TECHNICAL MEMORANDUM

143

Use of satellite vertical sounder data in the ECMWF analysis system

G. Kelly and J. Pailleux

Research Department

March 1988

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

1. INTRODUCTION

The satellite soundings obtained from the TIROS Operational Vertical Sounding (TOVS) instruments (called SATEMs) are the only observing system providing global 3-dimensional observations of the mass field in the atmosphere. The use of this satellite data in Numerical Weather Prediction (NWP) is a key research area for ECMWF because there is evidence that the present use of SATEMs is far from optimal: Information from the satellite soundings is used on vertical and horizontal scales which are chosen somewhat arbitrarily; the choice is not based on the information content of the radiance measurements. There are many problems in the retrieval algorithms which are not considered in the analysis. The quality control in the analysis does not cope well with erroneous satellite data.

In the recent past, several changes have been made in the operational analysis at ECMWF, to improve the use of satellite data:

- The average horizontal density of SATEMs was increased from approximately 500 km to 250 km (February, 1985).
- The first guess check was tightened by a factor of two for SATEMs (May 1985).
- 11 vertical layers were used instead of 14 in the SATEMs (October 1985).
- The operational SATEM blacklist (geographical regions where data is rejected) has been changed several times to cope with regional satellite data problems, especially in the polar areas. In the ECMWF operations at the beginning of 1987, the SATEMs are used only above 400 hPa near the poles (South of 60°S, North of 70°N), and above 100 hPa over land. Finally, the new analysis code, implemented in September 1986, allows a more extensive data selection, increasing the number of satellite soundings used for each analysis box (see Lönnberg et al., 1987).

In an earlier study (Pailleux et al., 1986) we carried out an interactive satellite retrieval experiment in which SATEMs were produced by a physical inversion method (Wisconsin University - International TOVS processing package), using the 6 hour forecast as first guess. The main purpose of this experiment was to investigate whether the quality of the satellite soundings (and of the subsequent analyses and forecasts) was improved by using the information from the ECMWF 6 hour forecast in the retrieval algorithm. The results of the physical retrieval experiment were then compared with the corresponding operational assimilation using the SATEMs produced at NESDIS (Washington) through a statistical retrieval technique. This experiment was run during a 6-day period in November 1979, using the FGGE-IIB data as described in Pailleux et al. (1986). The main conclusion drawn from this experiment was that it is dangerous to use a 6h forecast as a first guess to retrieve the SATEMs, if it is then assumed in the analysis that these SATEMs are not correlated to the first guess ("incest" problem in the analysis). We also found in this retrieval experiment that it was difficult to maintain a realistic stratosphere in the assimilations, suggesting that the forecast model's vertical resolution was too low in the stratosphere, with only 2 model levels above 100 hPa. The SATEM's vertical resolution, as used by the analysis, had to be chosen carefully. The model's description of the stratosphere has since been improved by implementing operationally a 19 level forecast model with 5 model levels above 100 hPa instead of 2 (May 1986).

The first part of this study explores the use of a reduced vertical resolution for the SATEMs in the analysis. The original choices of 14 or 11 layers were based on pure coding considerations (such as the necessity of describing all the WMO standard levels). We chose a vertical resolution (6 layers) based on physical considerations using the weighting functions of the TOVS radiance channels. Such a resolution change requires the derivation of new statistics for the SATEM observation errors. The second part of this study reassesses, with an up-to-date assimilation system, the overall impact of the satellite soundings on the analysis and the forecast.

Section 2 describes an experiment varying SATEM vertical resolution, run on the same FGGE period as the above mentioned physical retrieval experiment. A 6-layer assimilation is compared with the operational assimilation system as it was in October 1986 (using 11 SATEM layers; new analysis code; 19 model levels) except that the horizontal resolution in both assimilations was T63 instead of T106. These two experiments are also compared to a "NO-SATEM" run thus reproducing, with a current assimilation system, one of the FGGE Observing System Experiments (OSEs) carried out by Uppala et al. (1984). Our experiments highlighted some SATEM data quality problems which are discussed in Section 3. Section 4 explains the decision to add an extra SATEM layer in the stratosphere (7 layers instead of 6), and also describes the structure function computations and other adjustments in order to improve the response of the analysis to SATEM data. The full list of analysis modifications described in Section 5 has been used for an extensive benchmark run on a recent period (30 January 1987, 00Z, to 14 February 1987, 12Z) and which is described in Section 6. This modification set was implemented operationally at ECMWF in July 1987.

2. 6 - LAYER EXPERIMENT

2.1 Choice of SATEM layers

Currently U.S. vertical temperature soundings derived from passive satellite radiance measurements are sent to ECMWF as either 14 mean layer temperatures (GTS SATEMs) or 11 mean layer temperatures (250 km SATEMs). These mean layer temperatures were used in the analysis with specified horizontal and vertical error statistics. (See Table 1). Fig. 1 shows the vertical resolution of the HIRS instrument onboard the NOAA spacecraft. The diagram has been calculated using the HIRS weighting functions as input to a minimum information solution of the radiative transfer equation (Kelly, 1984). It is clear that the satellite radiance measurements only account for about six independent pieces of information in the vertical. Any extra structure in the satellite derived vertical profile is therefore introduced via the retrieval method, which currently is statistical regression. It was felt highly desirable, to remove extra statistical information from the satellite measurements by using thicker layers. Initially the six layers 1000/700, 700/500, 500/300, 300/100, 100/50, 50/10 hPa were chosen and are formed by addition of the SATEM layers.

The satellite vertical error statistics for these 6 layers were derived from the current 11-layer statistics and are shown in table 1(a), (b). The horizontal SATEM error structure function was not changed from the present analysis, i.e. an exponential with a length scale of 350 km (Lorenc, 1981).

A data assimilation was run with these modifications for six days during a FGGE period which was previously found sensitive to satellite data. After this experiment a small error was discovered in the calculation of the vertical correlation of satellite error and the recalculations are shown in Table 1(c). In general there is overall agreement with Tables 1(b) and (c) and it did not seem necessary to repeat the experiment.

2.2 Data Assimilation Experiments

Assimilation experiments were run during the period 00Z 14 November 1979 to 20 November 1979 using the latest analysis system (Lönnberg et al., 1986). The first experiment was an 11 layer satellite experiment (P73), the second was the 6 layer experiment (P71) and the third was a NO-SATEM experiment (P72). All experiments started with the same first guess six hour forecast.

Table 1(a)
6-layer standard error for SATEMs (meters)

Layer (hPa)	Clear	Cloudy	Microwave
1000-700	21	23	26
700-500	15	16	18
500-300	23	25	30
300-100	27	30	33
100-50	28	31	35
50-10	72	80	88

Table 1(b)
SATEM 6-layer Vertical Correlations
(used in experiment P71)

Layer (hPa)	1000-700	700-500	500-300	300-100	100-50	50-10
1000-700	1.0					
700-500	.27	1.0				
500-300	-.45	-.11	1.0			
300-100	-.01	-.01	-.38	1.0		
100-50	.0	.0	-.02	-.40	1.0	
50-10	.0	.0	.0	-.11	-.40	1.0

Table 1(c)
SATEM 6-layer Vertical Correlations
(recalculated)

Layer (hPa)	1000-700	700-500	500-300	300-100	100-50	50-10
1000-700	1.0					
700-500	.01	1.0				
500-300	-.46	.28	1.0			
300-100	-.12	-.43	-.29	1.0		
100-50	.0	.0	-.01	-.31	1.0	
50-10	.0	.0	.0	-.15	-.43	1.0

Temperature difference maps between the first-guess of the 6-layer and 11-layer assimilations are shown in Fig. 2 for the 1000-850 hPa and 850-700 hPa layers after 60 hours of assimilation. As expected the boundary layer structure over the ocean regions shows large areas of departure between the 11 layer (P73) and 6 layer (P71) experiments. In general there is a compensation between the upper and lower layers and the combined 1000-700 hPa mean layer temperature is very similar in both experiments. Similar difference were observed near the tropopause in 300-200 hPa and 200-100 hPa layers. It is presumed that the vertical structure in these thin layers is mainly controlled by the regression statistics of the SATEMs in one case and by the model structure in the other case.

2.3 Assimilation Statistics

In all these assimilation experiments the mean and RMS statistics are calculated for the differences between the observations and the six hour forecast, the analysis and the initialization. The level of the RMS differences in the Northern Hemisphere and tropics is generally similar between all experiments; only the NOSATEM experiment in the Southern Hemisphere has larger 6 hour forecast errors. In the biases there are more noteworthy differences between the experiments. The geopotential temperature biases, in all experiments, reveal that the forecast remains too warm in the stratosphere. This was a problem in the first guess and most probably due to the fact that the Northern Hemisphere SATEMs are too warm compared with stratospheric radiosondes.

The first guess for the first NOSATEM analysis was based on a previous satellite experiment. It is interesting that even without satellite data, the stratospheric bias thus introduced remains in the analysis after six days.

In the tropics the 6 layer run appeared to develop a larger stratospheric bias, compared with the 11 layer run, indicating a possible mismatch of the distribution of SATEM stratospheric layers and the analysis.

Generally the clear soundings are closer to the six hour forecast for both satellites (which were TIROS-N and NOAA6 in November 1979). There are biases in the tropics in the lowest layers; in almost all cases, these biases indicate the SATEMs are too warm with respect to the six hour forecast. The fit of the microwave NOAA-6 soundings in the Northern Hemisphere is particularly poor with biases of more than one degree in the middle troposphere. Only in the Southern Hemisphere and the Tropical Stratosphere is the SATEM data retained after analysis and initialisation.

2.4 6 and 11 layer assimilations

A series of seven T63 forecasts was run from the 6 layer and 11 layer assimilations. The results are summarized in Table 2. In general there was little difference in the Northern Hemisphere and a small increase in skill for the 6 layer forecasts in the Southern Hemisphere. On 16 November 1979, 12Z, the 6-layer forecast was clearly better in the Southern Hemisphere, judging from the objective scores, although the differences are very small on the Northern Hemisphere. The 500 hPa four day forecast for 18 November 12UTC (Fig. 3) shows clear synoptic improvements in the Southern Hemisphere: the deep low in the Southern Pacific near the Antarctic, the low south of Australia, and the trough and low south of South Africa, are all forecast better in the six layer run.

2.5 SATEM and NOSATEM

It had been found in a previous study (Uppala et al., 1984) that Northern Hemispheric forecasts in this November 1979 period were sensitive to SATEM data. These earlier experiments were run with the 1982 analysis system and the grid point forecast model. Fig. 4 shows similar impact of SATEM data using the current (1987) analysis for three SATEM/NOSATEM forecasts.

Table 2

Anomaly Correlation height

Mean 1000-200 hPa Area Mean 20° to 83°

No of days in which Anomaly Correlation reaches 0.6

DATE	Northern Hemisphere		Southern Hemisphere	
	6-Layer	11-Layer	6-Layer	11-Layer
791116 00Z	6.6	6.8	5.7	5.5
791116 12Z	6.8	6.7	6.3	5.5
791117 12Z	5.8	5.8	6.8	6.1
791118 12Z	6.5	6.1	5.4	6.0
791119 12Z	6.1	6.2	5.9	5.7
791120 00Z	5.8	5.7	6.6	6.5
791120 12Z	5.6	5.7	5.9	5.8
Average	6.1	6.1	6.1	5.8

3. OTHER REMARKS ON THE USE OF SATEMs

In a systematic study of the increment maps of the FGGE experiments discussed in Section 2, it was found that the response of the analysis to SATEM data in some areas was unexpected. The temperature increments were examined in detail for the layer 1000/850 hPa, and on several occasions the maximum increments were found outside or on the edges of the satellite swath.

