USE OF SATELLITE DATA
IN THE ECMWF ANALYSIS SYSTEM

Jean Pailleux
European Centre for Medium Range Weather Forecasts
Reading, U.K.

ABSTRACT

The first part of this paper describes the technique used to enter the satellite soundings in the ECMWF optimum interpolation (OI) analyis scheme.

Some experiments have been carried out on the horizontal and vertical resolutions of satellite data, and some results are presented in Section 2. Finally, in Section 3, alternative techniques for using the satellite data are discussed.

1. THE ANALYSIS SCHEME AND THE SATELLITE DATA

The ECMWF operational analysis scheme has three components: a mass/wind analysis, a humidity analysis and a surface analysis. In this paper we study the use of satellite data in the mass/wind analysis. The mass/wind analysis system has the following characteristics:

* It is a 3-dimensional OI scheme for wind and geopotential height increments (U, V and Z).
* The data used to evaluate the increments are wind, geopotential and thickness data on the 15 standard levels from 1000 to 10 hPa. The orbiting satellite soundings (SATEMS) are the only observations providing the analysis system with thickness data DZ.
* The increments are evaluated directly on the model levels.
* A "box technique" is used to set up a large correlation matrix and solve the OI equations: up to 255 pieces of data are used to analyse one latitude/longitude box.

More details about the analysis system are available in Lorenc (1981). Here we are interested in the way the satellite information is introduced in the right-hand part of the OI equations, which is:

$$\Sigma W_i \varepsilon U_i + \Sigma W_j \varepsilon U_j + \Sigma W_k \varepsilon Z_k + \Sigma W_l \varepsilon (DZ)_l .$$

$\varepsilon(DZ)_l$ are the thickness increments coming from the satellite soundings and W_l are the weights given to the different pieces of data. εU_i, εU_j and εZ_k are the wind and height increments coming from the other observations. For the current SATEMS available on the GTS (horizontal resolution: 500 km), the information content consists generally of 14 thicknesses between 1000 hPa and all the other standard levels.

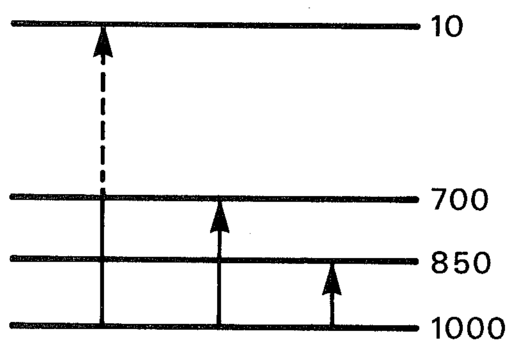


Fig. 1: Thicknesses available in the 500 km SATEMS.

A more comprehensive data set is also produced in Washington and available at ECMWF. The basic information is the same, but the horizontal resolution is 250 km instead of 500 km, and the presentation of the information is also different: one profile consists generally of 11 virtual temperatures for the following layers: 1000/850, 850/700, 700/500, 500/400, 400/300, 300/200, 200/100, 100/70, 70/50, 50/30 and 30/10. The ECMWF analysis generally uses the 250 km SATEMS, but because of telecommunication problems, it may happen that they are missing in some areas while the 500 km SATEMS are available. So both data sets are actually used.

Before being introduced in the OI equations, the satellite data are first converted to 14 adjacent thicknesses between the standard levels: 1000/850, 850/700, ... 20/10. Therefore the large thicknesses of the 500 km SATEMS are split into thin layers, while the virtual temperature data of the 250 km SATEMS are converted to the 14 thicknesses through the hydrostatic equation.

The following points must be noted:

* In the SATEM code the thicknesses are rounded to the nearest decametre, which gives a poor relative accuracy on thin layers such as 1000/850 and 850/700, in the 500 km SATEMS.

* Off-time data are used as if they were produced at the synoptic time, but the observation error standard deviation is slightly increased for off-time data.

* The observation time is rounded to the nearest hour in the 500 km SATEMS, which means that an observation made at 03.29Z is rounded to 03.00Z and goes into the 00Z analysis.

* The satellite soundings are not used over land below 100 mb and they are not used at all between 60°S and 90°S (decision based on past operational experience).

* A vertical and horizontal correlation is assumed in the OI system for the satellite observation error. For example, the correlation assumed between two soundings separated by distance r is $0.8 \exp(-r^2/2a^2)$ with a = 350 km.

