The use of raw TOVS / ATOVS radiances in the ECMWF 4D-Var assimilation system

A.P. McNally, E. Andersson, G. Kelly, R.W. Saunders
ECMWF Shinfield Park, Reading Berkshire, UK

Radiance data from the NOAA polar orbiting satellites have been assimilated at ECMWF for a number of years and represent a major component of the global observing system. The analysis schemes used to assimilate the data have changed significantly in this time from the first 1D-Var in June 1992 (Eyre et al. 1993), to 3D-Var (Andersson et al. 1998) in January 1996 and most recently 4D-Var in November 1997 (Rabier et al. 1998). However, the basic form of the radiance data used in the analysis has remained unchanged from the original NESDIS pre-processed cloud-cleared (TOVS) radiance products. These data have undergone a number of significant pre-processing stages (at NESDIS) before they are distributed to NWP centres and it is known that some of these stages can introduce complicated random and systematic errors in the data that are not present in the original raw radiance observations. There are good historical reasons why the pre-processing is applied, related to the fact that the radiances were originally intended to be used in linear retrieval schemes. However, most of the pre-processing is not necessary to use the data in analysis schemes such as 3D or 4D-Var and, since it can introduce errors, it is in fact undesirable. Furthermore, when a new satellite is launched (e.g. the NOAA-15 spacecraft in April 1998 carrying the new ATOVS instruments), raw radiance data may be available for some considerable time (up to a year) before the pre-processed radiance products are distributed. This paper describes the recent modifications to the ECMWF assimilation scheme that allow the raw TOVS / ATOVS radiance observations to be used instead of the pre-processed data. The results of experiments carried out to test the meteorological impact of the change are also presented.

Changes to the assimilation system

Use of unmapped instrument data

The TOVS instrument actually consists of three completely independent radiometric units. The High-resolution Infra-Red Sounder (HIRS), the Microwave Sounding Unit (MSU) and (on alternate NOAA spacecraft) a Stratospheric Sounding Unit (SSU). The new ATOVS instrument replaces the MSU with two Advanced Microwave Sounding Units (AMSU-A and B). All of these measure radiation leaving the top of the atmosphere in different parts of the electromagnetic spectrum, but the individual instrument characteristics are very different. The most obvious difference is in the scanning geometry (i.e. the number of measurements taken per scan and the size of the area or “field of view” from which radiation is obtained). The instrument characteristics are described in Table 1.

The NESDIS pre-processing attempts to map radiance data from all the different instruments on to a single scanning geometry (in practice to the geometry of the HIRS) in order to combine the multi-spectral information at a single geo-location. Unfortunately the mapping is difficult to achieve accurately (requiring a very precise knowledge of the relative scanning and orbital characteristics) and even when done correctly can introduce spurious signals into the data (e.g. when a low resolution instrument such as the MSU is mapped to a high resolution HIRS geometry). The analysis has been modified to deal with each instrument as an independent source of radiance data that is assimilated on its natural scan geometry. Information from the different instruments is still combined implicitly by the 3/4D analysis, but in a more optimal way respecting the true information content of the data.
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Spectral region measured</th>
<th>Size of field of view (FOV)</th>
<th>No. of FOV per scan line</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIRS</td>
<td>20 channels 3.7 - 15 microns</td>
<td>17 to 58 km (nadir to limb)</td>
<td>56</td>
</tr>
<tr>
<td>MSU</td>
<td>4 channels 50 - 57 GHz</td>
<td>109 to 323 km (nadir to limb)</td>
<td>11</td>
</tr>
<tr>
<td>SSU</td>
<td>3 channels at 15 microns</td>
<td>147 to 450 km (nadir to limb)</td>
<td>8</td>
</tr>
<tr>
<td>AMSU-A</td>
<td>15 channels 23 - 57 GHz</td>
<td>50 to 160 km (nadir to limb)</td>
<td>30</td>
</tr>
<tr>
<td>AMSU B</td>
<td>5 channels 89 - 183 GHz</td>
<td>17 to 55 km (nadir to limb)</td>
<td>90</td>
</tr>
</tbody>
</table>