In order to investigate the response of the analysis, a special assimilation experiment was run for a 4-day period which is included in the FGGE period: 14 November 12Z to 18 November 12Z, in 1979. This assimilation used SATEMs only, all the other observation types being discarded. The overall quality of this assimilation was poor, as expected. However the aim was to study analysis response to SATEMs using increment maps. Fig. 5 illustrates a typical analysis problem. Fig. 5a shows the 1000/850 hPa mean temperature increments on 14 November 1979, 12Z, in the Antarctic Ocean. There is a large negative increment (-3°C) in an analysis box close to (65°S, 85°W). The analysis box encompassing this point does not contain any data, but it is just to the East of a satellite track. Fig. 5b shows the 1000/850 hPa layer mean temperature first guess, and the departures "SATEM-first guess" at the observation points. Fig. 5c shows the SATEM observations used in the data selection for this analysis box, together with selected observed values in degrees (1000/850 hPa mean temperature). Several remarks can be made from these results:

a) The departures "SATEM - first guess" are often large close to the edges of the satellite swaths, suggesting some deficiencies in the limb correction in the retrieval algorithms (see the SATEM points to the West of the analysis box with departures around -4°C).

b) There are some large horizontal inconsistencies in the SATEM data. Fig. 5c shows an example: North-West of the analysis box there is one SATEM situated just on the edge of the swath reporting -5.1°C, and another one, 200 km away from the analysis box reporting +3.9° (a 9° gradient!) The SATEM reporting +3.9°C is obviously incorrect (much warmer than all its neighbours) while the one reporting -5.1°C is probably too cold. None of them has been rejected in the quality control. Moreover, as the analysis assumes a horizontal correlation between the observation errors associated to the previous SATEMs, the result is an "extrapolation" of the incorrect horizontal gradient outside the satellite swath. This caused the strange increment pattern, with a maximum in the no data region.

Similar SATEM quality problems have been found on recent operational cases. Although the SATEM quality has improved since FGGE, current operations are still affected by bad data near the edges of the swath and by horizontal inconsistencies within the data. The inconsistencies are probably related to the cloud-clearing algorithm, as they often occur in cloudy areas. As the number of rejected SATEMs is very small in operations (and in the FGGE experiments) the present quality control performed in the analysis is inadequate for SATEMs. There is a need to tighten the quality control.

The general quality of the SATEMs in the layer 1000/850 hPa during this FGGE period is illustrated in Fig. 6 which shows histograms of the departures "Observation-First Guess" in the 11-layer experiment, grouped according to geographical area (Northern Hemisphere, Tropics, Southern Hemisphere) and retrieval type (clear, cloudy, microwave). The statistics are accumulated on the period 14 to 20 November (00Z and 12Z only) and based on a large sample. We note the following points:

a) Existence of biases which are either due to SATEMs, the forecast or both. The present analysis system cannot filter such biases, but assumes that both the observations and the first guess are unbiased.

b) SATEM quality in the lower troposphere is dependent on the retrieval type. The standard deviations is lower for clear soundings, than cloudy soundings, and microwave soundings.

c) The assumption made in the analysis that the SATEM errors are gaussian does not handle correctly the non-gaussian error structures appearing on these histograms.

4. SATELLITE ERROR STATISTICS

In an OI analysis system a key factor for the quality of the analysis is the use of reasonable error statistics. In the context of this study, the satellite error statistics have been reexamined, adapted to the changes of SATEM vertical resolution and several aspects have been retained.

4.1 Horizontal correlation

The ECMWF analysis uses an exponential function for modelling the SATEM error horizontal correlation. The precise form is taken from work of Schlatter et al. (1979) using NIMBUS-6 data. The results appear still valid for the current TOVS instruments (Kelly, 1985). The length scale is presently 350 km, and following recent computations (Andersen, pers. communication) it was decided to keep this 350 km horizontal scale unchanged. With the use of better quality satellite retrievals these calculations will need to be recomputed. Until 1987 the horizontal error correlation of the SATEM error has been applied to all SATEM data, irrespective of satellite or instrument. Some developments have been made by Unden (1987): only SATEM data coming from the same satellite are assumed to be horizontally correlated; and the correlation is reduced according to the time difference between the two observations. These modifications have been used in the rest of this study. The correlation between SATEM errors should also be dependent on the retrieval type (in principle). This has not been considered so far.

4.2 Vertical correlation

In 1984 the vertical correlation structure SATEM errors were compared to previous error calculations in 1979 by Kelly (1984, 1985). Following this study some changes were made to the operational system in 1985 when going from 14 to 11 SATEM layers. The major aim of reducing the number of SATEM layers is to reduce the correlated error between layers and hence make those satellite measurements more independent.

Given the results on stratospheric bias in the FGGE 6-layer 19-level assimilation, it was decided, for operational implementation, to add an extra layer in the stratosphere and hence increase the number of SATEM layers from 6 to 7. This was considered necessary to control the top level of the forecast model as there is a tendency in the FGGE 6 layer run for biases to develop. It was also felt that a reevaluation of the SATEM errors statistics was necessary due to recent changes in NESDIS spacecraft performance and processing.

Table 3
Standard deviations (meters) of the differences between collocated SATEMs and radiosondes (±3 hr, ±150 km). The right columns contain the radiosonde error standard deviations as they were prescribed in the analysis program working with 11 layers

Layer (hPa)	Clear	Cloudy and Microwave	Radiosondes
1000-700	23	29	5
700-500	18	21	6
500-300	27	31	10
300-100	54	53	16
100-50	39	38	12
50-30	37	37	12
30-10	98	90	43

SATEM 7-layer Vertical Correlations
Clear

Layer (hPa)	1000-700	700-500	500-300	300-100	100-50	50-30	30-10
1000-700	1.0						
700-500	.23	1.0					
500-300	-.18	.27	1.0				
300-100	.10	-.07	-.16	1.0			
100-50	-.12	-.02	.16	-.11	1.0		
50-30	-.01	.08	.14	.15	.0	1.0	
30-10	.04	-.06	.03	0.	.11	.21	1.0

Cloudy and Microwave

Layer (hPa)	1000-700	700-500	500-300	300-100	100-50	50-30	30-10
1000-700	1.0						
700-500	.22	1.0					
500-300	-.34	.22	1.0				
300-100	-.03	.25	-.29	1.0			
100-50	.04	.06	-.02	-.06	1.0		
50-30	-.10	-.05	.06	.20	-.01	1.0	
30-10	-.01	-.13	-.06	.06	.0	.33	1.0

A global set of 2000 radiosonde/SATEM matches within ±3 hours and 150 km were collected during February 1987. These data were quality controlled using the six hour forecast to eliminate any gross errors in the data. The SATEM matches were then split into two classes: clear and cloudy/microwave, since little difference in accuracy could be found between cloudy and microwave profiles. The results of the radiosonde SATEM collocations are shown in Table 3. The clear soundings have somewhat less vertical correlation in the lower layers and a smaller standard error than the cloudy/microwave soundings. Generally biases in the data set were smaller. The correlation in the stratosphere was very similar for both classes. The corresponding radiosonde error standard deviation used by the analysis are also shown and these appear too small, in comparison with SATEM errors, particularly in the stratosphere (McMillin et al., 1987).

4.3 Response in the vertical

A special diagnostic package has been written (P. Lönnberg, pers comm), to allow the operational code to be used to study the effect of a single observation on the analyses at the observation point and at adjacent grid points. The analysis is performed in two slabs in the vertical (Lönnberg et al., 1986). Fig. 7(a) compares the response of the 11 layer SATEM statistics to 7 layer statistics, in the lower slab when the input is a uniform normalised departure. It can be seen that the 11 layer statistics give an undesirable response: the satellite data is less accurate near the surface and near the tropopause, and the analysis gives more weight to the SATEM at these layers and less in the mid-troposphere. In contrast the 7 layer statistics give a more even response in the vertical with a little less overall weight. Similar results are found in the stratosphere (Fig. 7(b)): the current 11 layer statistics gives an uneven response whereas the new 7-layer response is more even.

Fig. 7(c) shows the final effect of this SATEM increment at a nearby grid point in the analysis. The input normalized profile is shown as a dotted line and the responses in all model variables are also shown. The response to the model temperature profile is as expected with a minimum near the surface and near the tropopause. The minimum at the top could be avoided if the top SATEM layer was extended above 10 hPa. However, this minimum does not appear to affect the assimilation.

This diagnostic package has proved to be a valuable tool for tuning the analysis of satellite data. Some small adjustments were made to the calculated SATEM's statistics in order to achieve the smooth response. Table 4 shows the final statistics which are now used in data assimilation experiments.

Table 4
Standard errors for SATEM (Meters)

Layer (hPa)	Clear	Cloud and Microwave
1000-700	23	28
700-500	14	17
500-300	23	28
300-100	37	39
100-50	33	33
50-30	30	30
30-10	62	62

SATEM 7-layer Vertical Correlations
Clear

Layer (hPa)	1000-700	700-500	500-300	300-100	100-50	50-30	30-10
1000-700	1.00						
700-500	0.22	1.00					
500-300	-0.18	0.27	1.00				
300-100	0.00	0.00	-0.16	1.00			
100-50	0.00	0.00	0.16	0.00	1.00		
50-30	0.00	0.00	0.00	0.15	0.00	1.00	
30-10	0.00	0.00	0.00	0.00	0.00	0.27	1.0

Cloudy and Microwave

Layer (hPa)	1000-700	700-500	500-300	300-100	100-50	50-30	30-10
1000-700	1.00						
700-500	0.22	1.00					
500-300	-0.34	0.22	1.00				
300-100	0.00	-0.25	-0.29	1.00			
100-50	0.00	0.00	0.00	0.00	1.00		
50-30	0.00	0.00	0.00	0.20	0.00	1.00	
30-10	0.00	0.00	0.00	0.00	0.00	0.27	1.0

5. LIST OF MODIFICATIONS FOR REVISED USE OF SATEMs IN OPERATIONS

Using the experience accumulated in operations, in the FGGE experiment (sections 2 and 3), and also all the statistical computations on the SATEM observation errors (section 4), the following list of modifications to the analysis has been tested:

a) Use of 7 SATEM layers in the vertical instead of 11: 1000/700 hPa, 700/500 hPa, 500/300 hPa, 300/100 hPa, 100/50 hPa, 50/30 hPa and 30/10 hPa.

b) Rederivation and retuning of vertical covariance matrices for SATEM observation errors. The new statistical sets contain two 7x7 covariance matrices: one for clear soundings and one for cloudy or microwave soundings. The clear soundings are found to be of better quality. Quality is assumed the same for cloudy and microwave soundings.

c) Revision of the SATEM quality control checks. The quality checks are handled through analysis flags varying from 0 (good quality) to 3 (definitely bad data). General information on the quality control can be found in Shaw et al (1987). We have made the first guess check more stringent and the multi-level check against the first guess has been revised as follows: the complete satellite sounding is rejected whenever there are (i) one or more layers with analysis flag 3, (ii) two or more layers with analysis flag 2, or (iii) three or more layers with analysis flag 1. The previous operational multi-level check for SATEMs is based on the same principle as for radiosondes: the whole sounding is rejected when we find 4 adjacent layers (or more) with flag 1.

d) Remove the blacklisting of SATEMs below 400 hPa near the poles (South of 60°S, North of 70°N). Blacklisting of SATEMs South of 60°S and North of 70°N was introduced operationally in Autumn 85 following data quality problems in polar areas. In Spring 86 this blacklisting was found detrimental because the assimilation could not cope with a total absence of data in polar areas: the stratosphere became very noisy. A 400 hPa limit was then used as a compromise, and the satellite data have then been used only above 400 hPa South of 60°S and North of 70°N.