This correlation is applied to each pair of soundings, even if the soundings do not belong to the same satellite. This last assumption is a weakness when, for example, there is an overlap between a NOAA6 orbit and a NOAA9 orbit producing inconsistent data.

2. EXPERIMENTS ON HORIZONTAL AND VERTICAL RESOLUTION FOR SATELLITE DATA

2.1 Horizontal resolution : 500 km SATEMS / 250 km SATEMS

A data assimilation and forecast experiment (Experiment 1) has been carried out for the period (12Z, 5 February - 12Z, 7 February 1985).

* At that time the operational analysis scheme was using the 500 km SATEMS only.

* The purpose of the experiment was to compare the operational system with an experimental run using the 250 km SATEMS. The analyses of the parallel runs have been compared for the 2-day period and a 6-day forecast has been run from 12Z, 7 February 1985 to be compared with the operational forecast.

* At that time the operational orbiting satellites were NOAA6 and NOAA7. An instrument problem occurred on 5 February in the evening, producing several gaps in the data coverage.

* To use the 250 km SATEMS, an additional modification has been made in the data selection algorithm: the maximum matrix size used to analyse one box has been increased from 191 to 255. This means that up to 255 pieces of data can be used to analyse one given box and one given slab (in data-dense areas each latitude/longitude box is divided into 3 slabs in the vertical to perform the analysis). As an example Fig. 2 shows the observations which are used to analyse the middle slab of a box situated over Japan.

We can see that the data coverage is improved by the use of 250 km SATEMS, especially in the southern part of the anaiysis box.

When examining the analysis from the parallel run, no important differences are found. Another way to compare the two assimilations consists of looking at the fit of the observations to the 6 hour forecast (first guess). In the northern hemisphere the fit of the radiosondes to the 6 hour forecast is similar in the two parallel runs; in the southern hemisphere it seems to be slightly better when we use the 250 km SATEMS, as shown in Fig. 3. But, due to the small sample of radiosondes (about 35), it is difficult to be sure to what extent this difference is significant.

The fit of the SATEM observations to the first guess is improved when we use the 250 km SATEMS, especially in the southern hemisphere as shown in Fig. 4. But it is not a clear signal showing an improvement of the assimilation since it is expected that an increase of the number of satellite soundings gets the first guess closer to the satellite data.

When we compare the forecasts made at 12Z, 7 February 1985, no significant difference can be detected. The order of magnitude of the differences is illustrated by Fig. 5 which shows the differences between the 3-day forecasts (operations - 250 km SATEMS).

2.2 Vertical resolution: 14/11 thicknesses

Although we cannot conclude too much from the single forecast described in Section 2.1, at least the experiment does show that the impact of a drastic change in the satellite horizontal resolution is surprisingly small. It may mean that the redundancy of satellite data is very large in

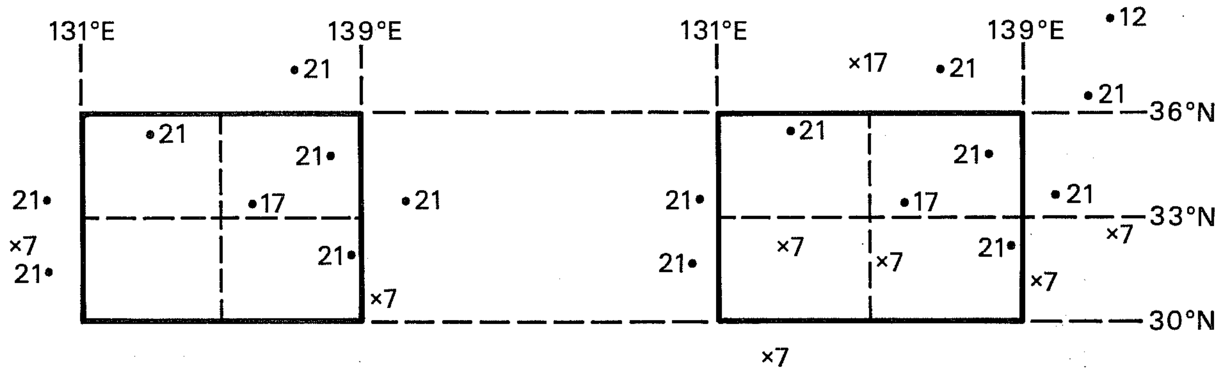


Fig. 2: Observations used to analyse the middle slab of a box over Japan.
. Radiosonde observations
x Satellite observations
The other observation types are not shown and the numbers indicate how many pieces of data have been used from each observation point.
Left: 500 km SATEMS with maximum matrix size = 191
Right: 250 km SATEMS with maximum matrix size = 255.

