS - Advanced Infrared Sounder- on board AQUA launched in 2002, IASI -Infrared Atmospheric Sounding Interferometer- onboard METOP from 2005 onwards, and CrIS on board NPP and NPOESS from 2006 onwards). In the microwave, improved sounding and imaging channel combination and scanning will be provided by the conical scanning SSMIS (Special Sensor Microwave Imager/Sounder) instrument onboard the recently launched DMSP-F16, and its operational follow-on CMIS on board NPOESS. More instruments providing information on the land surface but also on atmospheric composition (including trace gases) are already flying onboard R&D satellites (e. g. the ENVISAT payload) and more will be available in the future (SMOS from ESA, AURA from NASA). These new observations are not only useful for NWP application but also for climate, chemistry and environmental research. Active instruments (radars, LIDARS) will complement the current constellation to provide detailed vertical information on hydrometeors (Cloudsat, GPM), aerosols (EarthCare) and wind (ADM-AEOLUS). Last, limb sounding techniques (active and passive) will raise new challenges for data assimilation. These instruments should also contribute to improved temperature, moisture (and chemical species) vertical resolution.

ECMWF is currently devoting a lot of effort to tackle two scientific areas of particular interest: the exploitation of advanced high spectral resolution sounders and the assimilation of information on clouds and precipitations.

6.1. Assimilation of High Spectral Resolution Infrared Sounders

Those instruments measure radiation in several thousands of different channels, and therefore provide atmospheric temperature and composition information at a much higher accuracy and vertical resolution than what can be achieved with the current generation of current instruments such as HIRS. Among the list of challenges that NWP centres have to tackle in order to better absorb the massive amount of information available from space within the next decade, the exploitation of advanced sounders is currently probably the topic of most active research. Indeed, while individually channels from advanced sounders only provide a broad layer measurement, their multiple combination provide significantly higher vertical resolution than what is currently offered by conventional infrared instruments. Figure 9 displays the averaging kernels for the HIRS instrument (left panel) and the AIRS instrument (right panel). Averaging kernels correspond to the lines of the Model Resolution Matrix and ideal averaging kernels for a perfectly observed system would show 60 dirac curves corresponding to the 60 levels of the ECMWF model. It is obviously not the case even with the advanced AIRS instrument. However, it is clear from Figure 9 that AIRS data offer a much higher vertical resolution than HIRS. Data from the AIRS instrument onboard AQUA are now assimilated operationally at ECMWF and in research mode in many other NWP centres. Several key issues had to be (and are still being) addressed to use these data operationally.

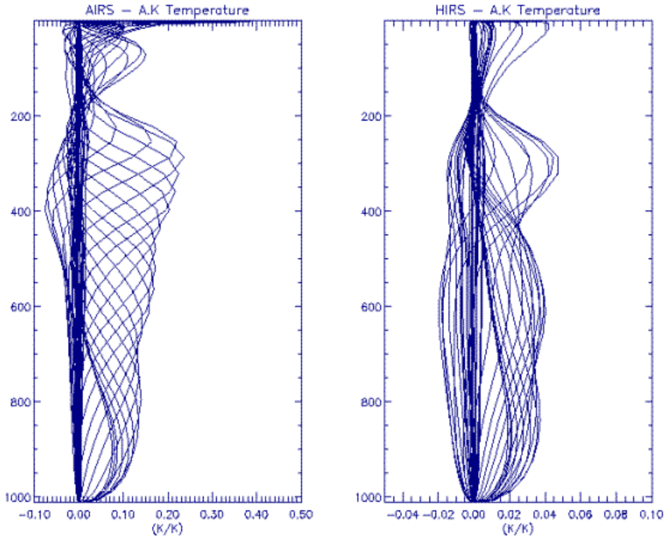

Figure 9. Averaging kernels (in K/K) for AIRS (left panel) and HIRS (right panel) instruments.

6.1.1. Monitoring and bias correction of high spectral resolution data

The monitoring system is extremely useful when a new instrument is launched. NWP centres have played a key role in the early checkout of the AIRS. The departure statistics provide vital information on the quality of the new data and how well the spectral characteristics of the instrument have been determined. The existing tools have been massively extended to simultaneously visualize several hundreds of channels, while keeping the facility to investigate the behaviour (in time and space) of individual channels. The monitoring statistics also provide the input to the bias corrections that have to be applied and to the characterizations of the observational errors. It is worth stressing that a comprehensive and consistent bias handling system will be a key element for the current and future exploitation of future satellite instruments. In particular, the extraction from advanced sounders of new quantities (such as trace gases) with a low signal-to-noise ratio will only be possible if different bias sources in temperature and water vapour are thoroughly understood and controlled. This requires a capability to disentangle systematic errors inherent to instrument, radiative transfer and NWP model deficiencies.

