A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts


The Met. Office, UK

(Received 4 October 1999; revised 17 March 2000)

SUMMARY

The Advanced TIROS Operational Vertical Sounder (ATOVS) was launched on the NOAA-15 satellite in May 1998. This provided a very significant improvement in the information available from meteorological polar-orbiting satellites compared with the previous TIROS Operational Vertical Sounder system, particularly for humidity and vertical resolution of temperature in cloudy areas. In preparation for assimilation of the observations into a three-dimensional analysis of atmospheric temperature and humidity, the observations have been compared with calculated top-of-atmosphere brightness temperatures computed from numerical weather prediction model profiles of temperature and humidity. Differences between observed and modelled brightness temperature are small. In some parts of the tropics and northern hemisphere the standard deviation of these differences for the tropospheric Advanced Microwave Sounding Unit sounding channels is only marginally higher than the radiometric noise of the observations. Early in 1999 a series of observation-system experiments were completed in which ATOVS observations were assimilated using a one-dimensional variational analysis. No use of the new humidity information could be made because of interference problems experienced by the microwave humidity sounder on ATOVS. Nonetheless, these experiments showed that the assimilation of the new temperature information provided by the radiance observations reduces forecast errors by as much as 20% in the southern hemisphere and 5% in the northern hemisphere. Further improvements have been found by assimilating more data over land. The major impact arises from the microwave channels. Whilst forward-model errors may be slightly lower for the microwave channels than the infrared channels the primary reason is the provision of sounding information in active weather systems, which are usually cloudy.

KEYWORDS: ATOVS  Data assimilation  Numerical weather prediction  Satellite sounding  1D-var  3D-var

1. INTRODUCTION

In recent years vertical sounding information provided by radiance measurements made on board satellites has become an increasingly important component of the global observing system (Andersson et al. 1994; Gadd et al. 1995; Derber and Wu 1998). Over the oceans there are very few radiosondes and over land the number of radiosonde observations is declining and expected to decline further (McPherson 1999). Even over land there are existing data-sparse areas (e.g. Antarctica). Therefore, it is important to provide some information in data-sparse regions and to balance the decline in radiosonde numbers in data-rich areas by making better use of the satellite observations. This can be achieved both through improved data assimilation methods and enhancement of the polar-orbiting sounding capability. In 1998 improved vertical resolution of satellite sounding information in cloudy areas was provided by the Advanced TIROS Operational Vertical Sounder (ATOVS) on the NOAA-15 satellite. The ATOVS consists of three instruments: the High-resolution Infrared Radiation Sounder (HIRS), which is an infrared temperature sounder; the Advanced Microwave Sounding Unit-A (AMSU-A), which is a microwave temperature sounder; and the Advanced Microwave Sounding Unit-B (AMSU-B), which is a new five-channel microwave humidity sounder. Compared with the TIROS Operational Vertical Sounder (TOVS) system on previous NOAA satellites, this represents a significant advance in microwave sounding

† Television InfraRed Operational Satellite.
‡ National Oceanic and Atmospheric Administration.
capability. The AMSU-A combines the role of the Microwave Sounding Unit (MSU) and the Stratospheric Sounding Unit (SSU), which is an infrared pressure-modulated radiometer. The MSU and the SSU have six sounding channels, whereas the AMSU-A has eleven, as well as four window channels. Therefore, as well as combining the roles of the SSU and MSU, the AMSU-A has the potential to provide more information. The TOVS has been described in more detail by Smith et al. (1979), and Saunders (1993) has provided a brief description of the ATOVS.

A major application of ATOVS and TOVS observations is the assimilation into numerical weather prediction (NWP) models in order to produce more accurate weather forecasts. There are three alternative methods to assimilate the information: (1) NWP-independent retrievals; (2) NWP-dependent retrievals; and (3) direct radiance assimilation. All three options are currently being used operationally at different NWP centres. Gadd et al. (1995) have described the transition from assimilating NWP-independent retrievals to NWP-dependent retrievals at The Met. Office, and Derber and Wu (1998) and Andersson et al. (1994) have described the process of direct assimilation of radiances in three-dimensional variational assimilation (3D-var) systems. In this paper, new (with respect to TOVS) scientific problems posed by ATOVS data are identified and solutions tested. The preprocessing of the radiances, the one-dimensional variational assimilation (1D-var) retrieval method and (briefly) the assimilation of the 1D-var retrievals in 3D-var are described in section 2. The intended scope of the paper is to evaluate, through experimentation, the success of new techniques for processing ATOVS data. This is achieved both through the fit of the observed radiances to short-range (six-hour) forecasts by an NWP model (section 3), and the impact of assimilating ATOVS data on the accuracy of NWP forecasts (section 4). The details of the techniques are usually referenced, rather than described, for the sake of brevity. The main focus of this paper is on the impact of the new observations from AMSU-A, although this is placed in context by also evaluating the impact of HIRS, MSU and SSU observations. The paper attempts to answer two questions: 'Do AMSU-A observations provide significant new information?' and 'How large an improvement in the accuracy of weather forecasts from an NWP system can be achieved through the assimilation of this new information?'.

2. PROCESSING OF ATOVS RADIANCES

(a) Preprocessing of radiances

In this study, calibrated and earth-located TOVS radiances are received from the National Environmental Satellite Data and Information Service (NESDIS). NESDIS maps the MSU and SSU observations onto the HIRS instrument grid. Furthermore, NOAA/NESDIS correct the data to nadir view in a process known as limb correction. The dataset distinguishes which HIRS radiances relate to clear air and which relate to cloudy air, and additional information (e.g. the total ozone and the surface type—land, sea, and sea ice) have also been included. This processing by NESDIS has been described by Reale et al. (1995). This dataset is referred to in this paper as the TOVS data. By contrast, the ATOVS data are received in what is known as level 1B format. Level 1B format includes raw instrument counts and calibration coefficients. The preprocessing of these raw counts is primarily achieved using the ATOVS and AVHRR* Processing Package (AAPP) described by Klæs (1997). This calibrates and earth locates the observations, preprocesses the radiances to the form required by the assimilation system and calculates some quality-control information. The AMSU-A and AMSU-B

* Advance Very-High-Resolution Radiation instrument.
data are mapped onto the HIRS-instrument grid, but no limb correction is applied. This dataset is referred to in this paper as the ATOVS data.

The mapping is a valuable step if radiances are not to be directly assimilated, because it allows a single profile of temperature and humidity to be retrieved using all instruments. However, even if radiances are assimilated directly, it still has some advantages. Firstly, microwave observations can be used to help detect cloud in the infrared observations (as described by English et al. (1999a)) and, secondly, the mapping may reduce noise by averaging several AMSU or MSU observations. This is of particular value if the observations are not to be assimilated at full resolution (e.g. because the computing cost is prohibitive).

