ministerie van verkeer en waterstaat

**koninklijk nederlands meteorologisch instituut**

# The use of satellite data in the ECMWF analysis system

Krisztina Labancz

Technical reports; TR-123

Technische rapporten; TR-123

de bilt, 1990

KNMI

## 1. Introduction

The main objective of a data assimilation scheme is to provide the initial state for a forecast model. The scheme that has been developed at ECMWF produces global analyses in numerical form using all appropriate types of available observations: conventional observations, cloud-track winds and satellite determined profiles (SATEMs). This paper gives a general survey of the use of satellite data. Quality control and selection of incoming observations are generally described in section 2. After a brief outline of data obtained from geostationary satellites, section 3 gives a more detailed description of polar orbiting satellite data. First the information content of these satellite data are accounted through the survey of the evaluation of horizontal and vertical resolutions of satellite sounding data sets. The radiance measurement instruments, the sensed parameters and the finally retrieved atmospheric profiles are summarized in the second part of section 3. Section 4 gives a brief description of the retrieval procedure, the quality controls and the data selection processes used by the ECMWF analysis system with regard to satellite sounding and radiance data. Finally a short outlook of further implementation on the use of satellite data is given in section 5.

## 2. Quality control and selection of data in the analysis scheme

In the ECMWF data assimilation scheme four analyses per day are produced. For each of these assimilation cycles a 6 hour forecast from the previous analysis is used as guess field. The global analysis, used as the initial state of the atmosphere for the forecast model, is a modification of this guess field by the observations using a three-dimensional multivariate optimum interpolation technique (Lorenc,1981). Hence an important part of an operational analysis system is choosing which data to leave out. Two types of data need to be identified (Lorenc, 1985); those which are grossly incorrect or misleading (quality control), and those which carry little extra information over others which are being used and which can therefore be disregarded to save time (data selection).

### 2.1. Quality control

It is essential to ensure that the data do not degrade the guess field but do correct its errors. This is performed by several quality control procedures at different stages in the assimilation scheme. The checks fall apart into seven different types (Lönnberg and Shaw,1983):

(i) Checks of code formats etc.

(ii) Internal consistency checks on the data within one observation.

(iii) Temporal consistency checks on observations from the same source.

(iv) Checks that the data are reasonably close to a climatological value.

(v) Checks that the data are reasonably close to a forecast value.

(vi) Spatial consistency checks between nearby observations.

(vii) Checks that an observation is reasonably close to an analyzed value, calculated without use of this observation.

At ECMWF checks (i)-(iv) are performed before the observations are inserted in the data-base (satellite soundings are used without temporal correction). Checks (v)-(vii) are performed during the analysis cycle (as a supplement to the other checks). Their purpose is to identify data which would degrade the analysis. These data are rejected.

### 2.2. Data selection

The data selection aims at supplying the analysis routine with checked data suitable for a specific analysis volume. The observations are sorted in the reports data-base into boxes of a same horizontal size of about 660kmx660km, and instead of choosing observations individually, a major part of the selection and checking is done while keeping the observations of one box as an entity.

Before the main analysis, the pre-analysis compresses the amount of data within one box at a time. Observations considered clearly erroneous by the quality control procedures are discarded and redundant information is compressed by combining close pairs of observations and forming "super-observations". At the end of pre-analysis stage not more than 40 observations per box can be passed to the main analysis program. In the rare event of the limit of 40 being exceeded, remaining observations are classified as "secondary" and placed in overflow records of the workfile. Such observations receive less consideration in subsequent selections, but still may be used in the final analysis (Lönnberg and Shaw,1983).

In the main analysis a reduced number of representative data are selected. The observations coming from only the neighboring boxes are sorted according to distance from the midpoint of the central box. Next, the observation set is truncated to contain observations which are within a certain distance Dmax from the center of the analysis box (for more details see Lönnberg and Shaw,1983). In data dense areas the data checking and analysis is done separately for three layers. The data selection proceeds separately for each slab of the atmosphere by using the observations ordered as described above. The slab boundaries are 1000-700, 700-150 and 150-10 hPa. Some serious effect can arise from this totally distinct selection for adjacent slabs (e.g. in the vertical profile of geopotential, where a discontinuity across a slab boundary transforms to an erratic profile of temperature). By permitting an overlap from one slab to the next in terms of data selected, one mitigates some of these effects.

