Recent improvements to the ECMWF 4D-Var data assimilation

Reprinted from ECMWF Newsletter Number 81

The 4D-Var assimilation system (Rabier et. al. 1998) was introduced operationally at ECMWF on 25 November 1997 (see ECMWF Newsletter 78, winter 1997/98). Since that date several improvements to the use of both conventional and satellite data have been made as part of the Cycle 18R6 package.

For conventional data the main changes are the use of hourly surface observations from frequently reporting stations (instead of six hourly), and the use of radiosonde temperatures at all reported levels in place of geopotential height, winds, and humidities on standard pressure levels. These variables are now used at significant as well as standard levels. Also the 10m-wind observation operator was unified for all observation types.

For satellite data SSM/I total column water vapour is now assimilated, more of the TOVS radiances over land and the new high-density cloud motion winds are used and the assimilation of the scatterometer winds has been revised. These changes are described briefly below and more details can be found in the relevant ECMWF memoranda. At the same time the wave model was coupled to the atmospheric model; the implementation and impacts of this were described in Newsletter 80, summer 1998.

The overall impact of these changes on the forecasts was clearly beneficial when compared with the operational system at the time, as can be seen in Figure 1 for 64 cases averaged over both hemispheres for the period 16 April to 18 June 1998 when the new package was run as an e-suite in parallel to operations. As a result of the positive impacts on the forecasts shown in Figure 1 these changes were introduced into ECMWF operations on 29 June 1998.

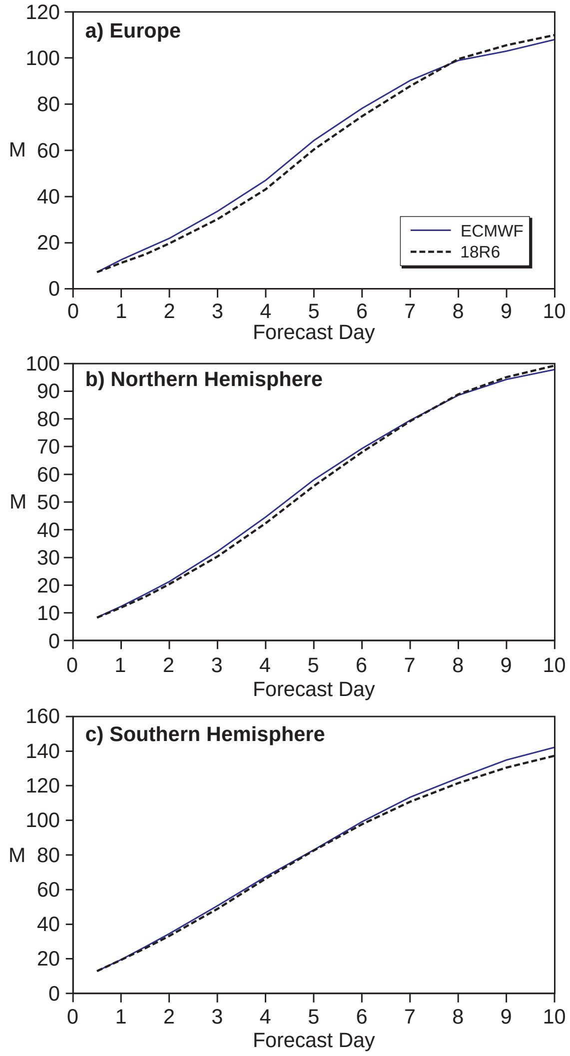

Fig. 1 500 hPa geopotential, root mean square of forecast error for Europe and the two hemispheres for the period 16 April to 18 June 1998. Both the 18R6forecasts and the corresponding operational forecasts are shown.

Use of hourly surface observations

There are many surface stations around the world (SYNOP and DRIBU) that report more than once every six hours. Most assimilation systems, however, use only one report from each station in each six-hour assimilation cycle. This has been the case in ECMWF's OI and 3D-Var systems as well as in its first operational 4D-Var system. The 4D-Var system has now been changed so that it makes use of this additional information.