The AAPP also has simple tests to detect which AMSU observations are affected by cloud and precipitation, and to determine which surface type to associate with the mapped observations (which primarily distinguishes land and sea ice from open water). These tests have been described by English et al. (1997) and have been validated for ATOVS by English et al. (1999b). The AAPP extracts a surface type from a high-resolution surface topography atlas and also retrieves a surface type from the AMSU window channels. The method of English et al. (1999a) is then applied to detect cloud-affected HIRS observations. Typically 80% of all fields of view (fovs) are classed as cloudy in the infrared, and for these fovs HIRS channels 1, 2, 3 and 12 are used. HIRS channels 1, 2 and 3 all peak above 200 hPa and HIRS channel 12, which is a water-vapour channel, peaks above 200 hPa in the tropics. If exceptionally cold cloud is detected, HIRS channels 3 and 12 are also rejected. Exceptionally cold cloud is defined as being where the brightness temperature measured by HIRS channel 8 is 55 K colder than the temperature calculated from the model under the assumption of cloud-free conditions. For AMSU observations 20–50% of fovs are classed as cloudy, which means that the liquid-water content of the cloud is probably in excess of 100 g m⁻². The rejection rate is lowest in the tropics (20%) and highest in polar regions (50%). When cloud is detected in the microwave-window channels, AMSU channels 6–14 (peaking at 400 hPa and above) can still be assimilated. AMSU channels 6 to 8 (which peak at 400, 250 and 150 hPa, respectively) are rejected for 5% of observations, as determined by tests using a simple scattering-index approach (English et al. 1997) that is sensitive to scattering from deep layers of ice cloud. This test flags fewer than 1% of the observations in the arctic, 1–2% in middle-latitude regions and 5–10% in the tropics.

The final stage of radiance preprocessing prior to assimilation is bias correction. Radiance measurements are not absolute and often have systematic biases. Sometimes the forward model used to simulate the radiances from the model profiles of temperature and humidity may also be biased. Radiative-transfer biases have the same effect as instrument biases, but are more prone to varying with air mass. Bias correction of radiances is a complicated issue and it is beyond the scope of this paper to discuss it in detail. The method of Eyre (1992) is followed, except that AMSU channels 5 and 9 are used to predict the bias for ATOVS, whereas MSU channels 2, 3 and 4 are used for TOVS. AMSU channel 5 is almost equivalent to MSU channel 2, both peaking close to 700 hPa, and AMSU channel 9 is very similar to MSU channel 4, both peaking around 85 hPa. So the main difference in the bias correction for TOVS and ATOVS data is that no equivalent of the MSU channel-3 bias predictor is used for ATOVS (this channel peaks close to 300 hPa). The bias-correction predictors provided by Eyre’s method function as a measure of air mass, so that those observation biases that vary with air mass can be removed. All biases between the NWP model and the observations are assumed to be observation biases or observation-operator biases (i.e. the NWP model is assumed to be unbiased). The assimilation of radiosonde data prevents the NWP model
analyses from drifting and so, indirectly, the absolute calibration of the radiances is tied to radiosondes.

(b) 1D-var retrievals

The 1D-var system is similar to that described by Gadd et al. (1995). It minimizes a quadratic cost function ($J$) composed of an observation cost ($J^o$) and a background cost ($J^b$),

$$J(x) = J^o(x) + J^b(x) = \frac{1}{2}(y^o - H(x))R^{-1}(y^o - H(x))^T + \frac{1}{2}(x - x^b)B^{-1}(x - x^b)^T,$$  

(1)

where the notation of Ide et al. (1998) has been used. The symbols have the following meaning: $y^o$ are the observed radiances; $x^b$ are the background values; $H(x)$ is the observation operator used to calculate top-of-atmosphere radiances, given the profile vector, $x$; $R$ is the sum of the observation and forward-model error-covariance matrices; $B$ is the background-error covariance matrix; and the superscript $T$ denotes the matrix transpose.

The background profile consists of the temperature on 40 levels, the humidity on 20 levels, the 1.5 m temperature and the surface skin temperature. In addition the 10 m wind speed and the 1.5 m humidity values are specified by the NWP model, but not analysed in the 1D-var. The microwave ocean emissivity is calculated from the skin temperature and the 10 m wind speed using the model described by English and Hewison (1998). The infrared and microwave land and sea-ice emissivities are all taken to be fixed constants. The total column ozone, which has a small effect on HIRS channels 3 to 9, is estimated from the 70 hPa model temperature using coefficients derived by comparing the ozone retrievals distributed by NESDIS with the 70 hPa temperatures between 1993 and 1998. For heights above the model top (10 hPa) a simple regression-based retrieval is used to provide values for $x^b$, using data from the SSU, HIRS channel 1 and MSU channel 4 for TOVS retrievals, or HIRS channel 1 and AMSU channels 10 to 14 for ATOVS retrievals.

The forward model $H(x)$ is calculated using either the radiative-transfer model described by Eyre (1991), or an updated version of the scheme described by Saunders et al. (1999). The Eyre model is referred to as RTTOV-2 and the Saunders model as RTTOV-5; RTTOV-2 is used to simulate TOVS data, whereas RTTOV-5 is used to simulate ATOVS data. If ATOVS and TOVS observations are processed together, RTTOV-5 is used for both. RTTOV-5 has several advantages when compared with RTTOV-2: more accurate water-vapour and ozone transmittances; variable sea surface emissivity (English and Hewison 1998); and the use of a more globally-representative set of profiles to generate the fast-model coefficients (i.e. the updated version of RTTOV gives accurate results for a more extreme range of profiles). The size of these improvements in RTTOV-5 relative to RTTOV-2 has been discussed in more detail by Saunders et al (1999), except for the ocean emissivity model. There was no previous emissivity model in RTTOV-2 to compare with, but the emissivity model agrees with aircraft observations at 23.8, 50.1, 89 and 150 GHz to within 1.0–1.5% (English and Hewison 1998).