This special blacklist is now removed. The quality problems observed from NOAA6 and NOAA9 have disappeared and the use of thicker layers in the vertical reduces the detrimental effect of biased data.

e) The horizontal correlation of SATEM observation error is put to zero when the two SATEMs belong to different satellites. It is also reduced by a factor which depends on the time difference between the two observations, when the observations belong to the same satellite (see Unden, 1987).

f) Modification of the data selection near the poles (see Unden, 1987).

g) Modify the weights given to different analysis boxes in the horizontal overlapping technique (see Unden, 1987).

h) Introduce a vertical overlap between the two slabs for wind analysis (in the present analysis it is done only for temperatures). (See Unden, 1987).

The list of modifications a) to h) is referred to as the "7-layer run" in this paper.

6. COMPARISON OF THE 7-LAYER ASSIMILATION WITH OPERATIONS AND WITH A "NO-SATEM" EXPERIMENT

The previous set of modifications (Section 5) has been tested in a 15½ day assimilation, from 30 January 1987, 00Z, until 14 February 1987, 12Z. The 7-layer run has been compared with the operational run (11 layers), and with a "NO-SATEM" assimilation. A series of 10-day forecasts have been run from each day (12Z) from the parallel assimilations during 31 January to 15 February. This February period was selected for the following reasons:

- The availability of two polar orbiting satellites, NOAA9 and NOAA10;

- The rapid evolution of the synoptic situation from 5 February onwards, especially over North Atlantic and Europe;

- The existence of a very deep vortex in the stratosphere over Russia, particularly at the beginning of the period;

- Several cases of poor medium range forecasts (e.g. 1st February 1987, documented in Radford (1987), Sommeria et al. (1987)).

- The availability of all the raw radiances, from NESDIS, for the period (15 January - 15 February 1987). This data will be used for several other satellite experiments for which the present experiments will provide controls.

6.1 Evaluation of the 3 parallel assimilations

Figures 8 evaluates the three parallel assimilations by comparing the different first-guesses, analyses and the initialisation to radiosonde data. All the statistics (RMS and biases) are accumulated in the last 5 days of the period (00Z only). From those results there are two main points to note:

- The operational assimilation (top) and the 7-layer assimilation (bottom) are extremely close to each other.

- The NO-SATEM assimilation (middle) is affected by a bias (first guess and analysis too warm) above 100 hPa, and even above 200 hPa in the tropics. The first-guess is about 1° too warm compared to radiosondes in the stratosphere.

This is consistent with what has been observed in the FGGE experiment. It suggests that the forecast model becomes too warm in the stratosphere. The only difference is that in this February 87 period, the satellite data managed to correct this bias in the assimilation. In the FGGE experiment they had a tendency to increase it, presumably because the FGGE satellite data were of lower quality than now, and by chance were also too warm in the stratosphere, like the forecast model.

The overall general quality of the SATEMs is illustrated in Fig. 9. which shows statistical plots for NOAA9, accumulated on a 6-day period for the 11-layer assimilation (Fig. 9a) and the 7-layer assimilation (Fig. 9b). Let us note the following points:

- The clear soundings are better than the others. This result was already found in the FGGE assimilations.
- There is no significant difference of quality between NOAA9 (shown) and NOAA10 (not shown);
- The microwave soundings contain strong biases near the surface. From this point of view, the 7-layer assimilation looks better than the 11-layer assimilation, because the vertical integration of the layers reduces the bias.

The number of rejected SATEMs in the assimilation is higher in the 7-layer assimilation than in the operations, especially at the beginning of the period when many SATEMs are rejected in the strong stratospheric vortex over Russia and Asia. This is due to the tightening of the quality control. The rejections were considered justified because the intense sudden warming in progress at the time rendered the SATEM regression statistics invalid for the meteorological situation. In spite of this more stringent quality control, the number of rejected SATEMs reduces towards the end of the assimilation to about 10 (0.2%) and it is still considered far too small, given the SATEM statistics (Fig. 9). In the future we will have to move towards a more severe (and maybe more intelligent) quality control.

Little difference can be seen between the experimental 7-layer and operational 11-layer analyses by studying synoptic maps at the surface, 500 hPa and

200 hPa. However, important differences can be seen in difference maps, even at large scales. Fig. 10 for the 200-300 and 200-100 hPa layers at 1200 UTC on Feb 1 shows the expected "compensation" effect in temperature near the tropopause: the differences have opposite signs at 300/200 hPa compared to 200/100 hPa. A maximum departure of 4°C is shown in the North Pacific. These maps mean that important changes are made in the description of the tropopause by moving from 11 to 7 layers: in the 7-layer assimilation, the internal structure of the layer 300/100 hPa is not driven directly by the SATEMs, but only by the forecast model.

The impact of changes in SATEM use on the Southern Hemisphere analyses is large. Fig. 11 shows the magnitude of the differences between "NO-SATEM" and "11-layer" on 1st February 12Z, for the layers 1000/500 and 500/100 hPa in the Northern Hemisphere. The differences are larger for the lower troposphere. Important differences are seen in all the eastern part of the Pacific Ocean in the layer 1000/500 hPa. The comparisons between maps (a) and (b) on one hand, and maps (c) and (d) on the other hand, show that most of the analysis differences are retained in the 6-hour forecast. An example of a vertical profile for this region is shown in Fig. 12. The SATEMS are often colder than the analyses in the surface boundary layer. This is a common problem with many SATEMs and the differences "NO-SATEM - 11-layer" indicate that the operational run may often be too cold in this cloudy region.

One expected to see larger differences than were actually seen between the NO-SATEM run and the other runs in the stratosphere, because of the bias problem discussed previously, particularly in the region of the strong anomalous vortex over Russia. Some stratospheric maps have been investigated in detail. In the Northern Hemisphere they are extremely close in the different assimilations, up to 10 hPa. The anomalous vortex, which is very strong at the beginning of the period, is described in the same way in all assimilations including the NO-SATEM. The vortex slowly fills and translates during the two week period, with a return to a more normal circulation.

Fig. 13 shows an interesting feature at the end of the assimilation period, at 10 hPa, on the Southern Hemisphere: the NO-SATEM height analysis is very noisy with unrealistic features near the Antarctic continent while the 7 layer analysis or the operational one (not shown) are very smooth. This confirms what was already shown in Pailleux (1986), that in the total absence of data the stratosphere has a tendency to become noisy and drifts away from reality, even from a realistic climate.

Another significant impact of the SATEMs was noted in the humidity analysis: the humidity fields are smoother in the assimilations using SATEMs, while the NO-SATEM assimilation tends to preserve sharper gradients.

6.2 Evaluation of the forecasts run on the 3 parallel assimilations

(a) Hemispheric scores

For 15 cases (31 January to 14 February, 12Z), three 10-day forecast sets have been compared for both hemispheres: operations, 7-layer, NO-SATEM. All the forecasts have been run using the operational model as of 20 February 1987.

Fig. 14, showing accumulated score and scatter diagrams at day 4, illustrates the overall impact of the SATEMs on the forecast:

- In the Southern Hemisphere this impact is very large.

- In the Northern Hemisphere, on average, the SATEMs have no impact on the quality of the forecast. This is a different result, compared with November 1979. It probably means that the level of impact of one observing system is highly dependent on the meteorological situation. It also indicates that the evaluation of all the satellite experiments in this period will be difficult in the Northern Hemisphere, so that our main conclusions will have to be drawn from the Southern Hemispheric impact.

Fig. 15 shows the overall evaluation of the 7-layer forecasts versus the 11-layer forecasts in a similar display, while Fig. 16 compares the RMS error of the 7 layer forecast with 11-layer operations in the Southern Hemisphere. Our conclusions from these figures are:

- Averaged over the Southern Hemisphere, there is a slight advantage for the 7-layer run over the 11-layer one, especially at days 3 to 5. This is more apparent in the RMS temperature error (Fig. 16): in spite of the initial handicap of 1°C due to the use of the 11-layer verification analysis, the 7-layer run manages to "overtake" the 11-layer run just before day 3.

- Averaged over the Northern Hemisphere, the results are again neutral, in spite of a very small signal in favour of the 7-layer run on the scatter diagram at day 4 (see Fig. 15).

When the 15 "triplets" of forecasts are considered, case by case, the only clear signal is the large positive impact of SATEMs in the Southern Hemisphere throughout the period. One might have expected the absence of SATEMs to be more and more detrimental from 30 January to mid-February, as the NO-SATEM assimilation started from the operational first guess (containing the SATEM information) on 30 January 00Z; it usually takes several assimilation cycles to "forget" the SATEM information. This actually did not happen.

Considering the scores, the largest differences between the NO-SATEM forecast assimilation and the 11-layer actually occurs on 2 and 3 February. In the Northern Hemisphere, if the 11-layer, the 7-layer and the NO-SATEM forecast are sorted according to the scores, all the possible orders are found but they vary from case to case. The scores, averaged over the Northern Hemisphere, should not be considered as a perfect indicator of the forecast quality. The scores for the forecasts from 14 February 12Z showed the NO-SATEM and operations (11-layer) scores to be very close, but the 7-layer score is significantly better than either between days 4 to 7. This is not confirmed by the day 6 forecast maps which show some synoptic differences, especially over Northern Atlantic, but from these differences it is almost impossible to judge which forecast is best.

(b) Regional scores

A total of 11 regional areas were used to evaluate the 15 forecasts. The operational runs of February 1987 (OPS) was compared first to the NO SATEM run, then to the 7-layer run.

The neutral impact of SATEMs in the Northern Hemisphere (NO SATEM vs OPS) results from a small positive impact of SATEMs on one side of the hemisphere (America-Atlantic-Europe), and a negative impact on the other side (East Asia and Pacific). A possible explanation is that the SATEM quality during this

February period was adversely affected by the anomalously strong stratospheric vortex over Asia: this could be an important weakness in the current operational satellite retrieval method, which cannot handle such extreme situation.

Fig. 17 shows the impact of the 7-layer modifications (7-layer vs 11-layer) in three different areas: North Atlantic, North America and North Pacific. Note the modest improvement brought by the 7-layer forecasts compared to the 11-layer forecasts in the North Pacific. This improvement tends to cancel the negative impact of SATEMs in this area, suggesting again the importance of data quality problems: by using 7 layers, we smooth a part of the noise and of the SATEM observation error, especially when the quality of the data is poor.

Fig. 18 shows the scores on the North Pole cap (North of 70°N). The NO SATEM assimilation performs slightly better than the 11-layer run (OPS - using only SATEMs above 400 hPa near the pole - see Section 5). And again the 7-layer run tends to cancel this negative impact (see "7-layers" vs "OPS"). The problem highlighted in Fig. 18 is the difficulty of performing a good retrieval in the polar areas covered with ice.

A possible explanation of this neutral situation over the Northern Hemisphere may be the following. The SATEMs contain some good information, but also a lot of noise and a lot of bad data. The analysis manages to retain a large part of the good information, but is also affected by the poor quality data. Depending on the weather situation, the good information may be predominant (then we have the expected positive impact of SATEMs), or the analysis may be degraded due to the poor SATEMs, especially if we use 11 layers. Finally the average impact during this February two week period is neutral.

In the Southern Hemisphere where there is a lack of conventional data, any SATEM data improve the analysis. By reducing the SATEM resolution from 11 to 7 layers, we reduce the level of noise in the data and we get on average a slight improvement in the forecast. Then the clear message coming out from this study is that an improved quality control is absolutely essential for making progress in the use of satellite data. We can reasonably assume that if we could reject noisy, inconsistent and poor quality data, the positive impact of satellite data would appear more clearly. The present quality control checks do not achieve that: they only reject the SATEMs which are obviously wrong.