In the observation screening for 3D-Var, observations are selected for assimilation in such a way that preference is given to the observations closest to the centre of the assimilation time window (3D-screening), i.e. closest to 00, 06, 12 and 18 UTC, for the 4 assimilation cycles. In 4D-Var, the model fields are made available hourly and the observations are accordingly organised into one-hour time slots. Observations are selected within each time slot independently. Preference is given to the observation closest to the centre of the time slot (4D-screening). The assimilation system has been prepared such that either 3D or 4D-screening can be performed in 4D-Var. The number of surface observations used increases with 4D-screening. Typically, the number of SYNOP surface pressure and wind observations increases by a factor of two, and for DRIBU observations by a factor of three. The number of TOVS radiances also increases with 4D-screening, as the satellite orbits overlap in the polar areas and the hourly thinning procedure retains more reports there.

With up to six observations from each station, 4D-Var with 4D-screening becomes vulnerable to isolated biased observations. This was found over the Antarctic region where the model first guess surface pressures can have large biases against some synoptic stations over the high orography and so cause deterioration of Southern Hemisphere forecast performance. We have addressed the bias problem by introducing serial correlation of observation error and by modifying the variational quality control of the time series of SYNOP and DRIBU surface pressure and height data. Serial correlation of observation error of a Gaussian form shifts the emphasis from the mean observed value to the tendency information in the time series. The variational quality control is modified to check all data from the time series for each station jointly, so that it rejects or accepts whole time series. These two developments successfully solved the initial difficulties with consistent biases between the first guess and observations, and also act as a conservative targeted bias correction scheme removing most of the detrimental effect of the biased time series. Experimentation has shown that the accuracy of the background 1000 hPa height fields has improved. A relevant measure for the accuracy of the background field is the rms of the analysis increments. The more consistent the background is with the observations, the smaller is the rms of analysis increments (Figure 2). The improvement is largest over mid-latitudes and polar areas. The forecast impact is small, yet significantly positive at 1000hPa for most areas up to forecast ranges of 4 days.

Revision to use of radiosonde and pilot data

To date only a limited amount of the available data from sondes (TEMPs and PILOTs) have been used in the ECMWF data assimilation schemes. A maximum of 15 levels has been extracted from the sondes, choosing the data at, or nearest to, the standard pressure levels. This practice has remained unchanged during the transitions from OI to 3D-Var to 4D-Var.

Radiosonde reports contain both temperatures and geopotential heights. In 3D/4D-Var there is no preference for using heights rather than temperatures, and the extra cost of additional levels is negligible. Reported standard level geopotential heights are considered more representative than temperatures, because the standard level heights are obtained by integration of the complete measured profile of temperature (which is only available to the observer). The reported 'significant level data' of temperature and wind are those data that depart significantly from the surrounding parts of the profile, i.e. the local extreme points. The significant level data are thus less representative than the standard level data are. On the other hand, they provide potentially important information on boundary layer structure, fronts, jet maxima and the tropopause, which will become more valuable as model resolution increases.

Data assimilation experiments with 4D-Var indicated a distinct benefit from using all reported wind, temperature and humidity data from TEMPs and PILOTs. The mean rms forecast scores for 500 hPa averaged over 35 days in two separate periods showed a clear improvement in the Northern Hemisphere. There are several advantages of using temperatures instead of heights:

- the observation operator for temperature is very simple, whereas height depends on surface pressure, and the vertical integration of temperature and specific humidity.
- Bias corrections may be simpler
- Every datum in a profile of height data depends on the surface pressure observation
- Variational quality control is cheaper for assumed uncorrelated data, such as temperature.
- It is much easier to detect (and reject) gross errors in the surface pressure and the temperature data separately, than in the heights, which is a combination of the two.