(c) The assimilation of ATOVS soundings in 3D-var

The retrieval increments (differences between the background $x^b$ and the final retrieved profile at convergence—usually taking 2–3 iterations) are interpolated back
to the levels of the NWP model, and this information is used by the 3D-var assimilation system to improve the new analysis. In this study all satellite observations are treated as if they were taken at the main synoptic times (0000, 0600, 1200 and 1800 UTC, i.e. all satellite observations between 2100 and 0300 UTC are assimilated as if they were measured at 0000 UTC, for example). It is not known how important this assumption is. As the cost of assimilating the full-resolution data is very high, and as the observations have horizontally correlated errors, the observations are thinned such that one sounding is presented to the assimilation in every 2°× 2° latitude–longitude grid box (this is similar to the analysis resolution). Clear soundings are selected preferentially to cloudy soundings for both TOVS and ATOVS data to allow use of both HIRS and AMSU (or MSU) data wherever possible. No account is taken of vertically or horizontally correlated errors in the assimilation of the ATOVS (and TOVS) soundings. To allow for this, the assimilation was tested for TOVS data with observational errors higher than would be estimated using the diagonal of the Hessian matrix. Three tests were completed with the observational errors multiplied by 1.0, 3.3 and 10.0. Lorenc et al. (2000) found that the increase of 3.3 gave the best result in terms of forecast accuracy. The same increase in observational error was used for ATOVS.

3. Evaluation of ATOVS in Orbit

(a) Using the NWP model as a reference

The quantity $y^o - H(x^b)$ is referred to as the ‘innovation’. In Fig. 1 the standard deviation of the innovations is compared with with noise-equivalent temperatures (NEΔT) for AMSU channels 4–11. The NEΔT is that measured post-launch
Figure 2. Plots (a) and (c) of the standard deviation of the innovations as a function of latitude and AMSU channel number, and (b) and (d) of the variation with latitude of the standard deviation of the innovations for AMSU channel 18 (dashed line) and HIRS channel 12 (continuous line) in July 1999 and January 2000.

(Atkinson, personal communication). Note, for several channels these are lower than the pre-launch measurements described by Saunders (1993). Figure 1 also shows the quantity $\sqrt{H'(x)BH'(x)}$, which we call the ‘background radiance error’ as it is simply the background error expressed in radiance space. $H'$ is the derivative of $H(x)$ with respect to $x$, often referred to as the Jacobian. In Fig. 1 the background radiance error is shown for two different $B$-matrices. The $B$-matrix used by Gadd et al. (1995) is the one currently used at The Met. Office and is denoted ‘Gadd95’. The equivalent using the $B$-matrix presented by Eyre (1993), denoted ‘Eyre93’, is also shown. These two $B$-matrices give very different background radiance errors because of the large difference in the size of the diagonal elements, especially in the stratosphere. The Gadd95 matrix has background radiance errors well above the standard deviation of the innovations and is, therefore, overestimating the background temperature errors. The Eyre93 matrix has background error variances that are closer to the observed standard deviations of the innovations. For AMSU channels 5 and 6 the $\text{NE}\Delta T$ is more than 50% of the global standard deviation of the innovations. This comparison implies that the background errors must be lower than those described by the Gadd95 matrix, and that the Eyre93 matrix appears closer to truth for AMSU channels 5 and above. For AMSU channel 4 both the Eyre93 and Gadd95 matrices overestimate the background error.

Figures 2(a) and (c) shows the standard deviations of the innovations for AMSU channels 4 to 13 as a function of latitude during July 1999 and January 2000. The AMSU channels increase in altitude from channel 4 (950 hPa) to 13 (5 hPa). In the stratosphere the standard deviation of the innovations is highest (0.5–0.6 K) in the
winter extratropics, lowest (0.25–0.35 K) in the summer extratropics, and in the range 0.3–0.4 K in the tropics. Note that, for AMSU channel 13, the standard deviation of the innovations is small because the parts of $\mathbf{x}^b$ that affect the calculated radiances are a regression retrieval from the measurements and are not obtained from the model’s short-range forecast. In the troposphere the standard deviation of the innovations is always lowest (0.2 K) in the tropics, and highest (0.35–0.40 K) in the extratropical southern hemisphere during the southern hemisphere winter. In the northern hemisphere winter there is no significant difference between the size of the innovations north and south of the equator. In conclusion AMSU-A observations have considerable opportunity for impact in the stratosphere, especially in the winter hemisphere and in the southern hemisphere.

Figures 2(b) and (d) compare the standard deviations of the innovations for AMSU channel 18 and HIRS channel 12, which are both moisture channels sensitive to upper-tropospheric humidity. In the northern hemisphere summer the largest innovations are found in the extratropical southern hemisphere. By contrast, in the southern hemisphere summer the values are comparable in the two extratropical regions and peak in the tropics. Apart from the relatively large innovations in the tropics, the pattern is similar to AMSU-A in the troposphere. The larger innovations in the tropics are more obvious in terms of bias (not shown) where there are large positive biases in the tropics (i.e. the observation implies less moisture than the model profile). It is noticeable that the innovations are larger for AMSU channel 18 than for HIRS channel 12, but this is true only for the random component of the difference.

There are some regions where persistently higher standard deviations of the innovations are found. We interpret these local increases as higher errors in the NWP model.
six-hour forecast. The standard deviation of the innovations were particularly high over north and east Asia in January 1999, as illustrated with AMSU channel 8 in Fig. 3. The cause of the higher than average NWP model errors over east Asia is believed to be due to subgrid-scale orographic (SGO) effects (Davies, personal communication). We know there is sensitivity to tuning of SGO effects, but there are many parameters and interactions involved and so optimizing the NWP model is difficult. We know that in this region the NWP model error is worse when there is a strong flow across the Himalayan region (flows in this region are not usually as strong in the summer months); there are insufficient conventional radiosonde observations to correct the error that results.

(b) Comparison of innovations in clear and cloudy conditions

In cloud-free conditions HIRS provides more temperature information than AMSU-A (Eyre 1990). However AMSU-A has the advantage that cloud detection is simpler (English et al. 1997) and less data are rejected globally. In this analysis AMSU observations were subdivided into clear and cloudy datasets using the method described by English et al. (1997). The HIRS observations were subdivided into clear and cloudy datasets using the method of English et al. (1999b). Fewer AMSU observations than HIRS observations were classified as cloudy. For upper-tropospheric temperature sounding channels 5% of the AMSU data and 80% of the HIRS data were classified as cloudy. For channels peaking at lower altitude there were three times more clear microwave observations than infrared observations.