## 3. Satellite observations

### 3.1. Satellite cloud track winds

Single level cloud track winds are obtained from geostationary satellites by tracking "target" clouds between successive images. There are currently five geostationary satellites in operation, each covering a circular field of view (of around 55 degrees) on the Earth from a height of approximately 3600km above the Equator. The temperature of the cloud top obtained from the infrared window radiances is used to determine the level of the cloud, and this information contributes to the height assignment of the resulting cloud track winds. The quality of these wind observations is monitored at ECMWF using statistics of observed minus first-guess differences accumulated over monthly periods. These differences are presented as bias wind vectors. In general the observed winds are lighter than those in the first-guess fields, mainly in areas of strong westerly flow, near the Northern and Southern Hemisphere polar jets and subtropical jet over Northern Arabia. This is an inherent problem with cloud track winds in the jet stream regions (Böttger and Radford,1989). In the ECMWF analysis scheme no cloud track winds are used over extratropical land areas (polewards of latitude 20) due to their minor quality, particularly over high orography. The root mean square observation error at e.g. 400hPa level is 3.7 m/s for radiosonde or pilot wind observations and 5.0 m/s for satellite-observed winds. The difference between these rms observation errors is a bit smaller below 400hPa and larger above this level (Lönnberg and Shaw,1983).

## 3.2. Satellite sounding and radiance data

### 3.2.1. Evolution of the horizontal and vertical resolutions

The data assimilation scheme at ECMWF uses satellite determined profiles (SATEMs) of polar orbiters. At the beginning of the 80's two data sets were used. The SATEMs, which were available on the GTS (Global Telecommunication System), comprised 14 adjacent thicknesses between the standard levels with 500km horizontal resolution. A more comprehensive data set was also produced in Washington and available at ECMWF. The basic information contents of these SATEMs were the same, but the horizontal resolution was 250km, and the presentation of the information was also different: one profile consisted of virtual temperatures for layers between standard levels. These virtual temperature data were converted to thicknesses through the hydrostatic equation. The ECMWF generally used the 250km SATEMs, but because of telecommunication problems it happened that they were missing in some areas while the 500km SATEMs were available (Pailleux,1986).

Since 1985, the ECMWF assimilation system has used only the SATEMs with 250km horizontal resolution. The NOAA TIROS-N series of satellites are high-quality observation platforms supplying a large amount of data. The TIROS-N Operational Vertical Sounder (TOVS) instrument packages on board of all these satellites allow vertical temperature and moisture structures (soundings) to be calculated between the surface and the tropopause. Soundings at 500km spacing are still considered to be the WMO standard sampling for exchange over the GTS. The higher resolution data are transmitted by separate arrangement between NOAA (National Oceanic and Atmospheric Administration of the United States) and Europe. Now ECMWF receives soundings from NOAA-9 and NOAA-10 polar orbiting satellites and the operational use of DMSP-8 and 9 (Defence Military Satellite Program Series) is currently under investigation (Böttger and Radford,1989). Instead of using 14 thin layers, the analysis scheme uses a reduced number of thicker layers (first 11 layers and now,since 1987 only 7 layers - 4 layers in the troposphere and 3 layers in the stratosphere). The main reason for the reduction of vertical resolution is that the satellite measurements only contain perhaps five or six pieces of information in the vertical (Kelly,1985), hence the less vertical layers better represent the real information content of SATEM profiles. Naturally at the same time some returning was made on the vertical covariance matrix for observation errors (used in the optimum interpolation scheme).

While there was a reduction in the vertical resolution, it became clear that there is far more information in the horizontal than was transmitted. At the end of 1988 a new TOVS data set was made available by NESDIS (National Environmental Satellite and Data Information Service) to Europe. This data set contains the soundings at roughly 80km resolution as well as the corresponding cloud cleared radiances.