7. CONCLUDING REMARKS

a) The impact of SATEM data on the current analysis system is always positive in the Southern Hemisphere. However, in the Northern Hemisphere the impact of SATEM data is often dependent on the synoptic situation.

b) SATEM data contain many large errors which are mostly due to problems in cloud clearing or to rain affected microwave data. If more use were made of sea surface temperature and the lower model level 6 hour temperature forecasts, it should be possible to improve the quality control of the SATEMs retrieval process and hence reduce their observational error. There is also a need to revise the quality control in the analysis since it is clear the current tests do not cope satisfactorily with bad satellite data.

c) Investigations of the effect of SATEMs in the analysis system reveal that they have only a modest effect because of their large observational errors. If large departures occur in the first guess then changes due to SATEMs, even if correct, will take a number of assimilation cycles to correct the analysis. The rather low weight assigned to SATEMs is a consequence of their variable quality.

d) In the stratosphere, the present assimilation cannot work properly with a poor data coverage. Stratospheric satellite data are important to keep the assimilation reasonable, especially in the Southern Hemisphere.

e) The use of 7 SATEM layers appears to give improved analyses, but the impact on forecasts is small. In the Southern Hemisphere, where the analysis relies greatly on SATEM data, there is a positive impact. The reduction of data in the vertical will enable more SATEMs to be used in the horizontal. The 7-layer version of the analysis (with all the other modifications described in Section 5) has been implemented operationally by ECMWF on 21 July 1987, 12Z.

References

Kelly, G.A., 1984: The accuracy and information content of satellite atmospheric vertical profile measurements. Proceedings of the ECMWF workshop on the Use and Quality Control of Meteorological Observations, 6-9 November 1984.

Kelly, G.A., 1985: Use of satellite radiance measurements in the ECMWF analysis system. Proceedings of the ECMWF workshop on High Resolution Analysis, 24-26 June 1985.

Lönnberg, P., et al., 1986: The new analysis system. ECMWF Technical Memorandum No. 125.

MacMillin, L.M., and M.J. Harris, 1987: Physical treatment of profile-dependent radiosonde biases: preliminary results (NESDIS internal paper).

Pailleux, J., S. Uppala, L. Illari and L. Dell'Osso, 1986: FGGE Re-analyses and their use at ECMWF. Technical Memorandum No. 126.

Pailleux, J., 1986: Satellite data quality and data assimilation problems near the poles. RD Memo R2313/1757 of 6 May 1986.

Radford, A., 1987: Large inconsistency between two consecutive forecasts. OD Memo 490 of 13 February 1987.

Schlatter, T.N. and Branstator, G.W., 1979: Estimation of errors in Nimbus 6 temperature profiles and their spatial correlation. Mon.Wea.Rev., 107, 1402-1413.

Shaw, D.B., P. Lonnberg, A. Hollingsworth and P. Undèn, Data Assimilation; The 1984/85 revisions of the ECMWF mass and wind analysis. Q.J.R. Met.Soc., 1987, 113 pp 533-566.

Sommeria, G., et al. 1987: Research work related to operational forecast performance during winter 86/87. RD Memo file 1997.

Unden, P., 1987: Analysis modifications concerning treatment of satellites, data selection and analysis grid point evaluation. RD Memo R2314/1996 of 17 July 1987.

Uppala, S., A. Hollingsworth, S. Tibaldi, P. Källberg, 1984: Results from two recent observing system experiments at ECMWF. Proceedings of the ECMWF Seminar/Workshop, 1984, Vol. 1.

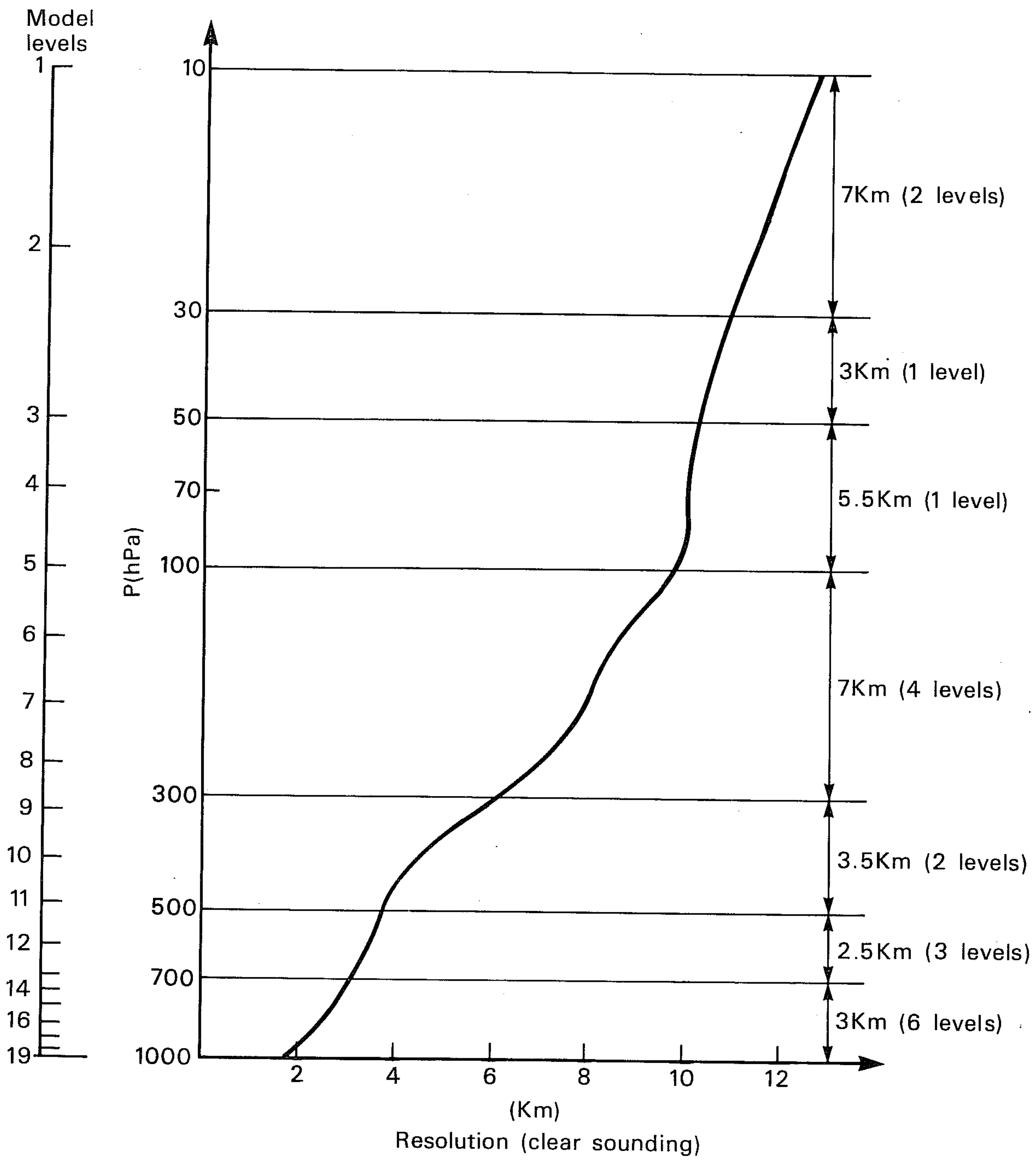


Fig. 1 Vertical resolution of the HIRS instrument with ECMWF 19 model levels and 7 SATEM thickness layers used by the ECMWF analysis.

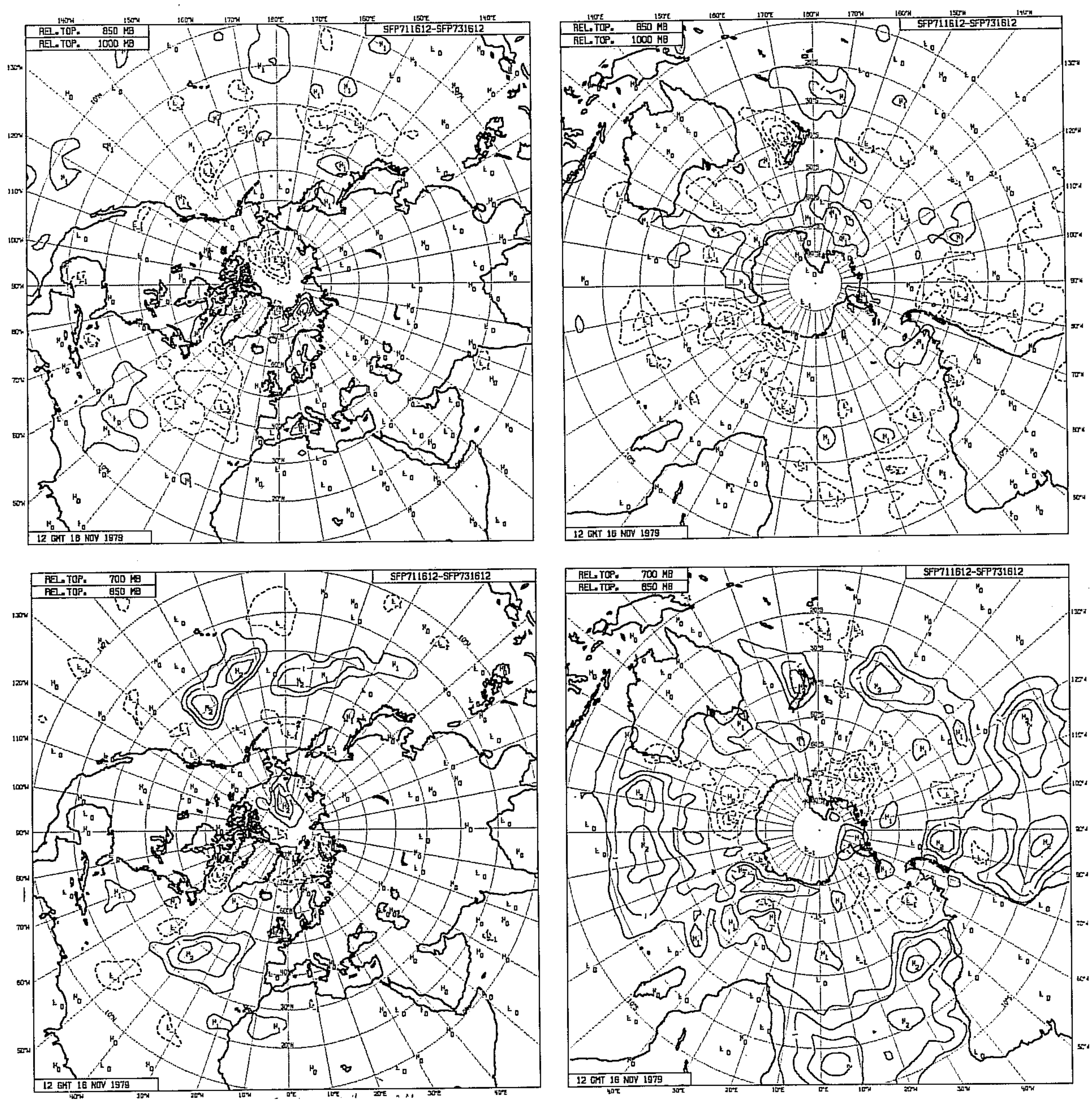


Fig. 2 First guess temperature differences between 6-layer and 11-layer assimilations on 16 November 1979, 12Z, for 1000/850 hPa (top) and 850/700 hPa (bottom).