Figure 4 shows the standard deviation of the innovations for three different cloud categories for November 1999. These cloud categories are: ‘clear’, for which neither
HIRS nor AMSU detect any cloud; ‘microwave-clear’, for which HIRS detects cloud but AMSU does not; ‘cloudy’ for which both HIRS and AMSU detect cloud. The number of observations in these three categories is, typically, 20% clear, 40% microwave-clear and 35% cloudy. For 5% of the observations, cloud is detected in the microwave channels but not in the infrared channels. This occurs for low-altitude broken marine stratocumulus where the cloud-top temperature is close to the sea surface temperature. For AMSU channels 6 to 10 (which peak at regular intervals from 750 hPa to 50 hPa) the clear innovations typically have a standard deviations of 0.3–0.5 K, whereas the microwave-clear innovations have a standard deviations of 0.5–0.7 K. For AMSU channels 6 to 10 there is also no significant difference between the standard deviations for microwave-clear and cloudy innovations. Therefore, the lower standard deviation for the clear sample is believed to be due to systematically lower errors in the background temperature profile. For AMSU channels 4 and 5 the effect of liquid-water emission increases the standard deviation for cloudy innovations compared with the microwave-clear sample. The standard deviation of cloudy innovations for AMSU channel 4 is nearly 2 K. Therefore, it is not possible to reach the same conclusions for AMSU channels 4 and 5 as for AMSU channels 6 to 10. These large errors are mostly found in the extratropical winter storm-track regions. The systematically higher standard deviation of the microwave-clear innovations, compared with the clear innovations for all channels, is also believed to be caused by higher errors in the background.

A study by Yu et al. (1997) also found that NWP model errors are highest in cloudy areas. If this is the case then the relative value of HIRS and AMSU-A cannot be judged on their clear-air information content alone. The relative importance of observations in clear and cloudy areas and the difficulties in solving the cloud detection problem are also important. Therefore, comparing the relative impact of HIRS and AMSU in data assimilation experiments is the most effective means of comparing their value. This is done in section 4.

(c) The radio-frequency interference problem for AMSU-B

As reported by Atkinson and Mclellan (1998), AMSU-B suffers from radio-frequency interference (RFI). This is caused by AMSU-B detecting the signals sent by the transmitters used for downloading data to ground stations, and by the search and rescue transmitter. Initially this caused a large, but constant, interference pattern. However, the problem became more severe when the helix transmitters themselves developed a fault which resulted in a less than perfect connection, leading to variable power output. This variable power output causes different interference patterns in AMSU-B data. Corrections for the radio-frequency interference have been developed for each individual interference ‘mode’ (Atkinson, personal communication). The AMSU-B observation statistics in this paper use the set of AMSU-B RFI corrections selected by automatic interference-mode tracking in the AAPP. On 28 September 1999 the transmitters which had been causing the variable interference pattern were switched off. Since that time AMSU-B data have been of useable quality for data assimilation, although no AMSU-B data were assimilated in any of the experiments described in this paper.

4. THE IMPACT OF ATOVS OBSERVATIONS ON THE ACCURACY OF NWP

The impacts of TOVS and ATOVS observations were evaluated from the results of five experiments. Firstly, an experimental forecast suite was run for four weeks in which NOAA-11 TOVS data were replaced by NOAA-15 ATOVS observations; in this paper we refer to this run as NEWATOVS (see Table 1). A control experiment was run
TABLE 1. TOVS CHANNELS USED IN 1D-VAR FOR EXPERIMENTS OLDTOVS AND NEWATOVVS (SEE TEXT)

(a) Experiment OLDTOVS

<table>
<thead>
<tr>
<th>Result of cloud test</th>
<th>Infrared</th>
<th>Microwave</th>
<th>Surface type</th>
<th>Levels assimilated in 3D-var</th>
<th>NOAA-11/14</th>
<th>NOAA-15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MSU</td>
<td>SSU</td>
<td>HIRS</td>
</tr>
<tr>
<td>Clear</td>
<td>Clear</td>
<td>Sea</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-8, 10-15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land</td>
<td>Above 100 hPa</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td>Cloud</td>
<td>Clear</td>
<td>Sea</td>
<td>Above 100 hPa</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land</td>
<td>Above 100 hPa</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
</tbody>
</table>

(b) Experiment NEWATOVVS

<table>
<thead>
<tr>
<th>Result of cloud test</th>
<th>Infrared</th>
<th>Microwave</th>
<th>Surface type</th>
<th>Levels assimilated in 3D-var</th>
<th>NOAA-11/14</th>
<th>NOAA-15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MSU</td>
<td>SSU</td>
<td>HIRS</td>
</tr>
<tr>
<td>Clear</td>
<td>Clear</td>
<td>Sea</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-8, 10-15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land</td>
<td>Above 400 hPa</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>'Siberia'</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-6</td>
</tr>
<tr>
<td>Low cloud</td>
<td>Clear</td>
<td>Sea</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3, 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land</td>
<td>Above 400 hPa</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>'Siberia'</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td>High cloud</td>
<td>Clear</td>
<td>Sea</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land</td>
<td>Above 400 hPa</td>
<td>2-4</td>
<td>1-3</td>
<td>1-2, 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>'Siberia'</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-2</td>
</tr>
<tr>
<td>Any cloud</td>
<td>Cloud</td>
<td>Sea</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land</td>
<td>Above 400 hPa</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>'Siberia'</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td>Any cloud</td>
<td>Scattering</td>
<td>Sea</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Land</td>
<td>Above 400 hPa</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>'Siberia'</td>
<td>All levels</td>
<td>2-4</td>
<td>1-3</td>
<td>1-3</td>
</tr>
</tbody>
</table>

Note: no channels were used for surface elevations above 1000 m, and over sea ice nothing was assimilated below 950 hPa in either TOVS or ATOVS systems.

for the same period which replicated the operational use of TOVS radiances to ensure that differences could only arise from the change in the satellite soundings; we refer to this run as OLDTOVS*. The NOAA-14 TOVS data were assimilated in both the NEWATOVVS and OLDTOVS runs. Over land the 1D-var retrievals were assimilated only at pressure levels above 100 hPa in the OLDTOVS run. In the NEWATOVVS run this was lowered to 400 hPa, allowing the observations to make much more significant changes to the analysis over land. This additional information would be expected to give most impact in the northern hemisphere simply because there is more land. In the NEWATOVVS run RTTOV-5 was used to compute radiances from the model profiles, whereas RTTOV-2 was used in the OLDTOVS runs.

* In later tests it was discovered that some radiosonde observations in Asia and South America, which were assimilated in the OLDTOVS control, were not assimilated in the NEWATOVVS experiment (Bell, personal communication). The complete radiosonde network gives a large positive impact, but the loss of a few radiosondes usually gives only a small impact. In this experiment this may have degraded the ATOVS trial results by around 0.5%. Therefore, the positive impact of ATOVS data is very slightly, but not significantly, underestimated. The 0.5% difference is small compared with the synoptic variability of the impact of the new observations.
In order to measure the absolute impact of the TOVS and ATOVS observations, a third run of the forecast suite was completed for the same time period in which no TOVS or ATOVS observations were assimilated; we refer to this run as the NOTOVS run. The forecast scores for the NEWATOVS run were compared with those from the NOTOVS and OLDTOVS runs. The relative impact of HIRS and AMSU-A in the extratropics was assessed by running the forecast suite for a fourth run with HIRS observations excluded from the assimilation except in the tropics*; we refer to this as the NOHIRS run. HIRS data were retained in the tropics because past experiments (e.g. McNally and Vesperini 1996) have shown the importance of satellite moisture information in that region. Finally, a fifth run of the forecast suite is described in section 5 in which additional channels were assimilated over Siberia; we refer to this as the SIBERIA run. The first four experiments were run only for the period 12 January to 9 February 1999. However, the SIBERIA experiment was also run in March and June 1999. RTTOV-5 was used in the SIBERIA and NOHIRS experiments.