### 3.2.2. The satellite-based radiance measurements

The TOVS package consists of three radiance measuring instruments: the High-resolution Infrared Radiation Sounder (HIRS), sampling at 20 infrared frequencies, the Microwave Sounding Unit (MSU), sampling at four microwave frequencies and the Stratospheric Sounding Unit (SSU), sampling at three additional infrared frequencies. Of these 20 HIRS channels, one is in the visible wavelengths, giving albedo, one is in the ozone band (sensing total ozone content). Three channels are in window regions giving essentially surface (or cloud top) temperature and three channels are sensing water vapor. This leaves 12 channels for temperature sounding, of which four are dimensioned to observing thick stratospheric layers. Of the 4 MSU channels one senses surface emissivity and one concentrates on the stratosphere. The radiance data as obtained by the TOVS instruments sense bulk, thermodynamic properties of the atmosphere: internal energy (i.e. layer thicknesses), mixing ratios for water vapor and total amounts of minor constituents like ozone (Prangsma,1989). Profiles of layer mean temperature are converted to profiles of geopotential thicknesses which are the inputs to the analysis scheme. The humidity data available from the SATEMs consist of precipitable water content for three layers 1000/700, 750/500 and 500/300hPa. They have been used operationally since March, 1986.

### 3.2.3. Retrieval of atmospheric profiles from radiance measurements

One of the main problems on the use of satellite radiance measurements is the retrieval of atmospheric profiles. All methods are based on the inversion of the integral equation describing the radiation transfer of the atmosphere for a number of wavelengths (channels) in both the microwave and the infrared part of the electromagnetic spectrum.

Current satellite-sounder inversion schemes, either for research purposes or for operational use, encompass the following three methods (Prangsma,1989):

(i) The operational NOAA/NESDIS scheme used to produce the so-called SATEM messages. These SATEMs contain geopotential thicknesses and precipitable water contents.

(ii) The methods based on the ITPP (International TOVS Processing

Package).

(iii) The 3I (Improved Initialization Inversion) method. The 3I produces geopotential thicknesses and relative humidities with 100km horizontal resolution. Cloud top heights, effective cloud amounts and surface temperature are also derived from raw counts of two satellites NOAA-9 and 10 (Chedin,1989).

The 3I scheme, which requires only one single iteration step, avoids radiative transfer computations all together by employing a pre-computed database, containing the radiances for all channels, all the historical temperature/humidity profiles, all viewing angles and all combinations of surface combinations. Since December 1986, the 3I code has been implemented, validated and run in global experiments using a data set of raw radiances at ECMWF.

Although the operational TOVS package produces radiosonde-like profiles close to their theoretical accuracy, these data cannot give the vertical detail of modern radiosondes. It is a problem to determine the temperature of N levels from radiance observations at say M discrete frequencies. Also many different temperature profiles will give rise to the same radiance measurements. Furthermore, the temperature profile solution tends to be unstable in the sense that small radiance measurement errors tend to produce large errors of temperature (Kelly,1985). In the ECMWF assimilation scheme there are strong selection criteria to choose a representative set of observations. No more than 255 data may be selected for one analysis box (see section 2.2) in the case of 250km horizontal resolution.

Because of the facts described above, complete satellite temperature or thickness profiles are used only over sea. Over land only the stratospheric thicknesses, i.e. above 100hPa are presented to the analysis.

## 4. The treatment of satellite radiation data at ECMWF

Before the observed quantities are inserted in the retrieval procedure, some calibration, transmission and navigation checks are performed as well as brightness temperatures are calculated from the sensed radiances by using the Planck function. The description of these checks and calculations are outside the scope of this paper. See e.g. "An update about calibration, navigation and formats" by G.Rochard, The Technical Proceedings of the Fourth International TOVS Study Conference for details.