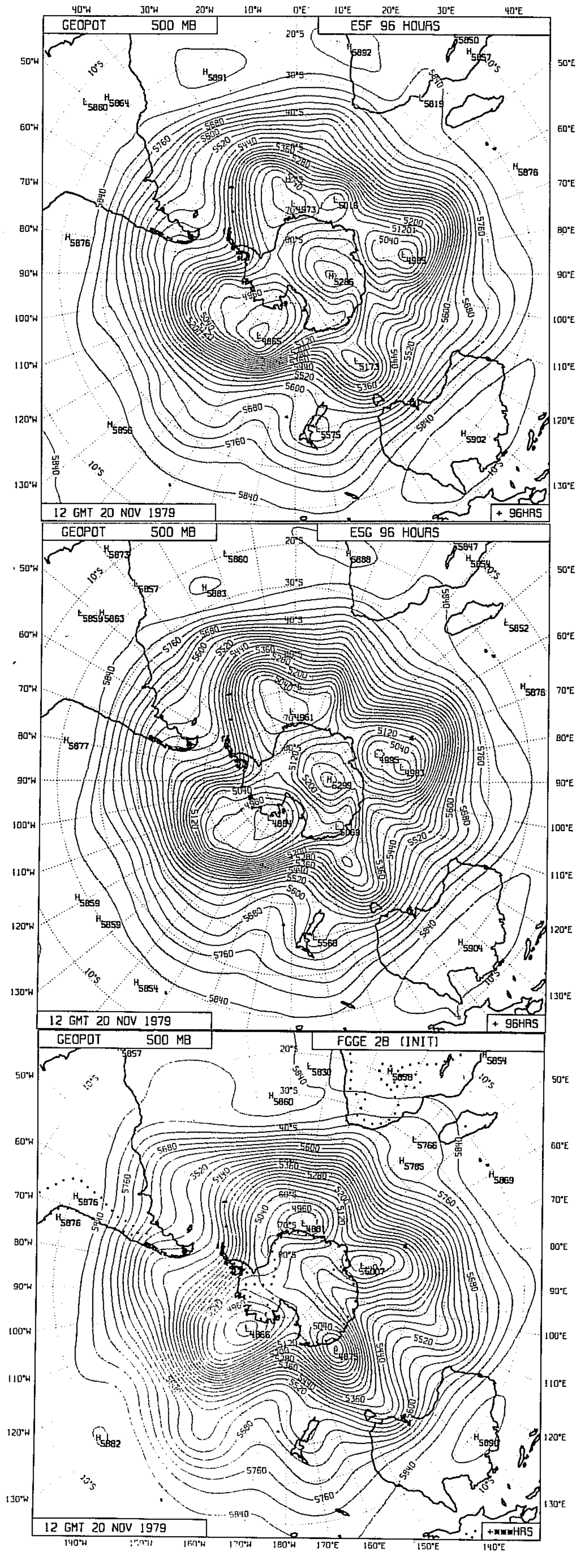


Fig. 3 Day 4 500 hPa Southern Hemisphere forecasts based on 16 November 12Z and comparison with the verification analysis. Top: forecast based on the 11-layer assimilation (E5F); Middle: forecast based on the 6 layer assimilation (E5G); Bottom: verification analysis for 20 November 1978, 12Z.

Northern Hemisphere

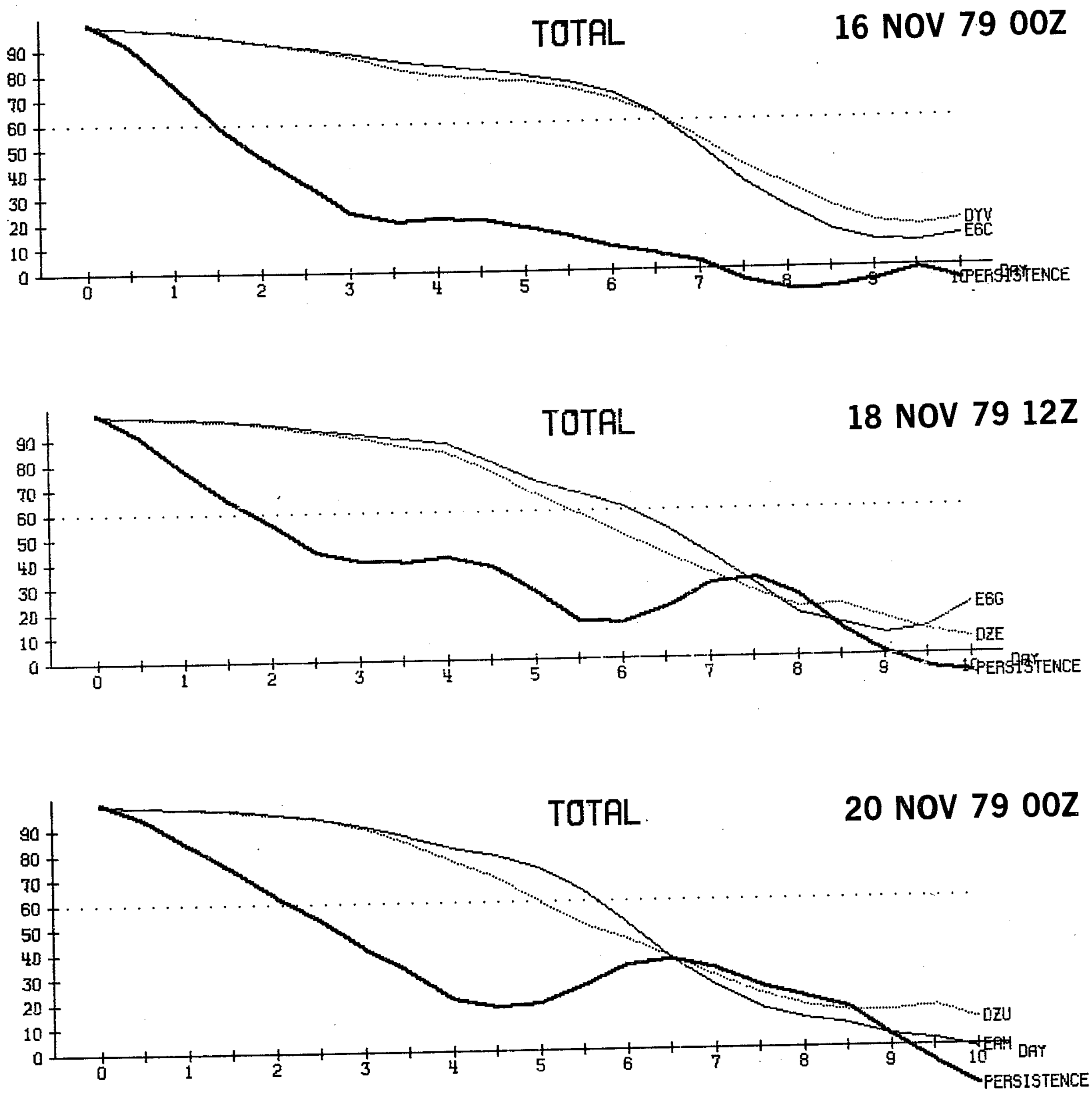


Fig. 4 Anomaly correlations for three pairs of forecasts from 11 layer assimilation and NO SATEM assimilation on Northern Hemisphere. Full line: 11 layer forecast. Dotted line: No SATEM forecast.

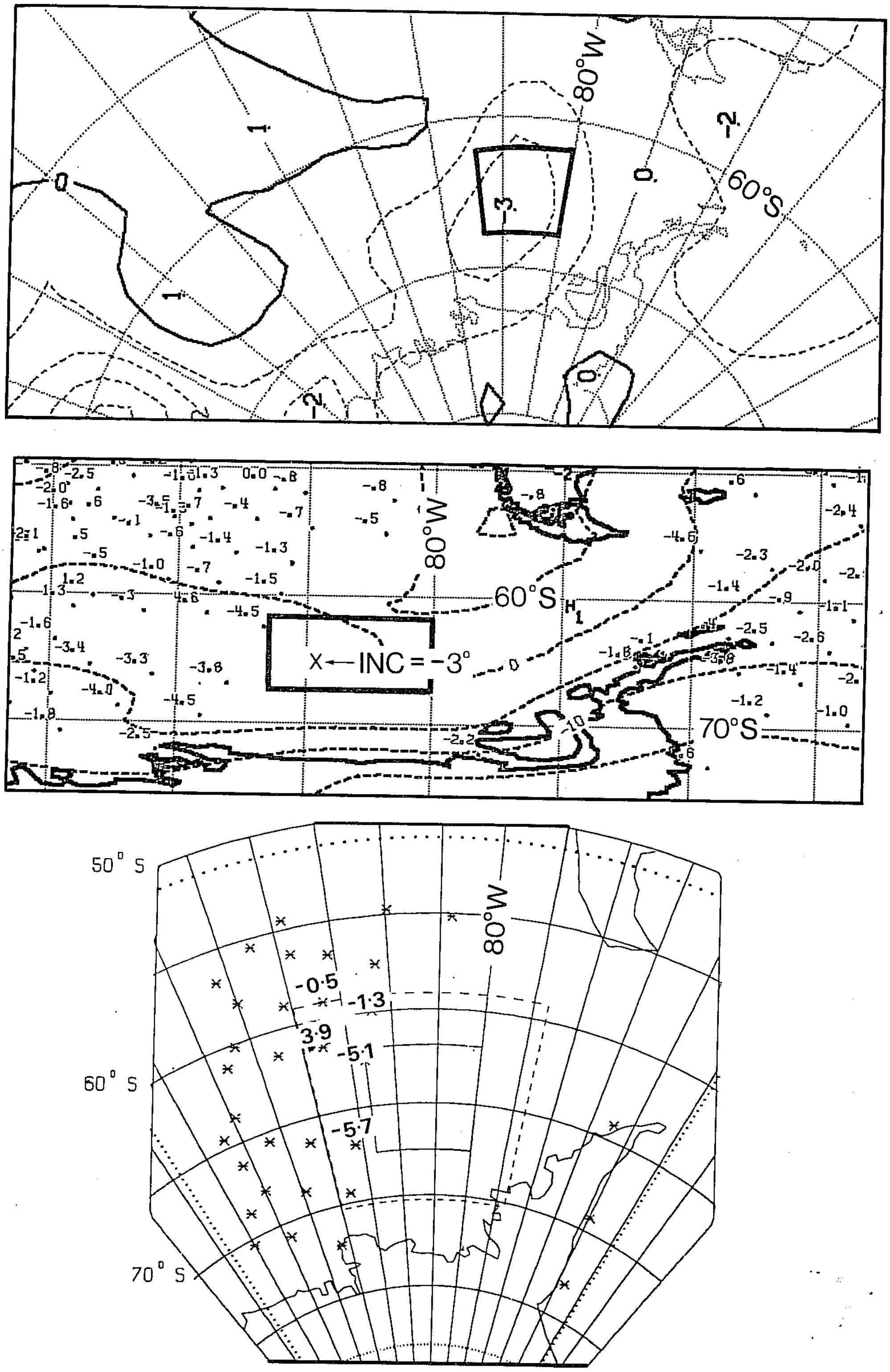


Fig. 5 "Satellite only" assimilation for 14th November 12Z. Maps showing large increments between satellite orbits illustrating problem in analysis extrapolation of bad satellite data.
Top: analysis increments for the 1000/850 hPa temperature.
Middle: the contours show the first-guess field and the increments are plotted at the observation points.
Bottom: the satellite observations retained by the data selection for inner box are shown by *, and five typical values are plotted (in degrees C).