*Table 1: TOVS / ATOVS instrument characteristics.*

**Limb and emissivity adjustment**

As a satellite instrument scans away from the vertical (or nadir) viewing position the radiation that is measured originates from a slightly higher level in the atmosphere due to the increased absorption of a longer (so called limb) atmospheric path. For channels sensitive to the troposphere this results in colder measurements at the limb relative to nadir and for stratospheric channels measurements can be warmer at the limb due to the often reversed lapse rate. Furthermore, channels sensitive to the surface display a scan variation due to the changes in surface emissivity with viewing angle. The NESDIS pre-processing attempts to remove the radiance variation with scan position by a process called limb adjustment. Non nadir measurements are statistically adjusted to estimate the radiance that would have been obtained from the same atmosphere and surface viewed at nadir. A significant proportion of the scan dependent systematic errors in the pre-processed data have been traced to problems in this limb adjustment process. ECMWF’s analysis has been modified to assimilate the raw radiance data at the scan position it was actually measured by performing radiative transfer and emissivity calculations along slant paths. While these calculations are currently less accurate than those for nadir paths (but should improve as more experience is gained) the associated systematic errors are generally smoother, and smaller, than the limb adjustment errors. More importantly, errors are found to be more stable in time allowing a more effective bias correction to be applied.

**Cloud and precipitation detection**

Radiance data are used to specify the atmospheric temperature and humidity structure within the analysis, but in many cases the measurements may be strongly contaminated by other atmospheric phenomena such as cloud (in the infra-red) and precipitation (in the microwave part of the spectrum). Analysis schemes are not yet sophisticated enough to extract only the temperature and humidity information from contaminated radiances so, at present, these data must be identified and removed before the analysis. If this is not done the analysis will attempt to interpret signals in the radiance data that are due to e.g. cloud by making erroneous adjustments to the atmospheric temperature and humidity. In the pre-processed radiances NESDIS attempts to identify contaminated situations using a
battery of tests that search for the characteristic statistical signatures of cloud and precipitation in the data. While these tests are generally quite effective and detect most contamination they are done in the absence of any up to date information about the particular state of the atmosphere and therefore cannot be applied too stringently. This has resulted in cases where significant residual cloud and precipitation are found in the NESDIS pre-processed data. The analysis screening (that uses a short-range forecast of the atmospheric state) has been modified to detect and reject contamination in the raw radiance data. The forecast is used to compute clear sky values of the window channel radiances (that are extremely sensitive to the presence of cloud and precipitation); if the measured values are found to differ significantly from these the situation is assumed to be contaminated. Results suggest that this approach is generally more stringent than the tests applied in the NESDIS pre-processing.

Figure 1: Zonally averaged monthly mean temperature analysis for the OPS assimilation in February 1999 (upper) and mean analysis differences (lower) defined as experiment minus control. Units are degrees Kelvin.

**Bias correction**

In the absence of the statistical adjustments applied to produce the pre-processed data and with a more stringent detection of cloud / precipitation contamination the raw radiance data are found to have smaller and more stable systematic differences from the background field that require less bias correction. Also, since many potential sources of systematic error have been eliminated it has been possible to gain a better understanding of the sources of bias that remain. It is now believed that the major source of systematic difference that remains is not in the data itself, but in our ability to simulate the radiances within the analysis. This is due to a combination of errors in the physics of the radiative transfer model and an insufficiently precise knowledge of the exact spectral intervals (or channels) in which the radiation is actually measured by the instruments. The exposure of these issues
as a significant source of bias has resulted in some considerable effort directed towards improving the radiative transfer model. A particularly encouraging feature of the new AMSU instrument on NOAA-15 is that only very small biases are observed, suggesting that the radiative transfer and channel characteristics are better known than on previous spacecraft.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{bias.png}
\caption{a) Mean fit of the background (solid lines) and analysis (dash lines) to radiosonde temperature observations (20°N to 20°S) for the OPS (black curves) and EXPT (red curves) assimilations. b) Mean fit of the background (solid lines) and analysis (dashed lines) to radiosonde observations of zonal wind (20°S to 90°S) for the OPS and EXPT assimilations.}
\end{figure}

**1D-Var**

Before being supplied to the main 3D or 4D-Var analysis the pre-processed radiances data were passed through a 1D-Var retrieval scheme. This was used to extract information about the atmospheric temperature outside the domain covered by the main analysis (e.g. the stratosphere) and to quality control the radiance data. The recent implementation of the 50 level model (extending up to 0.1 hPa) removes the need to pre-specify the stratosphere with 1D-Var and a new quality control scheme has been developed such that the 1D-Var is no longer used.