(a) Verification of NWP forecast accuracy

The accuracy of NWP forecasts can be measured in many different ways. Firstly, there is the choice of the measures of 'truth' against which forecast skill is judged; two were used in this study. The first uses observations to compare with the forecast results; the observations used were the standard synoptic radiosonde and surface observations of temperature, pressure, humidity and wind speed. The second used the analysis fields generated by the NWP system to represent the truth. In all the verification results discussed in this paper the forecasts were verified against the analyses produced by assimilation cycles based on the same experimental setups, i.e. NEWATOVS forecasts were verified against NEWATOVS analyses). The forecast errors are not independent of the errors in either verification dataset. For this reason subjective verifications of forecast accuracy by an experienced weather forecaster were carried out to complement the objective verification.

For the objective verification there are different methods of assessing how well the forecast fits the chosen measure of truth. The mean fit and the root-mean-square error (RMS error) are easily understood. However, the RMS error can be misleading if the variance of the model fields is much higher for one forecast than another. To try to overcome this two additional quantities were examined. These were (1) the anomaly correlation (AC), which measures the correlation between anomalies in the forecast and the analysis, and (2) the skill-score which normalizes the change in the forecast RMS error with the RMS error of a 'forecast' based on persistence. For the experiments described in this paper verification using RMS error, AC and skill-score usually gave similar results, and so the results shown in this paper are for the RMS error, unless stated otherwise. The change in the RMS error can either be expressed as a percentage reduction in the RMS error or as an equivalent gain in forecast time (i.e. the new forecast has the same RMS error as the old forecast had at an earlier time). In this study, verification against analyses and against observations usually gave similar results. This paper focuses on verification against analyses, only mentioning verification against observations where a significantly different result was obtained.

* In this paper the 'tropics' are defined to cover the area between 18.75°S and 18.75°N, with the 'northern hemisphere' occupying the area north of 18.75°N and the 'southern hemisphere' the area south of 18.75°S. This definition of the tropics is the standard one used for NWP verification by The Met. Office.
(b) Comparison of OLDTOVS and NEWATOVS forecasts

The NEWATOVS forecasts were expected to be more accurate in the extratropical southern hemisphere and to provide modest benefits in the extratropical northern hemisphere. Since no new humidity information was being used, no significant difference was expected in the tropics.

Bias corrections for both NOAA-14 TOVS and NOAA-15 ATOVS radiances were recalculated after one week of the experiment. These corrections were applied to the observations from 21 January to the end of the trial period on 9 February. Figure 5 shows the change in the mean state and RMS fit to the observations for three moisture channels (HIRS channels 10 and 11 and AMSU channel 18) and three temperature channels (AMSU channels 6, 9 and 11). The period shown is the first ten days following the first operational assimilation cycle using the configuration tested in the NEWATOVS experiment. The short-range forecast pulls the assimilation much closer to the observations for the moisture channels, especially HIRS channel 11, in just two days of assimilation. There is no obvious lowering of the random difference between the short-range forecast and the observations for the tropospheric temperature-sounding channels shown (AMSU 6 and 9). However, for the stratospheric temperature channel (AMSU 11), the short-range forecast fits the observations more closely after two to three days of assimilation. Several channels see large changes in bias, amongst the results shown in Fig. 5 the most obvious being AMSU channel 9. Figure 5 shows that the system had mostly responded to the change in data use, as measured by the fit to the radiances, within two to three days. However, the bias for AMSU channel 6 took four to five days to settle down. The bias corrections were recomputed one week into the experiments to allow for changes in model bias (results quoted in this paper refer to the whole trial period, unless indicated otherwise). The impact for the period after the bias correction was considerably more positive than for the first week.

(i) Impact of ATOVS in the southern hemisphere. In the southern hemisphere the RMS error (verified against analyses) of the mean-sea-level pressure was 7.3%, 11.1%, 10.1%, 7.4% and 4.9% lower in the NEWATOVS run than in the OLDTOVS run at days 1, 2, 3, 4, and 5, respectively. Similar reductions were found for other key meteorological variables (e.g. the 500 hPa height and the 250 hPa winds). Reductions in the errors verified against observations were very similar, being slightly smaller for the mean-sea-level pressure but exceeding 20% for the 500 hPa height. The impact was variable from day to day. The impact was largest for those days when forecast errors were largest (see Fig. 6). Considering all the individual verifications for the mean-sea-level pressure and 500 hPa height, positive impacts outnumbered negative impacts in the ratio three to one (where positive and negative are defined by a change of more than 1% in anomaly correlation). The errors in the individual forecasts are correlated, and so the number of independent verifications is less than the total number of verifications. The number of positive, neutral and negative impacts for 22 different forecast parameters is shown in Fig. 7 (which also shows these numbers for the tropics and southern hemisphere). The largest impacts of ATOVS data were for those forecasts which were particularly poor when ATOVS observations were omitted. For example on 23 January 1999 the control had an RMS error of over 70 m in the 500 hPa height, whereas the experiment had an RMS error of less than 30 m when verified against radiosonde observations.

(ii) Impact of ATOVS in the northern hemisphere. ATOVS observations also improved forecasts in the northern hemisphere although, as expected, the impact was smaller than in the southern hemisphere. The RMS errors for mean-sea-level pressure were 3.8%,
Figure 5. Time series of mean innovations before and after bias correction and its standard deviation for three moisture channels (HIRS channels 10 and 11 and AMSU channel 18) and three temperature channels (AMSU channels 6, 9 and 11). The continuous line denotes the mean innovation after bias correction, the dotted line the mean innovation before bias correction and the two dash-dot lines denote ± one standard deviation of the innovations around the mean. No bias correction was applied to AMSU channel 18.
Figure 6. The anomaly correlation of the mean-sea-level pressure in the southern hemisphere at day 3 of the forecast (the full line refers to forecasts from assimilations without ATOVS data and the dashed to forecasts from assimilations with ATOVS data).