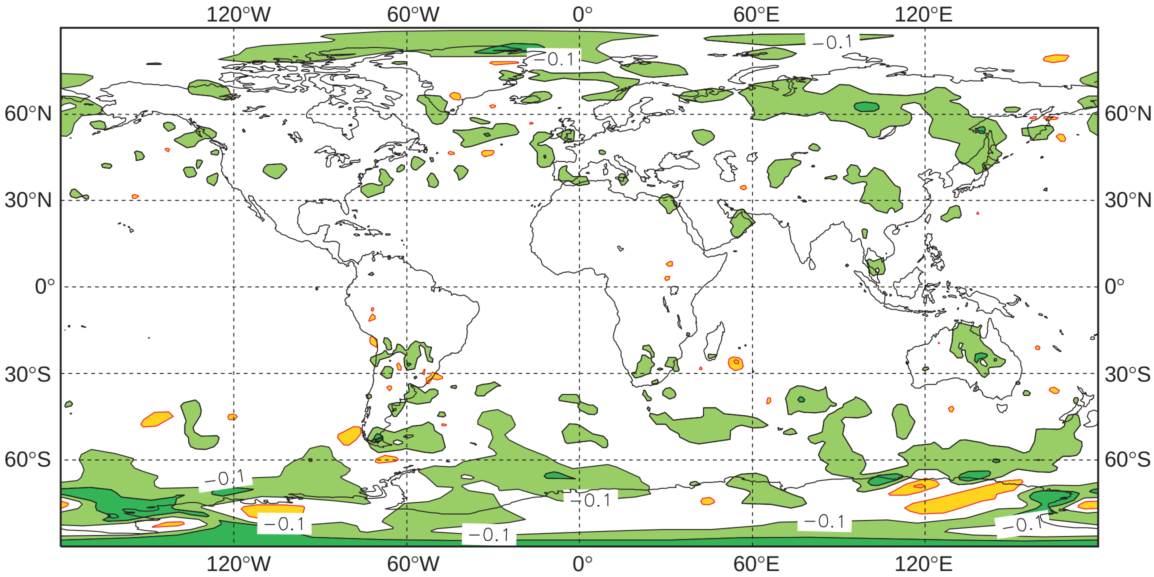

Fig. 2 1000 hPa geopotential height rms of difference in analysis increments between a 4D-Var experiment and its control. The experiment uses 4D-screening, serial correlation of observation error and joint variational quality control of time series of surface pressure and height data, and the control uses 3D-screening. The period is 11-24 Nov 1997. The contours are 0.35, 0.5, 0.75, 1.0, 1.5, 2.0 and 3.0 decametres. Green colour indicates areas where the rms of analysis increments is smaller in the experiment.

Use of more TOVS radiances over land

Traditionally only those TOVS channels, which are not sensitive to the surface, have been assimilated over land due to the difficulties of characterising the land surface emissivity at infrared and microwave frequencies. Over the ocean and sea-ice a much wider range of channels has been used as the sea and ice surface radiative temperatures are well estimated from the model fields. However the continued reductions in the radiosonde network prompted investigations into using more of the TOVS radiances over land to counteract the loss of upper air balloon data.

The use of TOVS radiances in 3/4D-Var since the introduction of Cycle 16R4 in August 1997 is given in Table 1 which lists in plain font the observation errors and surface types over which the TOVS radiances are used. The channel designations are listed for each of the TOVS instruments (HIRS, MSU and SSU). In summary the stratospheric sounding channels (HIRS 1-3, MSU-4 and SSU 1-3) were assimilated everywhere for clear and cloudy radiances and HIRS-12 is assimilated everywhere for clear radiances. These channels do not “see” the surface (with the possible exception of high mountains).

However there are further sub-sets of channels, which primarily sense the upper troposphere and are only slightly influenced by the surface. After some preliminary experimentation HIRS channels 4-6, 11, and MSU-3 were added to the list of channels assimilated over land and sea/ice. The observation errors assumed for the new channels used over land are given in italics in Table 1. The land/ice surface emissivity assumed in the forward model was unity for the HIRS channels and 0.9 for the microwave channels. However to ensure the surface only had a minor contribution to the radiance the data selection blacklist was modified to assimilate only HIRS channels 1-3 and SSU channels over topography higher than 1500m. In addition radiances from HIRS channel 8 (window channel) are now assimilated over sea and sea-ice in 4D-Var which only became possible with the new RTOVS radiance product available since autumn 1997. HIRS channel 8 radiances also allow a better detection of cloud contamination than was possible with HIRS channel 10, which was previously used. These changes resulted in an increase in the total number of radiances assimilated north of 20N by 10%.