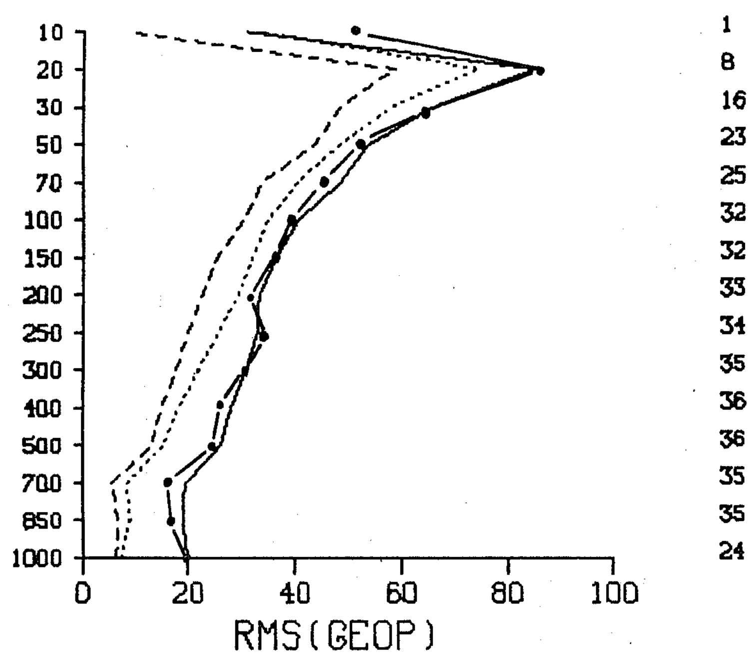
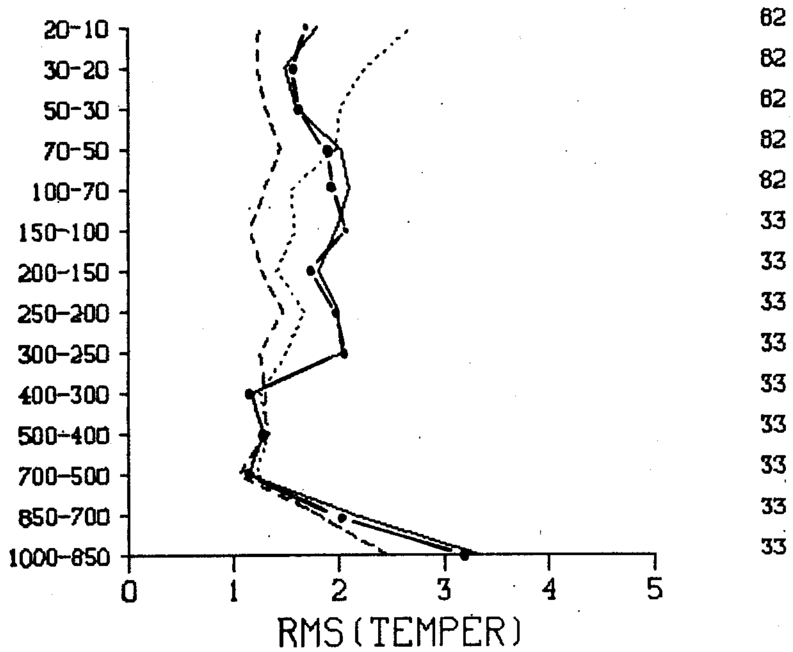
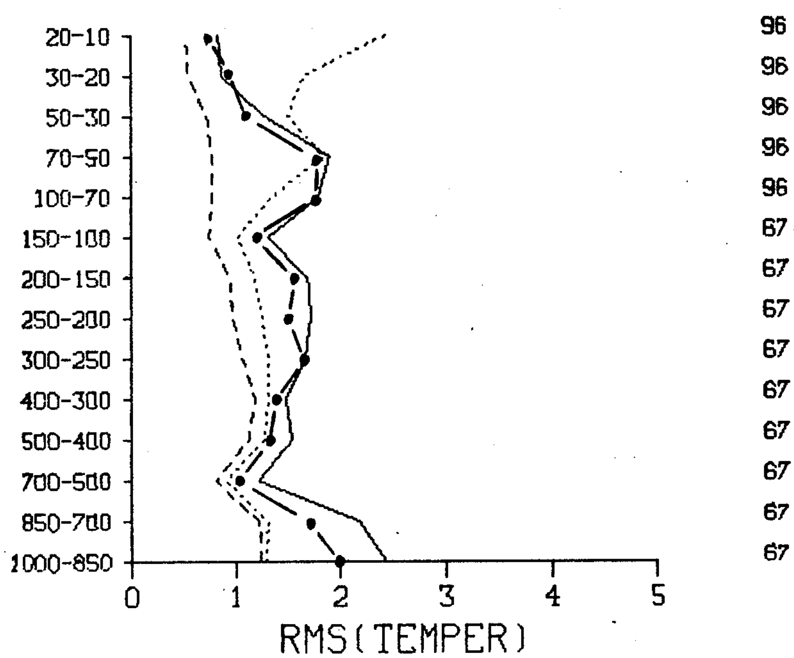