3.9%, 1.2%, -0.9% and 0.1% lower in the NEWATOVS run than in the OLDTOVS run at days 1, 2, 3, 4, and 5, respectively. The impact was more positive for the 500 hPa height, but again declined with forecast range: 6.6%, 5.1% and 1.8% at days 1, 2 and 3, and neutral thereafter. The verification against observations showed less of a decline in impact with increasing forecast range. The medium-range impact was most positive when compared with observations, whereas the short-range impact was most positive when compared with analyses. One possible explanation of this is that the short-range forecast in the NEWATOVS run had a stronger correlation with the analysis produced by its own assimilation cycle than the OLDTOVS run. However, this explanation seems unlikely to be valid as the TOVS and ATOVS 1D-var retrievals had equal weight in the 3D-var, but the ATOVS retrievals contain more radiance information and less background information. This would be expected to reduce the correlation between the forecast and analysis errors. For aviation flight-level winds, the forecast error was reduced by just under 2% when verified by observations and by just over 3% when verified by analyses. During this period, however, the impact was found to be very consistent. Considering all the individual verifications from one-day to six-day forecast ranges for the mean-sea-level pressure and 500 hPa height for all four weeks, 135 forecasts were improved, 50 were degraded and the remainder were neutral (as determined by a change of more than 1% in anomaly correlation). For the period after the bias correction update only five forecasts showed a negative impact, whereas 50 were positive (the rest being neutral) for one- and two-day forecasts.

(iii) Impact of ATOVS in the tropics. In the tropics a very modest fall in forecast error of around 0.5% to 1.0% occurred for winds at all heights when verified against analyses, but when verified against observations a rise in error of a similar magnitude
Figure 7. The percentage of forecasts showing positive, neutral and negative impacts to the inclusion of ATOVS data over a four week period, as assessed from a selection of important meteorological parameters which are identified by the numbers along the x-axis: 1 to 6—northern hemisphere (NH) mean-sea-level pressure at 24-hour intervals from the 24–144 hour forecasts; 7 to 12—the same as 1 to 6 but for the southern hemisphere (SH); 13 to 18—the NH 500 hPa height at 24-hour intervals from the 24–144 hour forecasts; 19 to 24—the same as 13 to 18 but for the SH; 25—the 250 hPa wind speed for the NH; 26—the same as 25 but for the SH; 27—the same as 25 but for the tropics; and 28–30 the 850 hPa wind speed in the tropics at 24-hour intervals from the 24–72 hour forecasts. The criterion for improvement (or degradation) is a change of more than 1% in anomaly correlation.

occurred. However, verification of short-range temperature forecasts in the tropics by aircraft flying at between 350 and 250 hPa showed that NEWATOVS temperatures were 0.1 K lower than OLDTOVS temperatures in the tropics. The RMS error was 10% higher in NEWATOVS than in OLDTOVS (at higher and lower altitudes this degradation was not observed). This degradation was thought to be caused by inconsistencies in the processing of TOVS and ATOVS observations, and further experiments have shown that assimilating either only ATOVS or only TOVS data gave a more positive impact than assimilating both TOVS and ATOVS data in the tropics. It is not possible to say whether the assimilation of TOVS radiances processed from level 1B data using the same system as ATOVS would have overcome this problem.

(c) Comparison of NEWATOVS and NOHIRS forecasts

An experiment was completed for the same 12 January to 9 February 1999 period without using HIRS data in the extratropics (HIRS data were still used in the tropics because humidity information is important in that region). The impact of HIRS data was neutral in the northern hemisphere but the RMS errors of the forecasts were typically 1% lower for NEWATOVS than for NOHIRS in the southern hemisphere. At longer ranges the impact of HIRS data was neutral. This experiment was repeated for March 1999 and the impact of HIRS data was again found to be neutral in the northern hemisphere, but NEWATOVS gave RMS errors 3% to 5% lower than NOHIRS in the southern hemisphere. Consequently, the impact of HIRS data was variable, but the impact of AMSU data was much larger than the impact of HIRS data. Given the apparent
advantage of HIRS radiances in terms of \( \Delta T \), this implies that either the forward-model errors were much higher for HIRS data or that the satellite data have more impact when assimilated in cloudy areas. The evidence from the monitoring of the innovations discussed in section 3 is that there is a correlation between higher forecast errors and cloudiness. It seems likely that the benefit of the assimilation of AMSU-A data arises from its near all-weather capability. However, as HIRS data did have a positive impact, albeit a small one, the lower \( \Delta T \) and the sharper weighting functions of HIRS must be providing some additional skill, even though only in clear areas. Nonetheless, it is clear that the major impact of ATOVS and TOVS comes from the information provided by the MSU and AMSU instruments.

(d) Comparison of NEWATOVS, OLDTOVS and NOTOVS experiments

Table 2 lists a selection of parameters illustrating the improvement of the OLDTOVS and NEWATOVS forecasts over the NOTOVS forecast. For some parameters the RMS error was nearly 30% lower in NEWATOVS than in NOTOVS; a fall in RMS error of this size is equivalent to gaining more than a day in forecast range (i.e. two-day forecast errors without the data are comparable with one-day forecast error with the data). Comparing the OLDTOVS run with the NOTOVS run, some of the parameters shown in Table 2 gave a negative impact when measured against analyses. However, in the comparison of NEWATOVS with NOTOVS there are no negative impacts.

For the 500 hPa height the impact of TOVS and ATOVS data in the southern hemisphere is in excess of one day of forecast range, and in the northern hemisphere around a quarter of a day. If the reduction in forecast RMS error is averaged over the range of parameters listed in Fig. 7 the impact of having data from one satellite carrying ATOVS and one satellite carrying TOVS is about twice that of having data from two satellites carrying TOVS. However, a little caution must be taken in interpretation of these results. Firstly, the TOVS instrument on NOAA-11 was old and pressed into service when the HIRS instrument on NOAA-12 failed. Not surprisingly, NOAA-11 was noisier than NOAA-14 and the MSU instrument failed on NOAA-11 shortly after the completion of the experiments. Furthermore, the increased use of the radiance information over land almost certainly contributed to the skill of the NEWATOVS forecasts, as did the use of a new and more accurate radiative-transfer model. Therefore, some of the impact arises from the new AMSU channels on ATOVS and some from changes made to the processing system. Nonetheless, further improvement is expected when NOAA-L is launched and a two-satellite system carrying ATOVS is available to NWP centres for the first time.