6.1.2. Channel selection and data compression

It is neither feasible nor efficient to assimilate all of the channels from high spectral resolution instruments (approximately 2400 in the case of AIRS and 8400 in the case of IASI) and a policy of channel selection has to be designed. The challenge is to find a set of channels that is small enough to be assimilated efficiently in a global NWP system (with operational time constraints), but which is still large enough to capture the important atmospheric variability. Most centres are designing channel selection strategies based on optimal theory or on a more pragmatic approach relying on a selection predefined at source (Collard 2000, Rabier et al. 2002, Fourrié and Thépaut 2003). The currently fairly limited impact of AIRS observed so far in NWP data assimilation centres may be partially explained by a waste of crucial information entailed by a too static channel selection that does not take into account the errors of the analysis of the day.

Other ways to reduce the huge data volume to be ingested in the assimilation are being considered. The purpose of data compression techniques is to reduce the data volume of the measured spectra without any distinguishable loss of accuracy. One candidate for this is to use a truncated set from the Empirical Orthogonal Functions (or EOFs) of the spectra with a suitably defined reproduction metric (Huang and Antonelli, 2001). Ideas of RT modelling directly in EOF space are coming up and will deserve attention within the next few years (Schlüssel, pers. com.).

6.1.3. Cloud detection

A key issue for the successful exploitation of advanced infrared sounders is the detection of the presence of clouds since they have a significant impact on the radiance measurements. Given that important atmospheric structures occur in cloudy areas, it is of utmost importance to take into account the effect of clouds when present and cloud detection is therefore needed to correctly interpret the data. A failure to detect the presence of clouds can result in a (possibly large) signal being wrongly interpreted by the assimilation system and force erroneous adjustments to temperature and humidity. An over-stringent cloud detection system can lead to very little data reaching the assimilation and thus reduce the impact of the new high resolution instrument. A number of strategies are being investigated by different NWP centres. McNally and Watts (2004) propose a technique based on the identification of characteristic spectral signatures or patterns in the observed radiances. The pattern recognition algorithm is used to identify which radiances at a particular location are free of cloud and which are not (it is therefore not designed to identify locations of completely clear spectra). In the longer term, cloud affected radiances could be exploited as they can, in principle, provide unique high-resolution information about the temperature structure near the cloud top (see the discussion below about the assimilation of cloudy and rainy information from satellites). In that respect, advanced sounders may indeed be able to determine the cloud top and cloud amount accurately.

6.1.4. Environment opportunities

Although this issue goes beyond strict NWP applications, it is worth mentioning that advanced sounders also provide information about minor constituents of the atmosphere (ozone, N2O, CO, CH4). It has now also been shown that CO2 information can be extracted from the high resolution sounder AIRS. This capability comes from the fact that by sampling the IR spectrum at very high resolution, one can measure radiation that is only dependent on temperature and the atmospheric CO2 concentration (small group of pure lines), while instruments with coarse spectral resolution sample radiation that is a mixture of absorbing species. If we have accurate temperature information (from NWP analyses driven by AMSU for example), CO2 signal can be separated out. Figure 10 displays the first map of stratospheric CO2 estimation (by the ECMWF data assimilation system) from the AIRS instrument (Engelen, pers. comm.). Although absolute values are still biased when compared with other independent data sources, the zonal distribution nicely represents the Brewer-Dobson circulation.

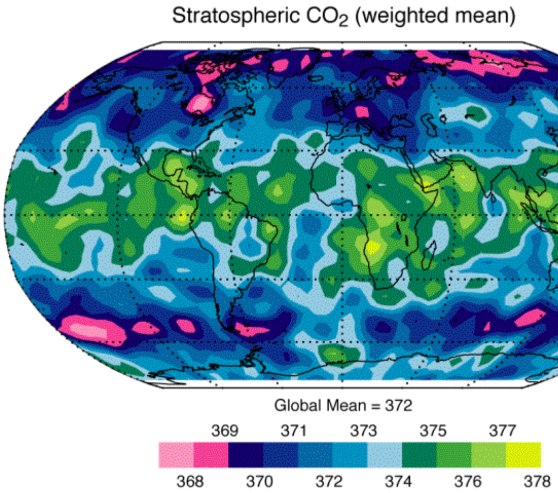

Figure 10. Analysed stratospheric CO2 distribution for 25-27 April 2003 using AIRS radiances.