Before operational implementation two 15 day assimilation experiments were run to test the revised use of TOVS radiances during May 1997 and December 1997. The impact of assimilating the TOVS radiances over land in terms of mean analysed 500hPa temperature differences (experiment-control) for the Northern Hemisphere is shown in Figure 3. The only significant differences are over Russia where the analyses with the use of TOVS data are on average 1 deg K colder in the mid-troposphere. Some differences in specific humidity were also seen with the main difference being an increase in total column water vapour in the Tropics of 2.5% for the May period but less for the December period. The forecast impacts for Europe were positive or neutral at all forecast ranges for both periods compared to operational forecasts. Over the N. Hemisphere the impacts were positive at all forecast ranges for the May period and neutral for the December period. For the S. Hemisphere and Tropics the forecast impacts were neutral.

Channel Number	4D-Var usage	Sea	Sea-ice	Land
HIRS-1	All	1.40K	1.40K	1.40K
HIRS-2	All	0.35K	0.35K	0.35K
HIRS-3	All	0.30K	0.30K	0.30K
HIRS-4	Only clear	0.20K	0.20K	0.20K
HIRS-5	Only clear	0.30K	0.30K	0.30K
HIRS-6	Only clear	0.40K	0.80K	0.80K
HIRS-7	Only clear	0.60K	1.20K	—
HIRS-8	Only clear	1.00K	2.00K	—
HIRS-10	Only clear	0.80K	1.60K	—
HIRS-11	Only clear	1.10K	1.10K	1.10K
HIRS-12	Only clear	1.50K	1.50K	1.50K
HIRS-13	Only clear	0.50K	1.00K	—
HIRS-14	Only clear	0.35K	0.70K	—
HIRS-15	Only clear	0.30K	0.60K	—
MSU-1	QC check	—	—	—
MSU-2	All*	0.30K	1.00K	—
MSU-3	All*	0.22K	0.22K	0.22K
MSU-4	All	0.25K	0.25K	0.25K
SSU-1	All	0.60K	0.60K	0.60K
SSU-2	All	1.00K	1.00K	1.00K
SSU-3	All	1.80K	1.80K	1.80K

* Only clear radiances used for latitudes < 30°

Table 1 TOVS radiance observation errors assigned in 1D-Var. The errors used in 4D-Var are currently 1.5 times these values. Figures in italics show where changes were made for CY18R6.

Assimilation of SSM/I total column water vapour

Since 24 February 1998 SSM/I (Special Sensor Microwave/Imager) 1D-Var (one dimensional variational method, Phalippou, 1996) retrievals have been produced as part of the operational suite, namely total column water vapour (TCWV), cloud liquid water path and surface wind speed. Comparisons of the first guess and SSM/I 1D-Var derived TCWV over the oceans has shown that the model tends to be too dry in the Tropics and Southern Hemisphere as previously shown by (Vesperini, 1998). The assimilation of TCWV derived from the SSM/I 1D-Var was tested within both the 3D-Var and 4D-Var assimilation systems and was included as part of the CY18R6 changes. Initially only radiance data from the DMSP-F13 satellite are being used. About 1300 to 1500 retrievals are assimilated at each cycle with a spatial sampling of about 250 km in latitude and longitude. It is intended to increase the coverage by also using radiances from the F-14 satellite in the near future.

In order to assess the impact of assimilating the SSM/I TCWV on the analyses and forecasts, two 15-day experiments were run in the 4D-Var assimilation system, as for the use of TOVS over land. The primary impact of assimilating SSM/I TCWV in 4D-Var is on the model humidity fields where more water vapour is included in the lower tropospheric humidity, particularly in the areas which are too dry, i.e. in the Tropics (2 to 3% increase) and Southern Hemisphere (5 to 6% increase) as shown in Figure 4.