The NEWATOVS and NOTOVS forecasts were also compared subjectively for the northern hemisphere by an experienced forecaster. For medium-range forecasts the subjective verification supported the objective verification. At shorter forecast ranges, particularly for one-day forecasts, the subjective verification was less positive than the objective verification. The largest analysis differences were over north-east Asia and Greenland (N.B. over Greenland the increments were implied by the horizontal correlation functions in 3D-var, as no ATOVS data were assimilated where the surface elevation was greater than 1000 m). The large increments over north-east Asia could, on several occasions, be traced forward to large improvements in forecast accuracy in the Atlantic and over Europe, emphasizing that significant analysis increments can be introduced over land by assimilating ATOVS above 400 hPa, and that this modified analysis can produce more accurate forecasts. The largest difference seen was in the analysis of the westward extent of the Alaskan vortex into north-east Asia on 6 February when NEWATOVS analysed heights some 100 m higher at 500 hPa than NOTOVS.
### TABLE 2. PERCENTAGE REDUCTION IN THE ROOT-MEAN-SQUARE ERRORS OF 24-HOUR AND 72-HOUR PREDICTIONS DURING THE PERIOD 12 JANUARY TO 9 FEBRUARY 1999 WHEN EXPERIMENTS OLDTOVS AND NEWATOVS ARE COMPARED WITH THE CONTROL NOTOVS, AND WHEN EXPERIMENT NEWATOVS IS COMPARED WITH OLDTOVS (SEE TEXT AND TABLE 1 FOR DESCRIBATIONS OF THE EXPERIMENTS)

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Control</th>
<th>MSLP</th>
<th>Z500</th>
<th>U250</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T+24</td>
<td>T+72</td>
<td>T+24</td>
<td>T+72</td>
</tr>
<tr>
<td>Verification against analyses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLDTOVS NOTOVS</td>
<td>-1.0</td>
<td>3.3</td>
<td>-5.3</td>
<td>5.2</td>
</tr>
<tr>
<td>NEWATOVS NOTOVS</td>
<td>2.8</td>
<td>4.5</td>
<td>1.3</td>
<td>7.0</td>
</tr>
<tr>
<td>NEWATOVS OLDTOVS</td>
<td>3.8</td>
<td>1.2</td>
<td>6.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Verification against surface and radiosonde observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLDTOVS NOTOVS</td>
<td>1.9</td>
<td>6.5</td>
<td>1.4</td>
<td>6.8</td>
</tr>
<tr>
<td>NEWATOVS NOTOVS</td>
<td>2.4</td>
<td>7.1</td>
<td>4.4</td>
<td>8.1</td>
</tr>
<tr>
<td>NEWATOVS OLDTOVS</td>
<td>0.5</td>
<td>0.6</td>
<td>3.0</td>
<td>1.3</td>
</tr>
</tbody>
</table>

(b) Southern hemisphere 18.75°–90°S

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Control</th>
<th>MSLP</th>
<th>Z500</th>
<th>U250</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T+24</td>
<td>T+72</td>
<td>T+24</td>
<td>T+72</td>
</tr>
<tr>
<td>Verification against analyses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLDTOVS NOTOVS</td>
<td>21.0</td>
<td>3.8</td>
<td>21.7</td>
<td>4.7</td>
</tr>
<tr>
<td>NEWATOVS NOTOVS</td>
<td>28.3</td>
<td>13.9</td>
<td>28.1</td>
<td>12.6</td>
</tr>
<tr>
<td>NEWATOVS OLDTOVS</td>
<td>7.3</td>
<td>10.1</td>
<td>6.4</td>
<td>7.9</td>
</tr>
<tr>
<td>Verification against surface and radiosonde observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLDTOVS NOTOVS</td>
<td>14.2</td>
<td>1.4</td>
<td>7.8</td>
<td>-3.6</td>
</tr>
<tr>
<td>NEWATOVS NOTOVS</td>
<td>17.3</td>
<td>2.6</td>
<td>24.8</td>
<td>16.8</td>
</tr>
<tr>
<td>NEWATOVS OLDTOVS</td>
<td>3.1</td>
<td>1.2</td>
<td>17.0</td>
<td>20.4</td>
</tr>
</tbody>
</table>

N.B. Positive values indicate reductions and negative values indicate increases in the root-mean-square errors.

MSLP: mean-sea-level pressure;
Z500: 500 hPa height;
U250: 250 hPa wind.

This appears to have had beneficial effects both downstream and in the same region at short forecast ranges. NOTOVS analysed the vortex to be too deep and the one-day forecast from this data-time carried forward these low values, thereby comparing poorly with the next analysis. By day two of the forecast, marked improvements were seen for ATOVS further downstream, reaching Japan. By day four of the forecast, this benefit had propagated into the Pacific. For the small sample of degraded forecasts, an attempt was also made to trace the forecast differences to analysis increments, but no clear link to any analysis increments was established.

### 5. IMPACT OF ASSIMILATING ATOVS OVER SIBERIA

There are many good reasons why centres have been more cautious about the the assimilation of radiances over land than has been the case over the sea. Some of these reasons are common to all NWP centres, e.g. the availability of other observations over land (and the associated danger of errors in satellite data interpretation degrading an otherwise accurate analysis), and the difficulties of modelling surface emissivity and of identifying cloud and rain effects on measured brightness temperatures. Other reasons
depend to a greater or lesser extent on details of different centres' own assimilation methods, such as the method of bias correction. In this study all observations were rejected where the surface elevation exceeded 1000 m because of difficulties with the bias correction technique.

Siberia has several features which make it a good test-bed for the assimilation of HIRS and AMSU radiance information. Firstly, it is mostly low lying. Secondly, the emissivity over the boreal forest is homogeneous during most of the year and usually takes one of two values, one corresponding to dry snow in winter and the other to snow-free conditions in summer. By contrast, during the spring-melt period the emissivity can become highly variable, but this period is short-lived. An impact trial was run during the spring-melt period to note any detrimental effects. A mean emissivity atlas generated from SSM/I data (Prigent et al. 1997) for the month of July shows that the summer emissivity is close to 0.95 in the test region (30°–130°E, 50°–70°N). Examination of mean differences of the observations minus the background after bias correction in the test region for AMSU channels 4 and 5 showed biases no higher than those typically found over the sea. The spring-melt region could be clearly observed in AMSU 89 GHz images to the west of the region during the January–February test period, progressing steadily eastwards during spring. English (1999) showed that, if emissivity is known to within 1–2%, useful temperature information can be obtained from ATOVS radiances over land. The third motivation for assimilating ATOVS in the Siberia region is that there are very few radiosondes reporting regularly, and those that do provide meteorological data are concentrated in the west of the test region. The absence of other observations, combined with the parametrization of subgrid-scale orographic effects discussed in section 3, gives rise to unusually high short-range forecast errors. Lastly, as has already been shown in section 4(d), increments in the analysis over north-east Asia arising from assimilating ATOVS above 400 hPa can improve forecasts.