### 4.1. Tests during the atmospheric profile retrieval procedure

The tests during the retrieval procedure depend on the type of retrieval method. Since 1986 the 3I code has been run in global experiments at ECMWF. The 3I procedure is based upon the creation, once and for all, of the synthetic TOVS Initial Guess Retrieval (TIGR) data set which comprises (among others) a description of a large number of atmospheric situation through radiosonde and rocketsonde data (pressure, temperature, moisture, ozone mixing ratio profile, latitude, longitude,date and so on) and computed values of the corresponding radiances and their associated brightness temperatures for the TOVS sounding channels. The observed radiances are first used to retrieve the "best" initial guess solution. This procedure makes use of the TIGR data set. The selected set of observed radiances is compared to each of the computed sets of radiances and the closest is retained. The second step is the inversion of the radiative transfer equation for retrieving meteorological parameters. The basis of this step is a maximum probability estimation procedure which obtains the final atmospheric profile solution from the initial guess solution (Chedin et al.,1985).

The presence of clouds over the observed area influences the retrieval process. Alternately, the characteristics of the cloud cover such as cloud top heights or cloud amounts are important parameters in a forecast model, hence some cloud detection tests are performed during the retrieval process (see Wahiche et al.,1986 for details).

For quality control purposes, all retrievals pass a series of tests aiming at detecting bad or doubtful retrieved profiles. The first class of tests consists of comparing the observations and corresponding initialization obtained from TIGR. The second class of tests are based upon comparisons between initialization and final retrieval, and separating the upper and the lower parts of the profiles. A retrieval is rejected if between the two profiles the distance for the upper part is too different from that for the lower part (Scott et al.,1988).

At the end of the retrieval procedure one gets a data set, including temperature profiles or geopotential thicknesses, cloud top levels, effective cloud amount, water vapor or relative humidity profiles, surface temperature and microwave surface emissivity. The use of these data in the data-assimilation scheme involves the replacement of the present SATEMs with the corresponding 3I products.

## 4.2. Quality control and data selection within the data-assimilation scheme

All the TOVS satellite thickness data are used completely over seas. However, over land only the stratospheric thicknesses, i.e. above 100hPa are presented to the analysis. These thickness data, calculated from the incoming retrievals, are treated in the same way as SATEM data in the analysis scheme of ECMWF. This is described in section 2. However, in data dense areas some new problems can arise in connection with the higher resolution of the TOVS retrievals profiles. The data-assimilation scheme acts as a scale dependent filter on the observations (Lönnberg and Shaw,1985). Observation densities beyond the analysis resolution (1.875 x1.875 i.e. 200kmx200km) may safely be compressed by forming average observations. In the ECMWF system this is done by averaging pairs of observations which are separated less than approximately 100 km from each other. The retrieval of temperatures (and thus also the thicknesses) from the radiance data is not a linear procedure. Hence in data dense areas it might be better to make super-observations directly from the radiance measurements to avoid partly the nonlinear growth of measurement errors which occurs during the retrieval procedure. Because the resolution of TOVS data is about 100km, the same value than the 100km limit of making super-observations, it is difficult to tell a priori how often TOVS observations will be "super-obbed".

## 5. The impact of satellite soundings and further developments

A positive impact of satellite data is observed under two types of situations. In data sparse areas (where "any" information is better than no information at all) positive effects are found - e.g. over the Southern Hemisphere. There is a positive impact in data rich areas when the mesoscale structure is not adequately resolved by the traditional observation network. In that case the satellite data can give considerably better defined horizontal gradients than traditional observations. This type of impact is therefore observed in high resolution models over time periods in the range of lifetimes of mesoscale weather systems. This type of impact is expected over e.g. the North East Atlantic ocean and its coastal areas (Prangsma,1989).

Generally speaking, the ECMWF analysis scheme still accepts too many satellite observations, many of them being poor or doubtful. SATEM profiles were discovered to be of poor quality over ice-covered seas. A potential improvement could be the consideration of SATEM data over ice as over land, i.e. used only above 100hPa. Also, for non-clear soundings, a more careful usage could be considered below 100hPa, especially in the Northern Hemisphere, where the impact of SATEMs is found to be negative in some cases (Pailleux et al.,1989).

In the longer term, the research in data assimilation is now "converging" towards using something which is as close as possible to the genuine observed quantity, rather than "interfaces". SATEMs or retrievals from the 3I process are interfaces, and the interface can never transmit the real information content as it is in the radiances. The integrated approach (using the radiances directly) is easier to achieve in the context of a variational analysis scheme than in the traditional optimal interpolation scheme.