6.2. Assimilation of Clouds and Precipitation

Sources of information on clouds and precipitation from satellites are already currently available. Satellite imagery (AVHRR, MODIS, METEOSAT,…) is already providing a lot of information on cloud fields, and rain information can be retrieved from microwave imagers. New missions such as Cloudsat-Calipso (to be flown together with AQUA) and Earthcare will enhance the validation of the cloud parameterizations of NWP models by providing for the first time information on cloud profiles thanks to the use of "active" instruments penetrating atmospheric hydrometeors. Global information on rain will be available for NWP with missions such as the Global Precipitation Mission (GPM) and its European component (E-GPM). A particular effort is already taking place in a number of NWP centres to assimilate cloud and rain information from satellites. A prerequisite to assimilate these observations is the ability of the numerical model to "look like" the observations. As it turns out, recent improvements in physical parameterizations (cloud description and convection) and radiative transfer modeling (in presence of absorption and scattering from hydrometeors) are such that the degree of agreement between numerical models and the observations is sufficiently good to be convinced that assimilation of such quantities is at hand. Several approaches are being developed to assimilate cloud and rain products in NWP. It is worth pointing out that a cloud/rain analysis is not the prime product that NWP centres are after. Cloud properties are forced in the model by the other atmospheric variables during the assimilation or the forecast integration. The goal is therefore to extend the analysis of temperature and humidity under all conditions, in order to provide an optimal coherent treatment of satellite observations under clear, cloudy and rainy areas, and subsequently to perform accurate cloud and rain forecasts. In the long term, the direct use of cloudy or rainy radiances in the ECMWF four-dimensional data assimilation system may be possible. However, as mentioned above, this technique requires an improved simplified physics in the incremental formulation of 4D-Var, a fast but accurate radiative transfer able to simulate cloud /rain absorption and scattering and a proper treatment of representativeness errors (in particular subgridscale variability errors due to the relatively coarse resolution of the ECMWF analysis grid). Intermediate approaches are being investigated for the rain assimilation. The assimilation of brightness temperatures (TB) has been developed as well as the assimilation of rain rate retrievals through a 1D+4D-Var approach. The system produces pseudo Total Column Water Vapour (TCWV) observations that are fed to the 4D-Var system (Moreau et al., 2003). Figure 11 displays the 12h accumulated precipitation field from a 24 (top) and a 48 hour (bottom) forecast performed from two experiments differing by the assimilation of TMI brightness temperatures (case of cyclone ZOE on December 26 2002). The difference of patterns between the "control" and the "radiance assim" runs illustrates that after 48 hours, the model has been able to maintain the rain and moisture modifications suggested by the radiance assimilation.

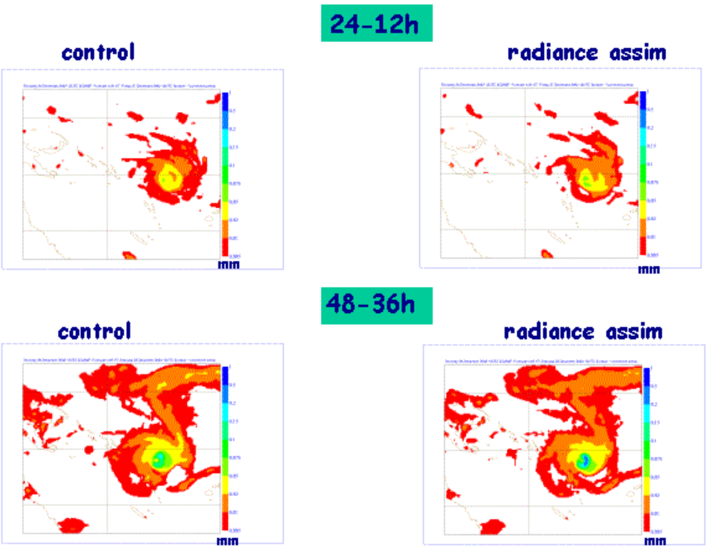

Figure 11: 12h accumulated precipitation field from a 24 (top) and a 48 hour (bottom) forecast performed from two experiments differing by the assimilation of TMI brightness temperatures (case of cyclone ZOE on December 26 2002).