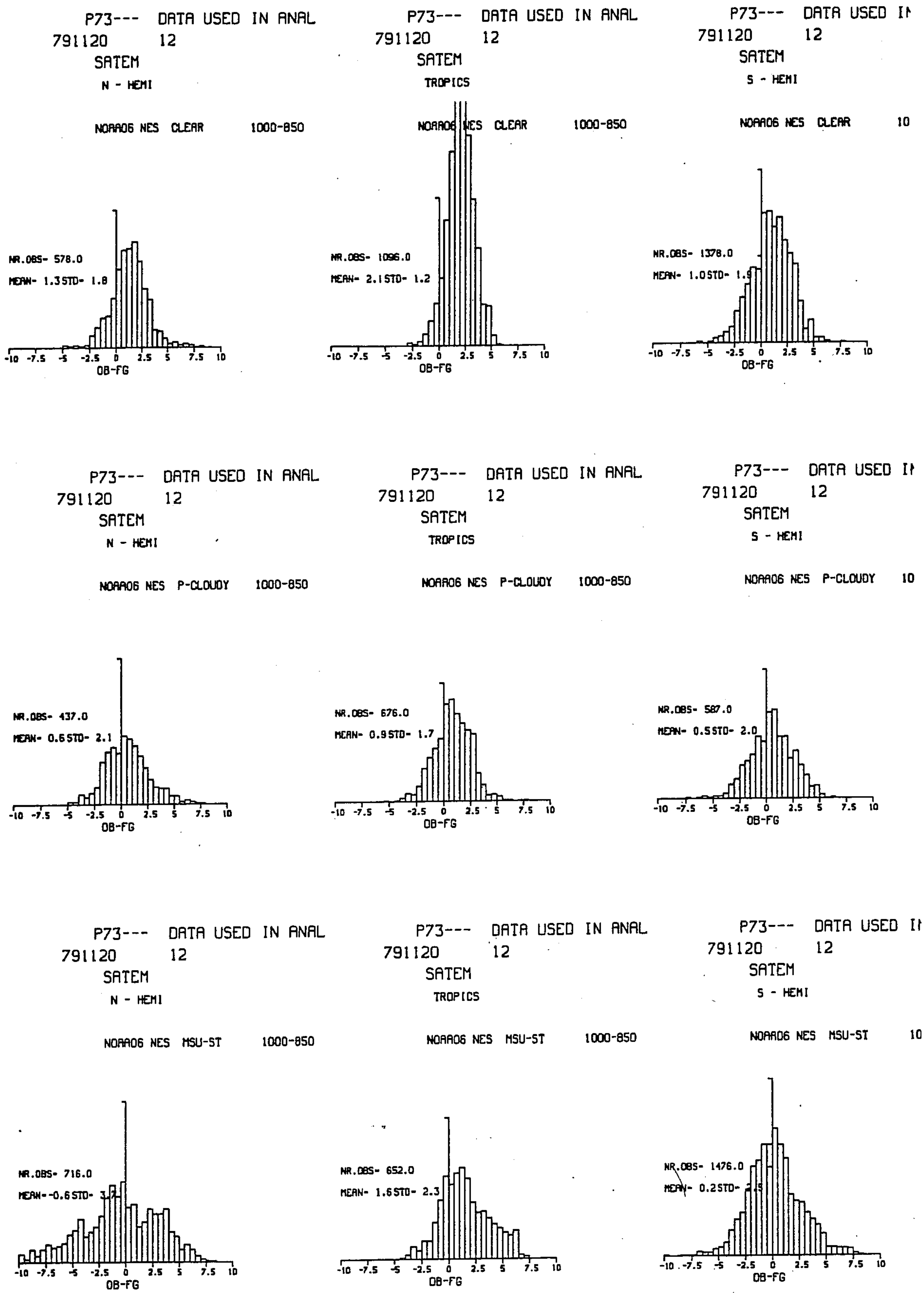


Fig. 6 Accumulated histograms of 11 layer SATEM assimilation for clear (top), cloudy (middle) and microwave (bottom) NOAA-6 SATEMs. The histograms are made on the departures "observation-first guess" for 1000/850 hPa mean temperature.

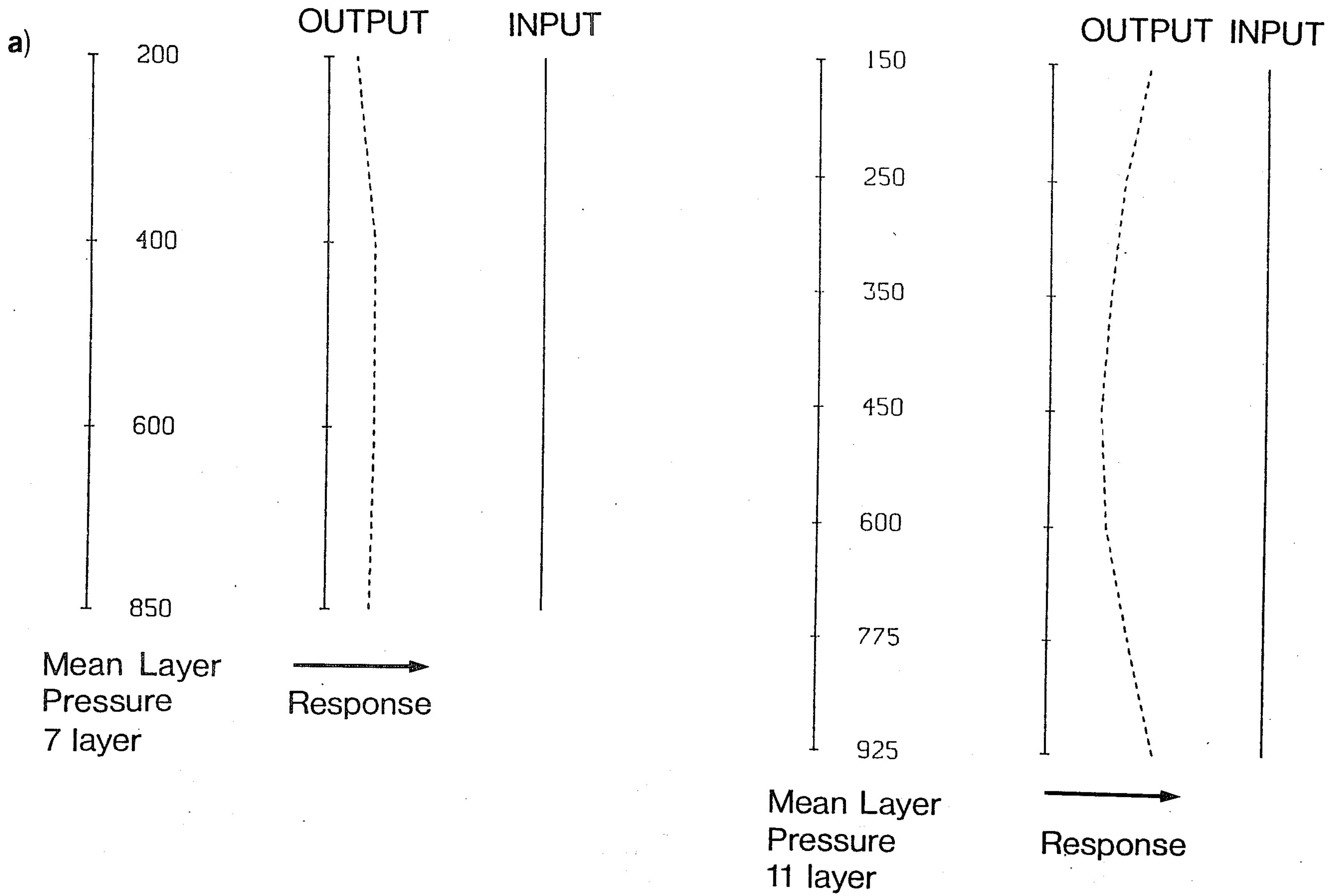


Fig. 7 Response of the analysis in the vertical to a normalized SATEM increment. (a) effect on lower analysis slab for 11 layer SATEMs and 7 layer SATEMs.

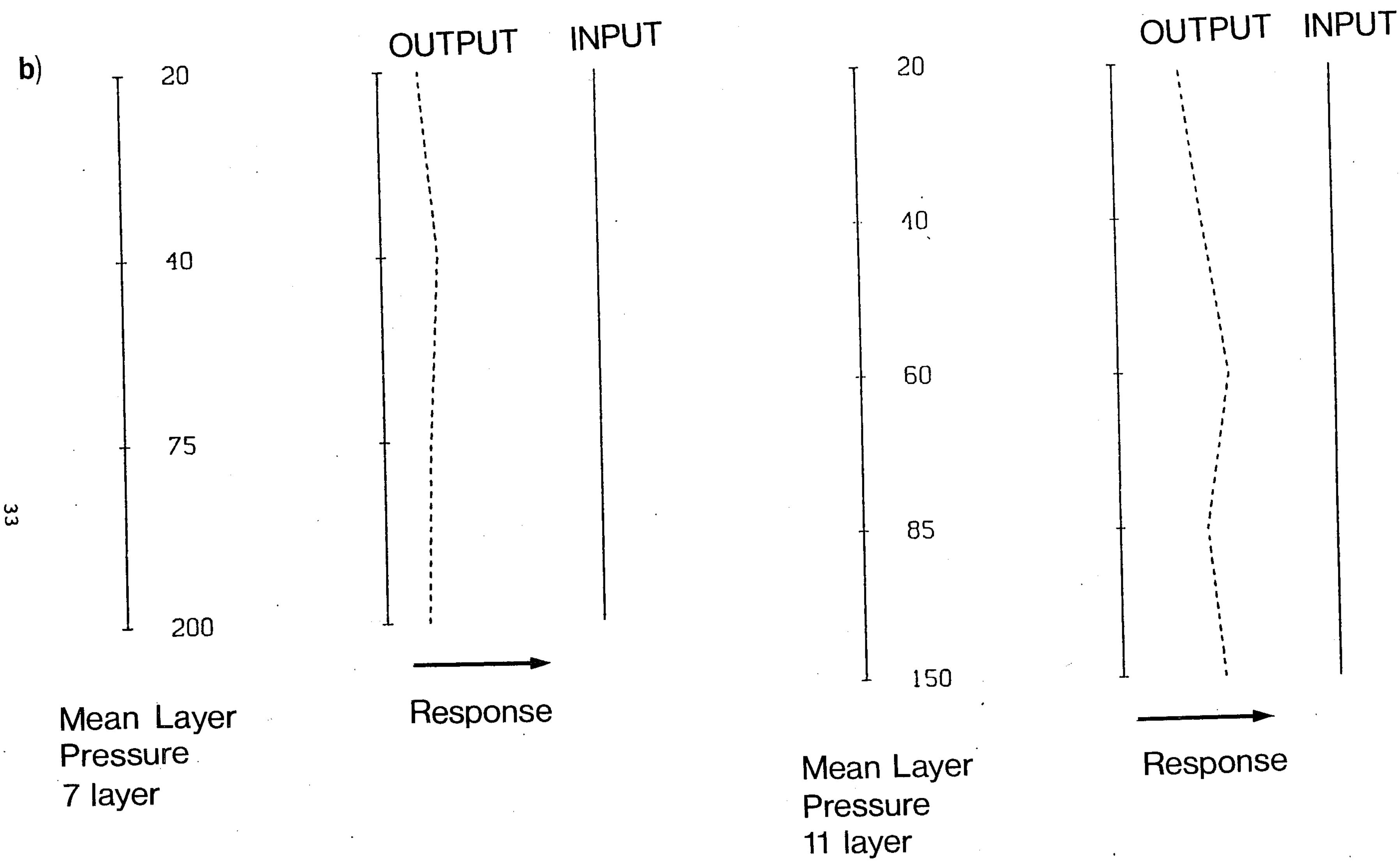


Fig. 7 Response of the analysis in the vertical to a normalized SATEM increment. (b) effect on upper slab.