Fig. 3: RMS fit of radiosonde observations to the operational first guess (———), the operational analysis (------) and the operational initialisation (.....). The first curve is to be compared with the fit to the first-guess using the 250 km SATEMS (.-.-.-). 12Z, 7 February 1985.



Northern
Hemisphere



Southern
Hemisphere

Fig. 4: RMS fit of SATEM observations to the first guess.
——: 500 km SATEMS (operations). .-.-.-: 250 km SATEMS.
The two other curves are the fit to the analysis and initialisation. 12Z, 6 February 1985.
Top: northern hemisphere
Bottom: southern hemisphere.

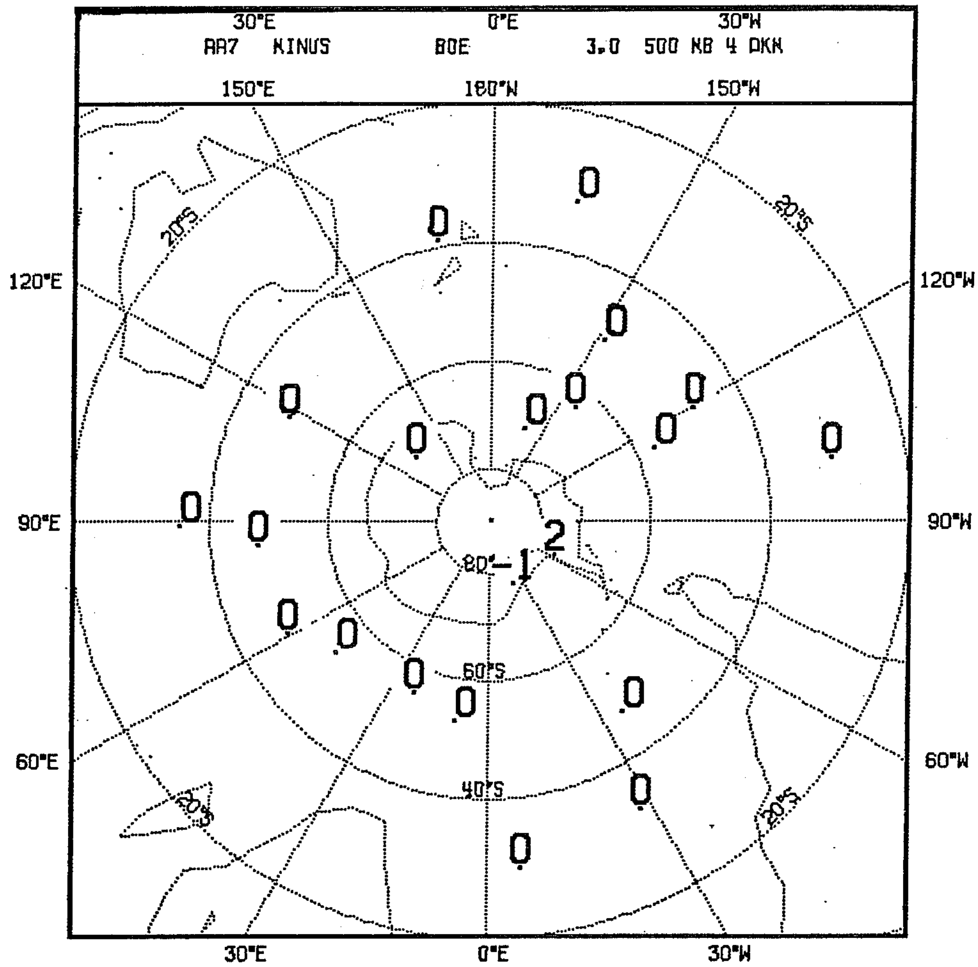


Fig. 5: Difference map between the operational forecast using the 500 km and the experiment using the 250 km SATEMS. 3-day forecast from 12Z, 7 February 1985. 500 hPa geopotential height in dam.

the horizontal, but the horizontal correlation of observation errors, $0.8 \exp(-r^2/2a^2)$ (a = 350 km), indicates that there should be additional information in the 250 km SATEMS, and the accuracy of this horizontal correlation is confirmed by Schlatter (1981) and recent studies carried out by Kelly at ECMWF. Another explanation is that the present ECMWF analysis system is deficient in extracting the information from high resolution satellite data.