**Analysis quality control of radiances**

Inside the analysis all observations (conventional and satellite) are checked against the forecast to detect and reject bad data. A desirable feature of such an approach is that an observation is less likely to be rejected in an area where we have less confidence in the quality of the short-range forecast (for example in data sparse regions or areas of rapid dynamic development). In the past this flexibility has not been possible for radiance observations since an estimate of forecast error in radiance space has not existed and the data had to be checked against essentially fixed thresholds. A method has now been developed to generate estimates of the forecast error in radiance space (Andersson et al. 1999) allowing the quality control of radiances to be dynamic and consistent with the treatment of conventional data.
Analysis and forecast impact experiments

Experiments have been conducted to test the performance of the new radiance assimilation scheme compared to that of the current operational configuration. It is important to note that there are two major differences between the experimental system (henceforth called EXPT) and the operational system (henceforth called OPS). The OPS system assimilates pre-processed TOVS radiance data from the NOAA-11 and NOAA-14 spacecraft. The EXPT system assimilates raw TOVS and ATOVS radiance data from the NOAA-14 and NOAA-15 spacecraft respectively. A cleaner comparison to test the raw radiance approach would have required pre-processed radiance data from NOAA-15 to be in the OPS assimilation, but this data is not expected to be available from NESDIS before the end of May 1999 (the raw data has been available since August 1998). Alternatively, the EXPT assimilation could have used raw TOVS radiances from NOAA-11 and NOAA-14 data, but as NOAA-11 was expected to become obsolete at the end of April 1999 it was not considered a worthwhile investment of effort to test a configuration with only a short anticipated lifetime (in fact the spacecraft failed at the end of February, 1999). Thus when comparing the results presented here it must be understood that there will be differences due to the extra microwave information provided by AMSU on NOAA-15 and differences resulting from the use of raw radiances rather than pre-processed data.

![Northern Hemisphere - 500 hPa geopotential](image1)

![Southern Hemisphere - 500 hPa geopotential](image2)

Figure 3: Root mean square forecast errors in the northern and southern extra-tropical regions for the OPS (blue curves) and EXPT (red curves) verified against radiosonde observations of 500 hPa geopotential. The sample consists of 127 cases.

Analysis impact

The changes to the assimilation system have resulted in some significant differences in the resulting analyses. Figure 1 shows the zonal mean temperature analysis for OPS, and the difference OPS-EXPT. The largest systematic differences are in the stratosphere and are due to the use of the AMSU-A radiances by the EXPT assimilation. This instrument has six channels that peak above 100 hPa, the highest being sensitive to the atmospheric temperature around 2 hPa. This is a significant addition to the information previously provided by the uppermost channels of the HIRS and SSU instruments can be exploited in the recent extension of the ECMWF forecast model in to the stratosphere. Mean changes in the troposphere are generally small, but are significant over the polar areas where there have been changes to the radiance data usage. The interpretation of raw radiance data in the analysis and the quality control procedures described earlier (that aim to detect cloud and precipitation) rely
upon an accurate estimate of the underlying surface skin temperature and emissivity. Over the poles little is known about the quality of the model skin temperature (and even less about the surface emissivity) so it was decided to be cautious (at least initially) and not use any channels that have a significant sensitivity to the surface. Unfortunately most of the channels that provide mid to lower tropospheric information are also sensitive to the surface and this policy results in some considerable loss of data. However, results (from forecast comparisons presented later) suggest that it is safer not to use tropospheric radiances in these areas than to attempt to use radiance data that we do not fully understand.

Figure 4: Root mean square forecast errors in the northern and southern extra-tropical regions for the OPS (blue curves) and EXPT (red curves) verified against radiosonde observations of 30 hPa temperature and vector wind.