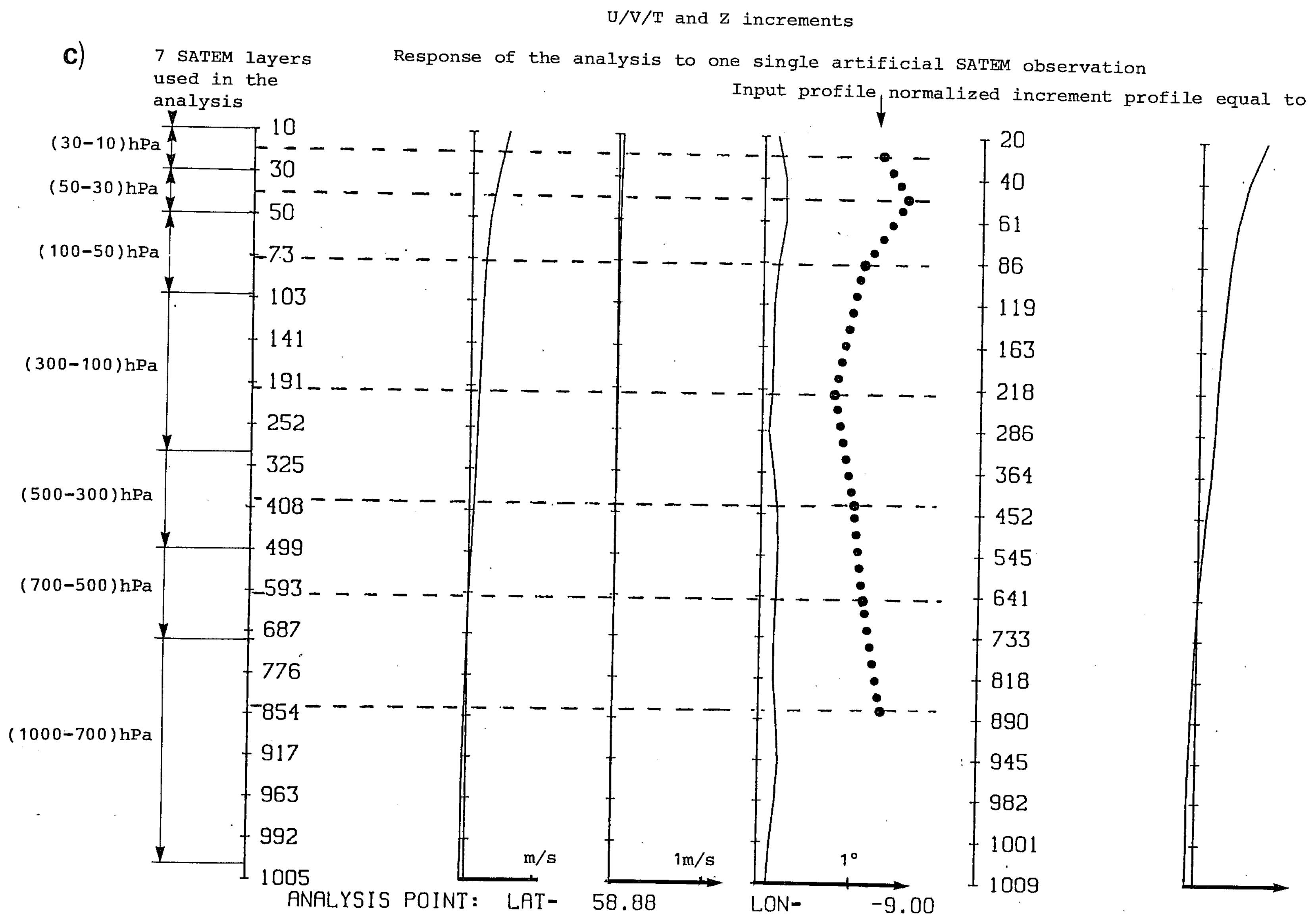


Fig. 7 Response of the analysis in the vertical to a normalized SATEM increment. (c) effect on nearby grid point of 7 layer normalized increment.

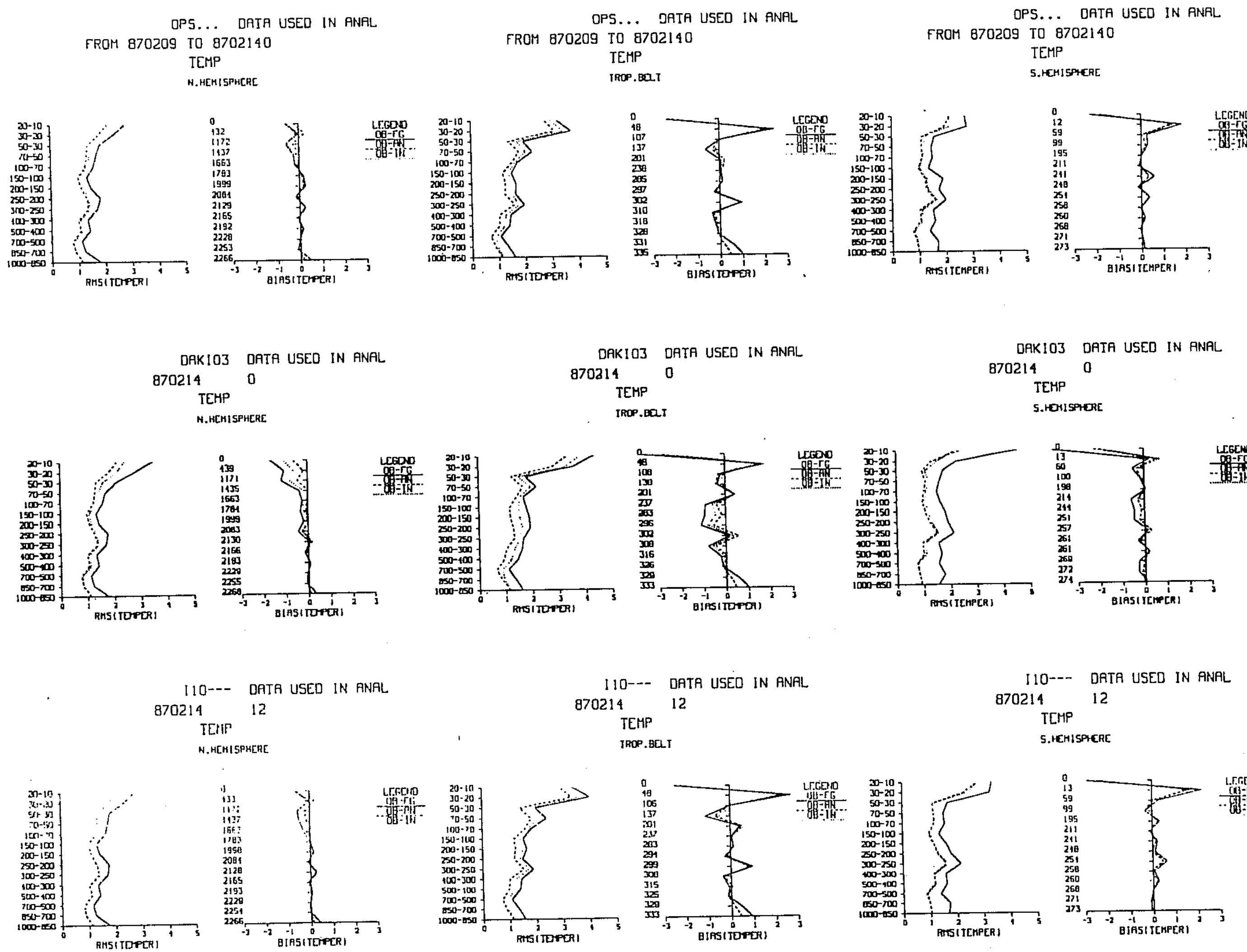


Fig. 8 Accumulated statistics plots for 11 layer assimilation (OPS), NO SATEM assimilation (IO3) and 7 layer assimilation (I10), for temperature differences between radiosonde observation and first-guess, analysis, initialisation.

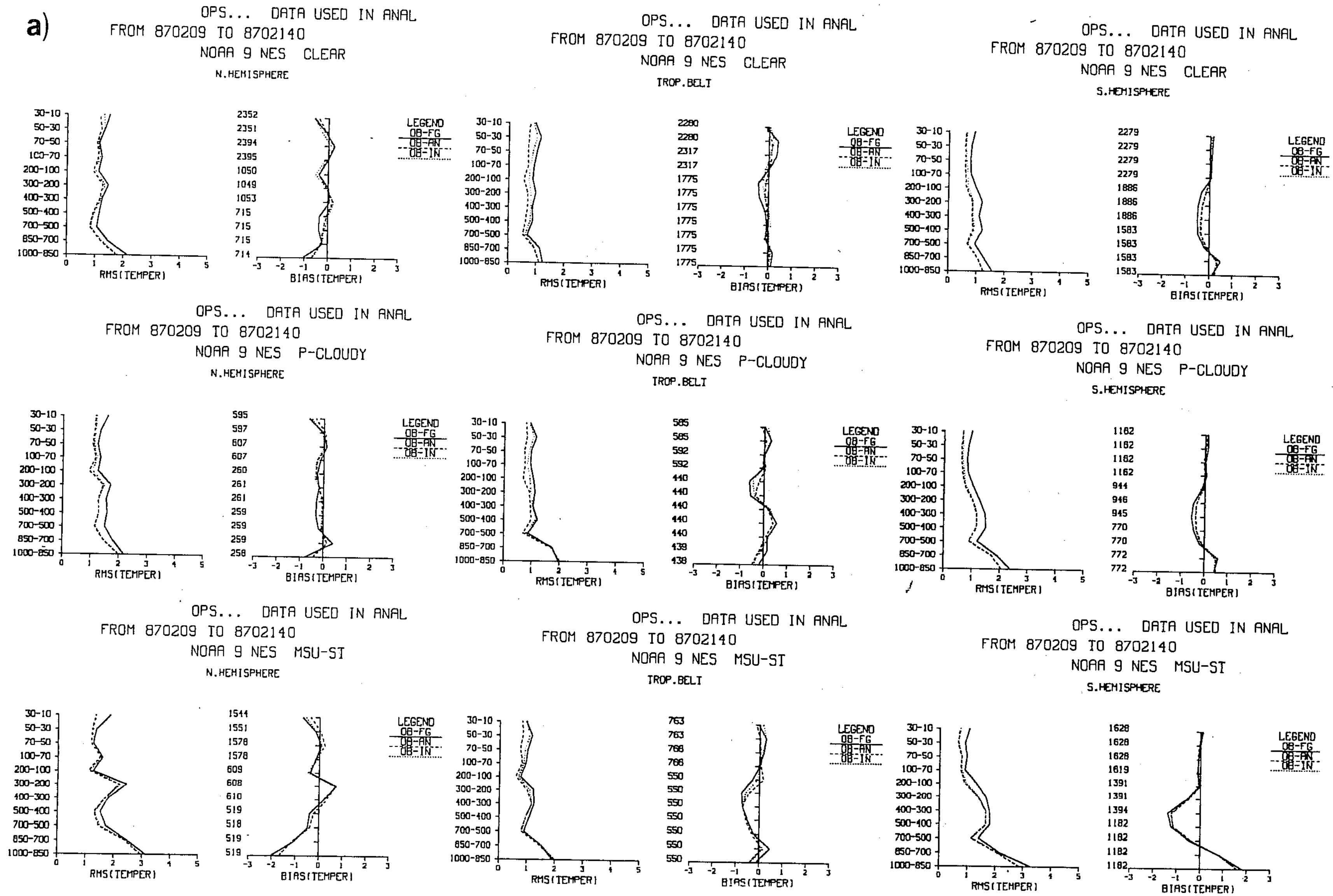


Fig. 9 Accumulated statistics plots for NOAA9 SATEMs (clear, cloudy and microwave). (a) NOAA-9 11-layer (OPS)