One known deficiency is related to the vertical resolution of satellite data which is used in the analysis. Kelly has shown that the 14 thicknesses introduced in the OI analysis are highly redundant. It is not realistic to use more layers than the 11 contained in the 250 km data set, and also the analysis should not fit small thicknesses which are given in the 500 km SATEMS with a very low relative accuracy (see Section 1).

In a recent experiment carried out by Andersen at ECMWF (Experiment 2), 11 layers instead of 14 were used in the analysis:

* 300/200 instead of 300/250 + 250/200
* 200/100 instead of 200/150 + 150/100
* 30/10 instead of 30/20 + 20/10

To do this, the observation error statistics were re-derived, as we do not use the same observed quantities. At the same time the vertical correlation matrix of observation errors was tuned in order to let the analysis draw to large thicknesses rather than to small thicknesses (e.g. 700/500). The other characteristics of the experiment are as follows:

* Period: 12Z, 15 March 1985 - 12Z, 19 March 1985. At that time the operational analysis was run with 14 layers in the vertical and 250 km in the horizontal (the 250 km SATEMS were introduced operationally at the end of February 1985).
* The operational satellites were NOAA6 and NOAA9 (NOAA9 replaced NOAA7 at the end of February 1985).
* Two 10-day forecasts were run (12Z, 17 March and 12Z, 19 March) to assess the impact of these modifications in the vertical.

Although the differences between the two analysis sets are generally small, they are larger than those described in Section 2.1; locally they reach 4° for the 700/500 layer. The same conclusion can be drawn from the