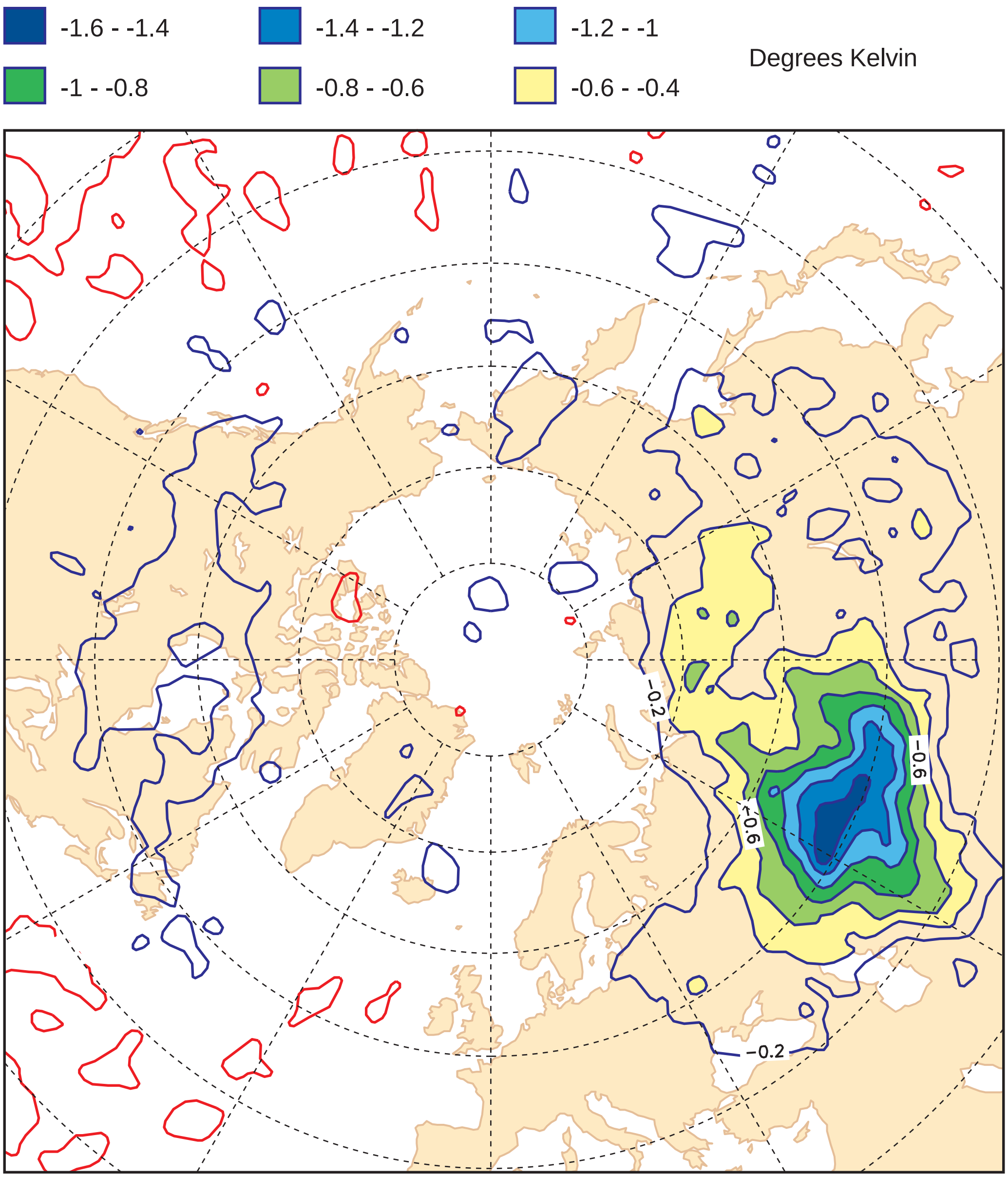

Fig. 3 The difference in temperature at 500hPa between 15 analyses in May 1997 with and without the extra TOVS channels used over land. The contours are in intervals of 0.2 K and the colours start at differences greater than 0.4 K according to the legend. Red contours are positive differences and blue negative differences.