**Verification of analysis changes**

A key quality indicator for any data assimilation system is the extent to which it draws to radiosonde and other conventional observations in the presence of (generally more numerous) satellite data. Experience has shown that the fit to conventional data (e.g. the root-mean-square temperature
differences between the background/analysis and the radiosondes) are generally very stable quantities, but are often adversely affected when there are problems with the use of satellite data. An example of the fit of the EXPT and OPS assimilations to radiosonde temperature data is shown in figure 2a). It can be seen that the EXPT assimilation fits the tropical radiosonde data better than OPS at many levels but there is some degradation around 100 hPa and 20 hPa. There are improvements in the extra-tropical regions, but they are generally much smaller and are not shown here. Thus the radiosonde temperatures suggest an improved temperature analysis in the vicinity of the radiosonde data. It is also useful to examine wind statistics. These are sensitive to the correct specification of the horizontal gradient of temperature (particularly in the extra-tropics assuming a geostrophic balance) and thus represent a less local measure of the assimilation quality. Statistics for the southern hemisphere are shown in figure 2b) and show a small, but consistent improvement in the wind fit of the EXPT assimilation compared to that of OPS.

Forecast impact

It is always difficult to relate the changes that have been made in the assimilation system to changes in the quality of forecasts. After a certain time (estimates vary between three and five days) it becomes impossible to trace forecast differences back to differences in the initial conditions (i.e. the analysis). Furthermore, before this so called non-linear stage of the forecast the choice of the truth against which we measure the quality of the forecast is significant. Usually the analysis is used, but if there are significant changes to the analysis for OPS and EXPT it must be remembered that there are two possible versions of the truth. An obvious compromise is to verify the forecasts from OPS and EXPT against radiosonde observations (which are not perfect but are the same for both experiments) and it is these results that are presented here. Figure 3 shows root-mean-square errors for the forecasts of 500 hPa height. It can be seen that the tropospheric impact of the EXPT is generally neutral in the extra-tropical northern hemisphere (although clearly positive after day 6) and positive at all ranges in the southern hemisphere, note that the differences in spatial and temporal coverage of southern hemisphere radiosondes give different statistics at 00 UTC and 12 UTC. No clear consistent impact was found in the tropical wind forecasts in the troposphere.

It was reported in Newsletter No. 82 that the stratospheric performance of the L50 forecast model (equivalent to the OPS in this context) showed a significant improvement over the previous L31 system. A small additional improvement has been found with the temperature and wind forecasts from the EXPT system relative to OPS (shown in figure 4 at 30 hPa as this was the level quoted in the previous newsletter).

Summary and future work

The ECMWF data assimilation has successfully been converted to use raw TOVS / ATOVS radiance data and has resulted in some useful improvements in the quality of analyses and forecasts. In the troposphere these are most likely due to the improved bias correction and quality control that are possible with the raw data. In the stratosphere the gain is almost certainly due to the extra information provided by the AMSU instrument.

The system described here must be considered only a first step in the direction of raw radiance assimilation. Since 1992 we have gained a great deal of experience with the use of pre-processed NESDIS data and it will take some time before our understanding of the raw radiances reaches maturity. The next step will be to extend our use of the raw data to the channels that are sensitive to the lower troposphere and surface. This is currently hindered by uncertainties in our knowledge of the physical characteristics of the surface and phenomena such as cloud and precipitation. Further in the future the challenge will be to extract valuable information on these processes and not regard them as contaminants to be removed.
The development of the raw radiance assimilation system is timely for the next ECMWF re-analysis project ERA-40. The use of raw radiance data will not be subject to the many changes that have occurred over the years in the NESDIS pre-processing and thus allow a greater degree of time consistency in the analyses.
TECHNICAL PROCEEDINGS OF THE TENTH
INTERNATIONAL ATOVS STUDY CONFERENCE

Boulder, Colorado
27 January - 2 February 1999

Edited by
J. Le Marshall and J.D. Jasper
Bureau of Meteorology Research Centre, Melbourne, Australia

Published by
Bureau of Meteorology Research Centre
PO Box 1289K, GPO Melbourne, Vic., 3001, Australia

December 1999