Acknowledgements

I am grateful to my colleagues at KNMI (Royal Netherlands Meteorological Institute) for help and comments. In particular I am indebted to Ad Stoffelen for valuable suggestions and comments which improved this paper.

References

Böttger, H. and Radford, A., 1989: *Monitoring of satellite data*. In: Proceedings of the ECMWF Seminar on Data Assimilation and the Use of Satellite Data (ECMWF: 5-9 september 1988), pp.235-250, Published by ECMWF, Reading, UK.

Chedin, A., Scott, N.A., Wahiche, C. and Moulinier, P., 1985: *The Improved Initialization Inversion Method: A high resolution physical method for temperature retrievals from satellites of the TIROS-N series*. Jour. Clim. Appl. Met., Vol.24(2), pp.128-134.

Chedin, A., 1988: *The 3I retrieval method. Recent local and global applications*. In: Proceedings of the ECMWF Seminar on Data Assimilation and the Use of Satellite Data (ECMWF: 5-9 september 1988), pp.181-214, Published by ECMWF, Reading, UK.

Delsol, F., 1985: *Monitoring the availability and quality of incoming GTS observations at ECMWF*. In: Workshop on the Use and Quality Control of Meteorological Observations (ECMWF: 6-9 november 1984), pp.55-89, Published by ECMWF, Reading, UK.

Kelly, G.A., 1985: *The accuracy and information content of satellite atmospheric vertical profile measurements*. In: Workshop on the Use and Quality Control of Meteorological Observations (ECMWF: 6-9 november 1984), pp.55-89, Published by ECMWF, Reading, UK.

Lorenc, A.C., 1981: *A global three-dimensional multivariate statistical interpolation scheme*. Mon.Wea.Rev., Vol.109, pp.701-721.

Lorenc, A.C., 1984: *Analysis methods for the quality control of observations*. In: Workshop on the Use and Quality Control of Meteorological Observations (ECMWF: 6-9 november 1984), pp.397-428, Published by ECMWF, Reading, UK.

Lönnberg, P. and Shaw, D.(Eds.), 1983: *ECMWF Data Assimilation Scientific Documentation*. ECMWF Meteorological Bulletin M1.5/1, Research Manual 1, Published by ECMWF, Reading, UK.

Lönnberg, P. and Shaw, D., 1985: *Data selection and quality control in the ECMWF analysis system*. In: Workshop on the Use and Quality Control of Meteorological Observations (ECMWF, 6-9 november 1984), pp.225-254, Published by ECMWF, Reading, UK.

Pailleux, J., 1986: *Use of satellite data in the ECMWF analysis system*. In: Workshop on High Resolution Analysis (24-26 June 1985), pp.15-26, Published by ECMWF, Reading, UK.

Pailleux, J., Kelly, G., Flobert, J.F. and Andersson, E., 1989: *Current work at ECMWF on the use of satellite data*. In: Proceedings of the ECMWF Seminar on Data Assimilation and the Use of Satellite Data (5-9 september 1988),pp.215-234, Published by ECMWF, Reading, UK.

Prangsma, G.J., 1989: *Basic characteristics of methods for the retrieval of atmospheric profiles from space-based radiation measurements*. KNMI Personal Memorandum FM-89-03, Unpublished manuscript available at KNMI, de Bilt.

Scott, N.A., Chedin, A., Breon, F.M., Claud, C., Flobert, J.F., Husson, N., Levy, C., Tahani, Y., Prangsma, G.J., Rochard, G., 1988: *Recent advances in the retrieval of meteorological parameters through the 3I system*. In: The Technical Proceedings of the 4th International TOVS study Conference (Igls, Austria, 16-22 march 1988), pp.289-322, edited by Menzel, W.P..

Wahiche, C., Scott, N.A. and Chedin, A., 1986: *Cloud detection and cloud parameters retrievals from the satellites of the TIROS-N series*. Annales Geophysicae, Vol.4B(2), pp.207-222.