The experiments with SSM/I TCWV gave neutral to positive impacts on the Northern Hemisphere geopotential forecasts during the two periods; the impact in the Southern Hemisphere seems to be dependent on the season, slightly positive in winter, neutral in summer. The forecasts of geopotential height are generally improved in the lower troposphere. In the Tropics the vector wind scores are neutral to slightly positive when compared with observations.

Changes to use of cloud motion winds

Meteorological observations from geostationary satellites are primarily wind observations provided by the tracking of clouds or water vapour features in subsequent satellite images. One recent change on 3 March 1998 was a revised method by NESDIS to derive GOES satellite winds (Velden et al., 1998). The changes included a smaller processing segment used for the selection of suitable tracers, their assignment to a vertical level, and the actual tracking scheme. This has led to a marked increase of the observation density by a factor three to ten depending on the tropospheric level. The smaller processing segment not only enhances information in areas with previously good coverage but also enables the extraction of winds which were not accessible at the coarser resolution providing entirely new information (e.g. cyclonic systems in the South Pacific). These wind observations can have a direct impact on the geopotential analysis at 300 hPa as demonstrated in Figure 5.

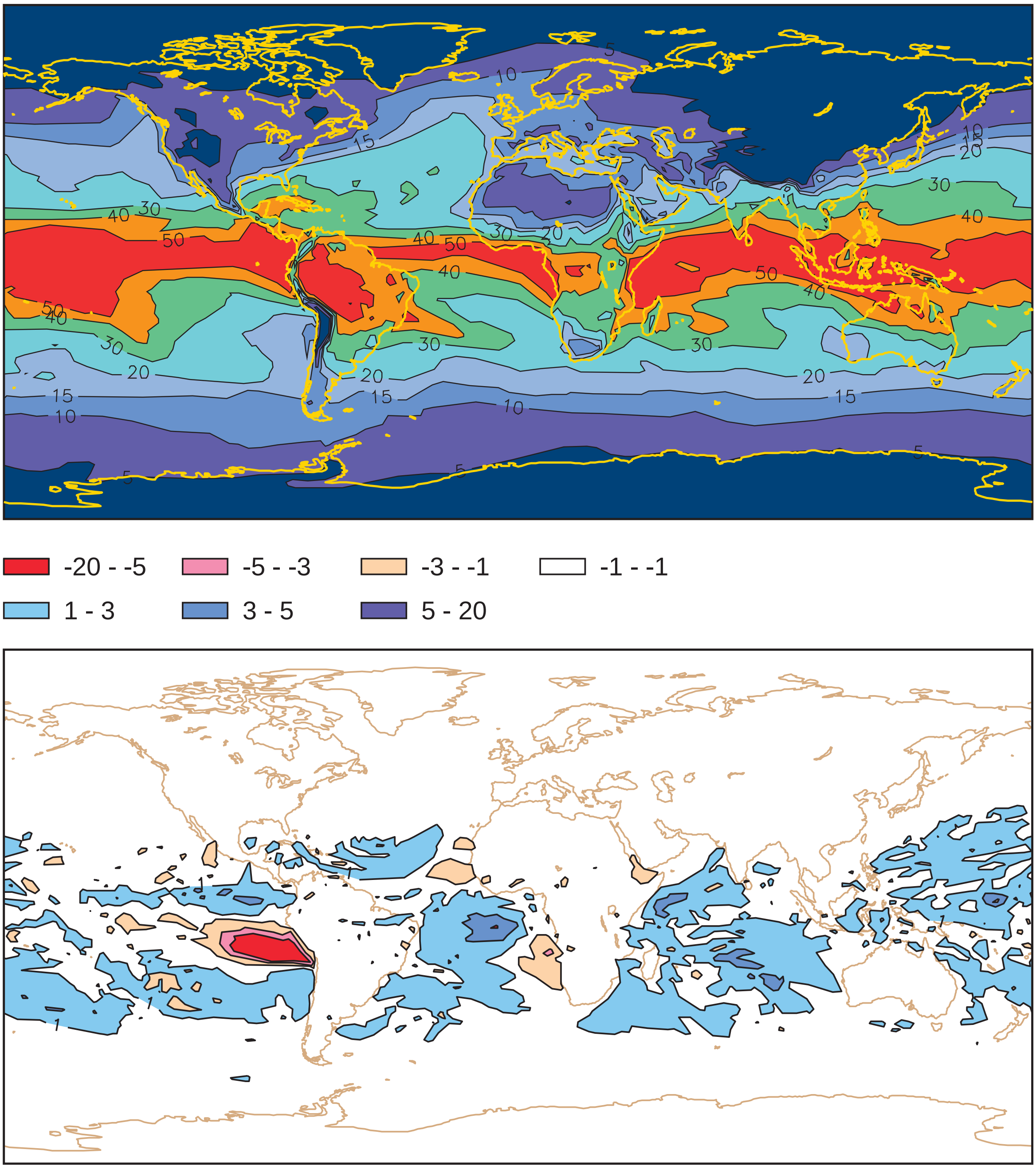

Fig.4 Mean fields computed over the 28 November - 15 December 1997 period: (top) TCWV from the experiment analysis (24.70 kg.m-2 global mean). (bottom) TCWV difference between the experiment and the control analyses (0.42 kg.m-2 global mean). Units are in kg.m-2.

However, the assimilation of the high-density GOES winds initially led to a slight degradation of the forecast in the Northern Hemisphere. One possible explanation for this is the mismatch between the resolution of the observations and the current T63 spectral truncation of the variational analyses. Therefore an additional thinning step was implemented which ensures a minimum horizontal and vertical distance between satellite winds. The impact of both the full data set and the thinned dataset during the observation screening was assessed by assimilation and forecast experiments. They showed the thinning eliminated any negative forecast impact in the N. Hemisphere and therefore the enhanced data set is used everywhere.