It is also hoped that a more aggressive use of infrared radiances in presence of clouds is achievable (see the previous section), and may lead to an improved description not only of T/q parameters but also of cloud top height. However, the different linearity properties and the sensitivity of infrared radiances to smaller-scale vertical structures than microwave imagers may imply different strategies for the assimilation of clouds.

NWP data assimilation systems currently reject about 80% of satellite data because they are rain or cloud-contaminated. This number clearly indicates the potential benefit for NWP of investing in the assimilation of cloud and rain-affected satellite radiances.

7. Conclusions

This paper has briefly described the Space Observing System, as well as the different techniques to observe the Earth/Atmosphere. The increasing importance of satellite data in Numerical Weather Prediction has been reviewed. A special emphasis on what makes the satellite data specific has been presented. The advantages of assimilating raw information versus products have been reviewed. Finally, a number of future challenges to better exploit an increasing amount of information from satellites has been described.

Today, satellite data are successfully exploited (at least in clear sky and over sea) by modern data assimilation schemes, with the consequence that introducing *well characterized* satellite data improve the quality of NWP models almost systematically. Combined with improvements that make up numerical models (e. g. dynamics, physical parameterizations of sub-grid scale processes, heat and moisture exchanges with the surface,…), advanced data assimilation methods will make it possible to extract the maximum information content from existing and future satellite instruments in progressively more weather conditions.

This optimism should be tempered by that the future proliferation of new satellite instruments will make it hard for end-users to monitor/quality-control/assimilate efficiently all what becomes available and choices may have to be made. In any case, a massive effort in data handling and monitoring will have to be funded in a sustainable way. A short-loop dialogue between the NWP community and the Space agencies is therefore crucial to optimize the exploitation of future platforms.

Acknowledgements. The author is indebted to the research department at ECMWF, in particular the Satellite Section from where most of the material was borrowed. The contribution from Erik Andersson and Mike Fisher is greatly appreciated.

List of acronyms

ADM:	Atmospheric Dynamics Mission
AIRS:	Atmospheric Infrared Sounder
ATOVS:	Advanced TIROS Operational Vertical Sounder
AMSU:	Advanced Microwave Sounding Unit
AQUA:	NASA research satellite
AURA:	NASA research satellite
AVHRR:	Advanced Very High Resolution Radiometer
CALIPSO:	Cloud Aerosol Lidar and Infrared Pathfinder Satellite Observations
CLOUDSAT:	NASA research satellite
CMIS:	Conical Microwave Imager/Sounder
CRIS:	Cross-track Infrared Sounder
DMSP:	Defense Meteorological Satellite Program
DWL:	Doppler Wind Lidar
ENVISAT:	ENVIronment SATellite
ERS-2:	Earth Remote Sensing satellite -2
EarthCARE:	Earth Clouds Aerosols Radiation Explorer
ESA:	European Space Agency
EUMETSAT:	European organization for the exploitation of METeorological SATellites
GEO:	Geostationary Satellites
GOES:	Geostationary Operational Environmental Satellite
GOME:	Global Ozone Monitoring Experiment
GPM:	Global Precipitation Mission
GPS:	Global Positioning Satellite
HIRS:	High resolution Infrared Radiation Sounder
IASI:	Infrared Atmospheric Sounding Interferometer
JAXA:	Japan Aerospace eXploration Agency
JMA:	Japanese Meteorological Agency
LEO:	Low Earth Orbiting satellite
METOP:	METeorological OPerational satellite
MODIS:	MODerate resolution Imaging Spectroradiometer
MSG:	Meteosat Second Generation

NASA:	National Aeronautic and Space Administration
NOAA:	National Oceanic and Atmospheric Administration
NPOESS:	National Polar-orbiting Operational Environment Satellite System
POLDER:	POLarization and Directionality of the Earth's Reflectance
QuikScat:	NASA satellite
SAR:	Synthetic Aperture Radar
SBUV:	Solar Backscatter Ultra-Violet (ozone instrument)
SMOS:	Soil Moisture and Ocean Salinity Mission
SSM/I:	Special Sensor Microwave Imager
SSMIS:	Special Sensor Microwave Imager/Sounder
TMI:	TRMM Microwave Imager
TRMM:	Tropical Rainfall Monitoring Mission
UARS:	Upper Atmosphere Research Satellite

References

Andersson, E., J. Pailleux, J.N. Thépaut, J.R. Eyre, A.P. McNally, G.A. Kelly and P. Courtier, 1994: Use of cloud-cleared radiances in 3D and 4D variational data assimilation. *Quart. J.R. Meteor. Soc.*, **120**, 627-653.