The experiment was to assimilate 1-D-var retrievals from ATOVS at all model levels using AMSU channels 4 to 14 and HIRS channels 1 to 6. This configuration is compared in Table 1 with that used operationally. The main difference for land surfaces is the assimilation of information from the surface to 400 hPa, this being made possible by the use of AMSU channels 4 and 5 and HIRS channels 4, 5 and 6. These additional channels were only used when the cloud test described by English et al. (1999b) detected no cloud. A very conservative cloud threshold was used to detect as many cloudy fovs as possible. The rejection rate was over 85%. This configuration was tested in winter (12 January to 9 February 1999), spring (4 March to 26 March 1999) and summer (7 to 17 June 1999). It was considered necessary to conduct the test in three seasons because: in summer the surface is mostly free of snow and the microwave surface emissivity is high (at or above 0.95); in winter it is covered with dry snow which has a much lower microwave emissivity (as low as 0.7–0.8 at 50 GHz); and in spring the snow is melting and the emissivity is variable. In each period a control experiment using the same forecast-suite configuration as the NEWATOVS experiment was run.

In the January–February experiment, the short-range (1–2 day) RMS errors were 5–10% lower in the SIBERIA run than in the NEWATOVS run for mean-sea-level pressure and 500 hPa height in and just downwind of the test area. Thereafter this benefit propagated eastwards until at days 4 to 6 the northern hemisphere the improvement amounted to a 1–2% fall in the RMS error when verified against both observations and analyses. The 850 hPa temperature analysis in the test region was 0.2 K colder in the SIBERIA run than in the NEWATOVS run. In the spring and summer experiments the impacts were slightly smaller, especially when verified against analyses, the RMS error being lower in the SIBERIA run by 1.0%. However, verification against observations
at days 4 to 6 showed a very similar result to the winter experiment, with the RMS errors being 1–2% lower in the SIBERIA run. The use of ATOVS in this region helped to correct the effects of a systematic error in the analysis. There were insufficient conventional observations to have a comparable effect to that obtained by using ATOVS data. The impact of assimilating ATOVS observations over Siberia has been further validated by comparing how closely short-range forecasts fit the observations over Siberia in January 2000 compared with January 1999 (shown in Fig. 3). In January 2000 the short-range forecast fits the observations more closely over Siberia than over northern Canada, but in January 1999 the opposite was true. This gives some confidence that the assimilation improves the accuracy of short-range forecasts.

6. Conclusions

The assimilation of observations from TOVS on NOAA-14 and ATOVS on NOAA-15 was found to produce forecasts with significantly lower RMS errors and higher anomaly correlations than those achieved by assimilating TOVS on NOAA-11 and TOVS on NOAA-14. Averaged over a four-week period the largest impacts found amounted to a fall of more than 20% in the RMS error; some individual forecasts were improved by more than this. The impact was largest in the southern hemisphere. Compared with assimilating no TOVS or ATOVS observations, the impact of assimilating ATOVS and TOVS data was equivalent to a reduction of close to 30% in RMS error for 500 hPa height in the southern hemisphere at the one- to four-day forecast range. The impacts in the northern hemisphere were smaller, but still important, with RMS errors typically 5% lower when ATOVS observations were assimilated than when they were not. Further reductions in the RMS error of the forecasts in the northern hemisphere were achieved by assimilating more radiance information over Siberia.

These impacts are large. Why is such a large improvement seen? There are many aspects of ATOVS data that are beneficial compared with TOVS data. The AMSU-A is a very accurate instrument achieving an in-flight NEΔT of 0.1 K in the tropospheric temperature-sounding channels. The influence of cloud on microwave observations is much smaller on the AMSU-A temperature-sounding channels than on the equivalent HIRS channels. Furthermore, over the ocean, clouds with high values of liquid water are easily detected using the AMSU-A window channels. This allows far more microwave data than infrared data to be used, and these additional data are in cloudy areas where background errors tend to be highest (in the absence of new observations, if background errors are high the analysis errors will also be high). The experiments in January and February 1999 showed that, in the extratropics, AMSU data can add skill to a forecast system already using HIRS data, whereas HIRS observations add very little extra skill to a forecast system already using AMSU data. This is consistent with other experiments (e.g. Yu et al. 1997) that have demonstrated that the value of satellite data is greatest in active weather systems (which tend to be cloudy). Therefore, it is likely that a substantial part of the strong impact of the new microwave data arises from its near all-weather capability. Other improvements, such as a more accurate radiative-transfer model and the increased use of observations over land, will also have contributed to the measured benefit. In the northern hemisphere the increased use of satellite data over land may be the most important factor, as subjective forecast verification has traced the biggest improvements to analysis increments over north-east Asia. As AMSU-B data have not yet been assimilated, no comment can be made on the relative value of the HIRS and the AMSU-B humidity information in the tropics. However, HIRS humidity information has been demonstrated to be very important in the tropics (McNally and Vesperini 1996).
The impacts of assimilating TOVS radiances directly in 3D-var or 4D-var assimilation systems have been reported by Kelly (1997) and Derber and Wu (1998). These studies have shown a larger impact from TOVS data than has been reported in this paper. The relatively poor performance of TOVS data described in this paper is believed to be due to the assimilation of 1D-var retrievals rather than the direct assimilation of the radiances. The impact of assimilating radiances directly, instead of using 1D-var retrievals, has been described by Eyre et al. (2000) and, although it is beyond the scope of this paper to discuss this in detail, the results appear to confirm that direct radiance assimilation gives substantial new benefits. Further increases in the impact of this type of data are expected by extending the use of radiance assimilation over land, for which encouraging results have been reported here over Siberia. Despite the strong impact of AMSU-A observations, 40% of the radiance observations from AMSU channels 4 and 5 were classed as cloudy and, consequently, were not assimilated. If cloud effects could also be analysed there would be further potential for providing a substantial increase in temperature and humidity information for the analysis in areas of rapid synoptic development.

ACKNOWLEDGEMENTS

This work would not have been possible without the contribution of numerous staff in many sections of The Met. Office. We would particularly like to thank Nigel Atkinson, Stuart Bell, Andrew Lorenc, Dave Tinkler and Phil Evans and their teams. The radiative-transfer model, RTTOV-5, was provided by the European Centre for Medium-Range Weather Forecasts, and we are particularly grateful to Roger Saunders for his help. Terry Davies provided useful comments on the assimilations over Siberia.

REFERENCES


English, S. J. 1999 Estimation of temperature and humidity profile information from microwave radiances over different surface types. J. Appl. Meteorol., 38, 1526–1541