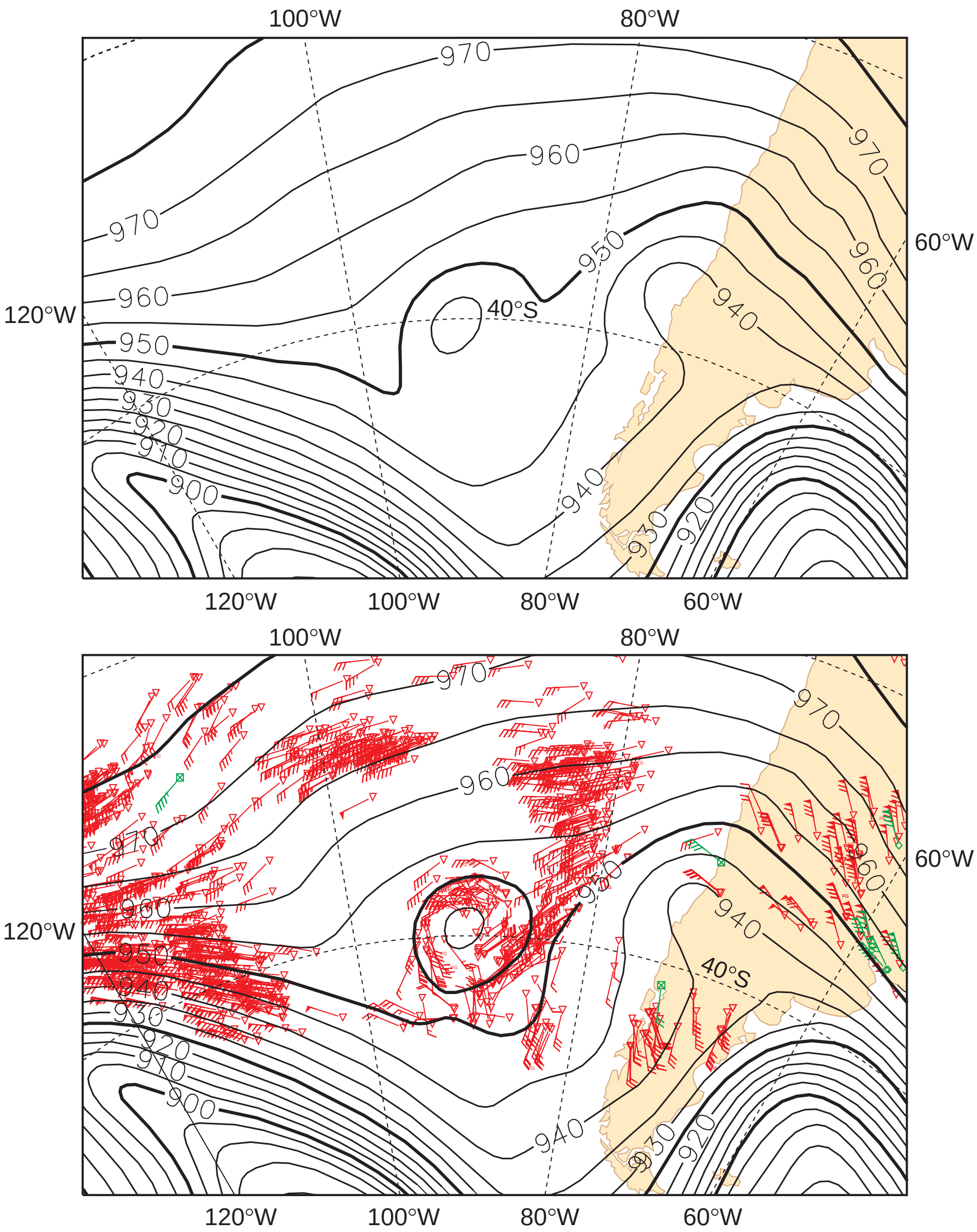

Fig. 5 Impact of high-density GOES wind observations on the 300 hPa geopotential analysis for 00UTC on 10 February 1998. The top panel was the operational analysis and the lower panel the revised analysis using the high density GOES winds. Also plotted are the GOES observations between 250 and 350 hPa and a few sparse TEMP, SHIP, PILOT observations, plotted in green.

Use of Scatterometer Winds

An inconsistency in the IFS has been that the assimilation of low level winds from different observation types (e.g. buoys, SYNOPs and scatterometer) has employed different observation operators. For the scatterometer winds, instead of using a simple logarithmic wind-law (neutral stratification), we now use an operator based on the Monin-Obukhov surface layer which takes stratification into account (Cardinali et al.; 1994). There is then just one single observation operator for 10m winds for all observation types. The scatterometer bias corrections have been recalculated to reflect this.

Another change is that the roughness length, z0, used for the calculation of 10m scatterometer winds now uses the model z0 rather than the observed winds used previously. This change together with the coupling of the wave model to the atmospheric model has resulted in a reduction in the standard deviation between scatterometer and model first-guess wind speed from 1.80 ms-1 to 1.67 ms-1.