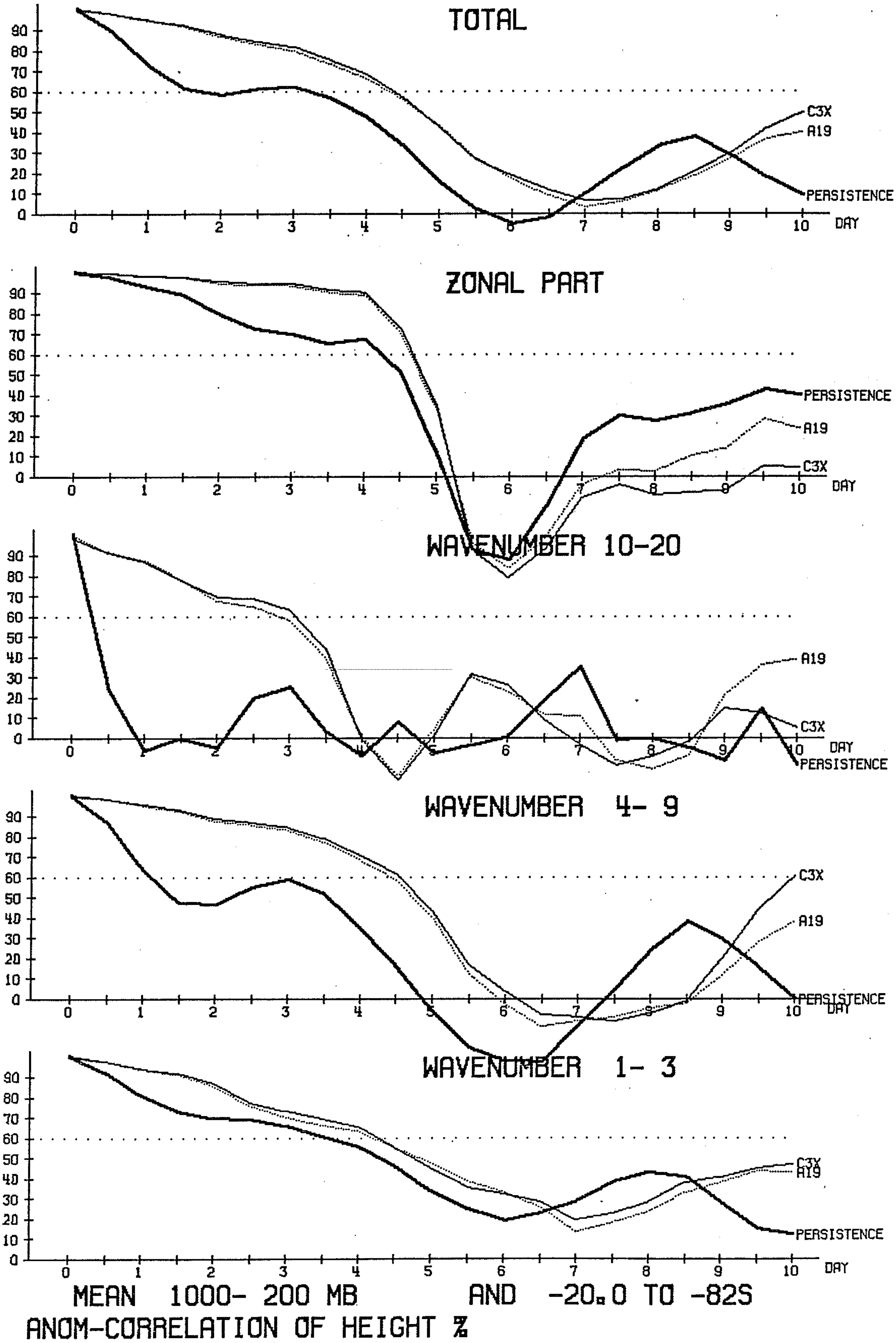


Fig. 6: Anomaly correlation of height as a function of forecast time for the extratropical southern hemisphere for forecasts starting on 12Z, 19 March 1985.
A19: 14-layer run (operations)
C3X: 11-layer experiment.

comparison of the forecasts. In one case (12Z, 17 March) the impact is negligible, but in the second (12Z, 19 March) the forecast is somewhat better in the 11-layer assimilation, especially in the southern hemisphere (see the anomaly correlation curves given in Fig. 6). From these results it seems we can deduce two conclusions:

* A confirmation that 14 layers in the vertical is too much, as a reduction to 11 layers has no detrimental effect.
* We should rely more on the observed values of thick layers. In this matter the modification made on the vertical correlation function may be more important than the vertical resolution itself.

2.3 Horizontal resolution

The experiment comparing the 500 km SATEMS with the 250 km SATEMS (Experiment 1) was repeated using the 11-layer system with the modified vertical structure function used for Experiment 2. The period was the same as the one used in Experiment 2, and an assimilation experiment was carried out on the period (12Z, 17 March - 12Z, 19 March 1985). Then a forecast was run on 12Z, 19 March 1985 to assess the impact of horizontal resolution in SATEMS.

The results are similar to those of Experiment 1. No improvement can be seen in the first guess and the medium-range forecast; the differences are very small, sometimes in favour of the 250 km SATEMS, sometimes in favour of the 500 km SATEMS!

3. CONCLUSIONS

* When we double the horizontal resolution of satellite soundings (250 km instead of 500 km) in the present ECMWF assimilation system, the impact on the forecast is small and cannot be properly assessed by the two forecast experiments which have been done. A much larger sample of experiments would probably be necessary to see the improvement.
* A modification in the vertical resolution (11 layers instead of 14) combined with a retuning of the vertical correlation function seems to have a somewhat larger beneficial effect.
* Instead of using 14 thin layers, the analysis scheme should use a reduced number of thicker layers.

* More generally, the exact information content of satellite data needs more investigation in order to use the appropriate resolutions in the horizontal and the vertical.

* In order to use the correct information which is contained in the satellite soundings, it is not enough to know the appropriate resolutions: correct tuning of the observation error statistics is necessary and this is especially true for off-time data which are used very crudely at the moment.

* Finally it must be stressed that the retrieval technique itself has not been studied in these experiments. All the SATEMS used come from a statistical retrieval made in Washington. One basic problem is to determine the best combination of a retrieval system and a data assimilation system. Several alternative methods are possible, such as the direct use of radiances in the 3-dimensional OI system, or a physical retrieval method using the first guess of the data assimilation system.

REFERENCES

Lorenc, A.C., (1981): A global three-dimensional multivariate statistical interpolation scheme. Monthly Weather Review, 109, 701-721.

Schlatter, T.W., (1981): An assessment of operational TIROS-N temperature retrievals over the United States. Monthly Weather Review, 108, 110-119.