Bouttier, F., and G. Kelly, 2001: Observing system experiments in the ECMWF 4D-Var data assimilation system. *Quart. J.R. Meteor. Soc.*, **127,** 1469-1488.

Chédin, A., N. Scott, C. Wahiche and P. Moulinier, 1985: The Improved Initialization Inversion method: a high resolution physical method for temperature retrievals from the TIROS-N series. *J. Climate Appl. Meteor.*, **24**, 124-143.

Collard, A., 2000: Assimilation of IASI and AIRS data: Information Content and Quality Control. *Proceedings of the ECMWF seminar on the exploitation of the new generation of satellite instruments for Numerical Weather Prediction*. Pp 201-224. available from the ECMWF library.

Derber, J., 2000. Assimilation of TOVS, GOES and ATOVS radiances. *Proceedings of the ECMWF seminar on the exploitation of the new generation of satellite instruments for Numerical Weather Prediction*. Pp 47-56. available from the ECMWF library.

Eyre, J. R., 2000: Planet Earth seen from space: Basic concepts. *Proceedings of the ECMWF seminar on the exploitation of the new generation of satellite instruments for Numerical Weather Prediction*. Pp 5-20. available from the ECMWF library.

Fisher, M., 2003: Estimation of entropy-reduction and degrees of freedom for signal for large variational analysis systems. ECMWF Tech. Mem. 397. available from the ECMWF library.

Fourrié, N. and J.-N. Thépaut, 2003: Evaluation of the AIRS near-real-time channel selection for application to numerical weather prediction. *Quart. J.R. Meteor. Soc.*, **129**, 2425-2439.

Huang, H. L., and P. Antonelli, 2001: Application of Principle Component Analysis to High Resolution Infrared Measurement Compression and Retrieval, *J. Appl. Meteor.,* **40**, 365-388.

Joiner, J., and D. Dee, 2000: An error analysis of radiance and suboptimal retrieval assimilation, *Quart. J.R. Meteor. Soc.*, **126**, 1495-1514.

Kelly, G. and J.-N. Thépaut, 2004: Impact of various Observing Systems in the ECMWF NWP model. *in prep.*

Liou, K. N., 1980: An introduction to atmospheric radiation. *Internal Geophysics Series*. **Vol 26**, Academic Press.

Liu, Z. Q. and F. Rabier, 2002: The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. *Q. J. R. Meteorol. Soc.*, 128, 1367-1386.

McNally, A. P. and P. D. Watts, 2004: A cloud detection algorithm for high spectral resolution infrared sounders. *Q. J. R. Meteorol. Soc.*, **130**, 1-13.

Moreau, E., P. Bauer and F. Chevallier, 2003: Variational retrieval of rain profiles from spaceborne passive microwave radiance observations. *J. Geo. Res.*, 108(D16), 4521, doi:10.1029/2002JD003315.

Phalippou, L., 1995: Variational retrieval of humidity profile, wind speed and cloud liquid water path from SSM/I: potential for numerical weather prediction. *ECMWF Tech. Memorandum **216**.*available from the ECMWF library

Rabier, F., J.-N. Thépaut and P. Courtier, 1998: Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Quart. J.R. Meteor. Soc.*, **124**, 1861-1887.

Rabier, F. , N. Fourrié, D. Chafai and P. Prunet, 2002: Channel selection methods for infrared atmospheric sounding interferometer radiances. *Q. J. R. Meteorol. Soc.*, **128**, 1011-1027.

Reale, A., C. Novak, M. Chalfant, H. Drahos and D. Gray, 1989: NOAA/NESDIS sounding products. *Proceedings of ECMWF/EUMETSAT workshop on the use of satellite data in operational numerical weather prediction*. Pp 153-172. available from the ECMWF library

Rodgers, C. D., 1976: Retrieval of Atmospheric Temperature and Composition from Remote Measurements of Thermal Radiation. *Rev. Geophys. and Space Phys*., **14**, 609-624.

Simmons, A. and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **128**, 647-687.

Stephens, G. L., 1994: Remote sensing of the lower atmosphere. An introduction. Oxford University Press.

Stoffelen, A., 2000: A generic approach for assimilating scatterometer observations. *Proceedings of the ECMWF seminar on the exploitation of the new generation of satellite instruments for Numerical Weather Prediction*. Pp 73-100. available from the ECMWF library.

Thépaut, J.-N., 2003: Assimilation of remote sensing observations in numerical weather prediction. *in Data Assimilation of the Earth System, Nato Science Series, Kluwer Academic.*