References

Cardinali, C., E. Andersson, P. Viterbo, J-N. Thépaut and D. Vasiljevic: 1994 Use of conventional surface observations in three-dimensional variational assimilation. ECMWF Technical Memorandum 207

Phalippou, L., 1996: Variational retrieval of humidity profile, wind speed and cloud liquid water path with the SSM/I: Potential for numerical weather prediction. *QJ Royal Met. Soc.,* **122,** 327-355.

Rabier, F., J-F. Mahfouf, M. Fisher, H. Järvinen, A. Simmons, E. Andersson, F. Bouttier, P. Courtier, M. Hamrud, J. Haseler, A. Hollingsworth, L. Isaksen, E. Klinker, S. Saarinen, C. Temperton, J-N. Thépaut, P. Undén and D. Vasiljevic. 1998 The ECMWF operational implementation of four-dimensional variational assimilation. *QJ Royal Met. Soc.*

Velden, C. S., T. L. Olander and S. Wanzong, 1998: The Impact of Multi-spectral GOES-8 Wind Information on Atlantic Tropical Cyclone Track Forecasts in 1995. Part I: Dataset Methodology, Description, and Case Analysis. *Monthly Weather Review,* **126,** 1202-1217.

Vesperini, M. 1998 Humidity in the ECMWF model: Monitoring of operational analyses and forecasts using SSM/I observations. *QJ Royal Met. Soc.,* **124,** 1313-1328.

R. Saunders, E. Andersson, H. Järvinen, É. Gérard, M. Rohn, L. Isaksen.