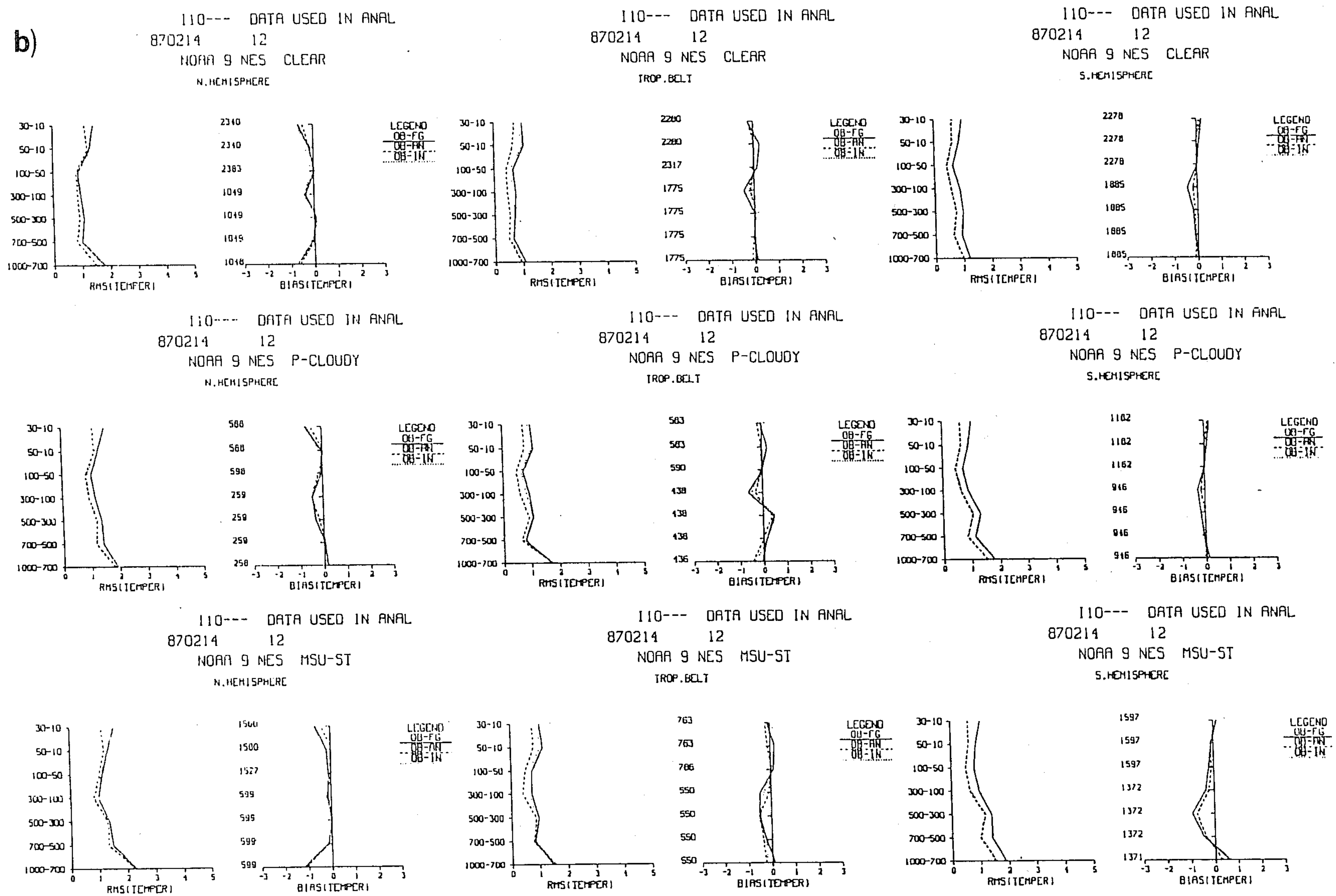


Fig. 9 Accumulated statistics plots for NOAA9 SATEMs (clear, cloudy and microwave). (b) NOAA-9 7-layer (I10).

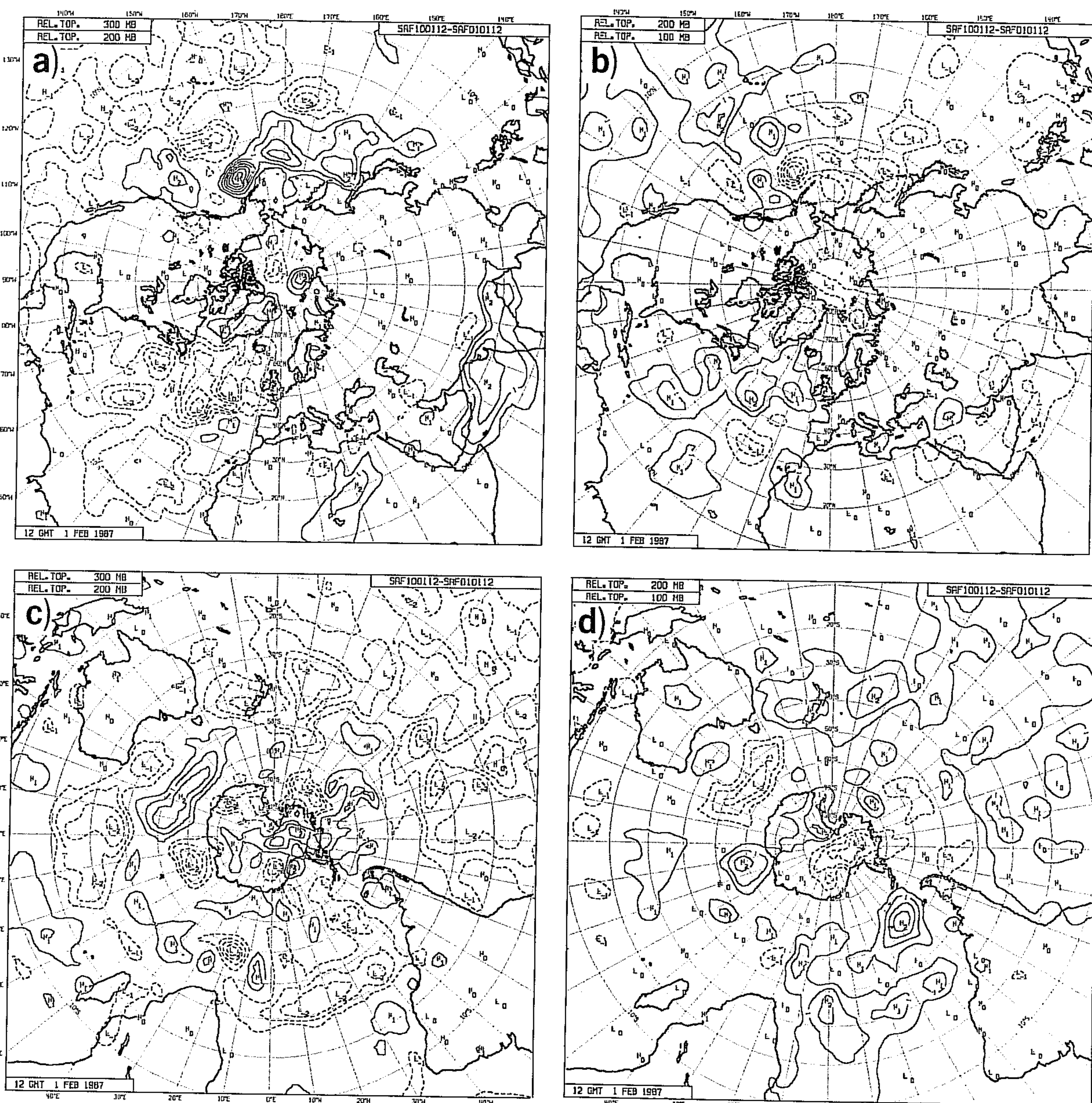


Fig. 10 Temperature difference maps between 7-layer analysis and 11-layer analysis for February 1987 12Z. (a) Northern Hemispheric 300/200 hPa mean temperature differences. (b) 200/100 hPa as (a). (c) same as (a) Southern Hemisphere. (d) Same as (b) Southern Hemisphere.

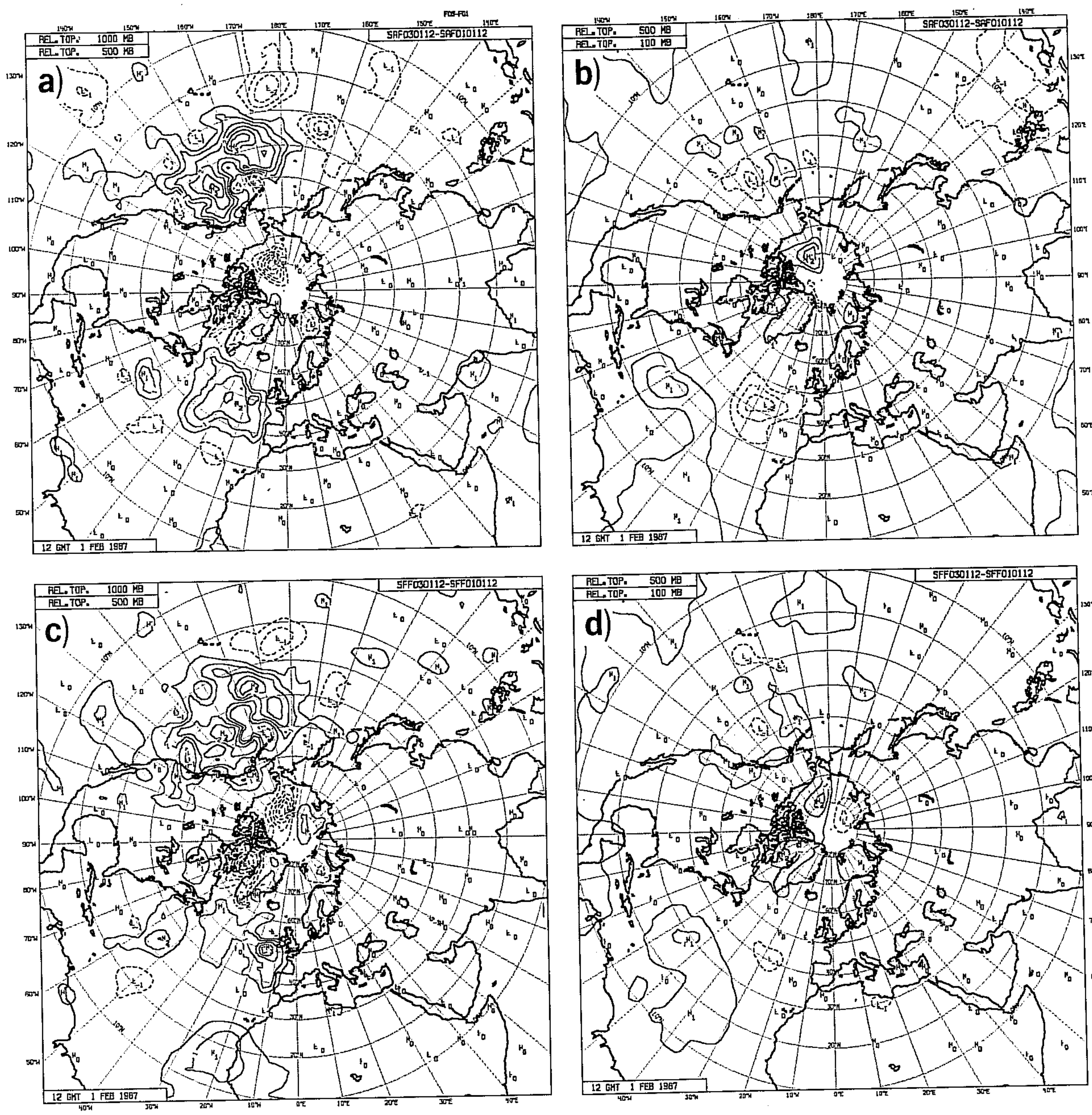


Fig. 11 Difference maps for mean temperature for Northern Hemisphere between the NO SATEM assimilation and the 11 layer SATEM assimilation. (a) Analysis differences for 1000/500 hPa layer. (b) Analysis differences for 500/100 hPa layer. (c) 6 hr forecast for 1000/500 hPa layer. (d) 6 hr forecast for 500/100 hPa layer.

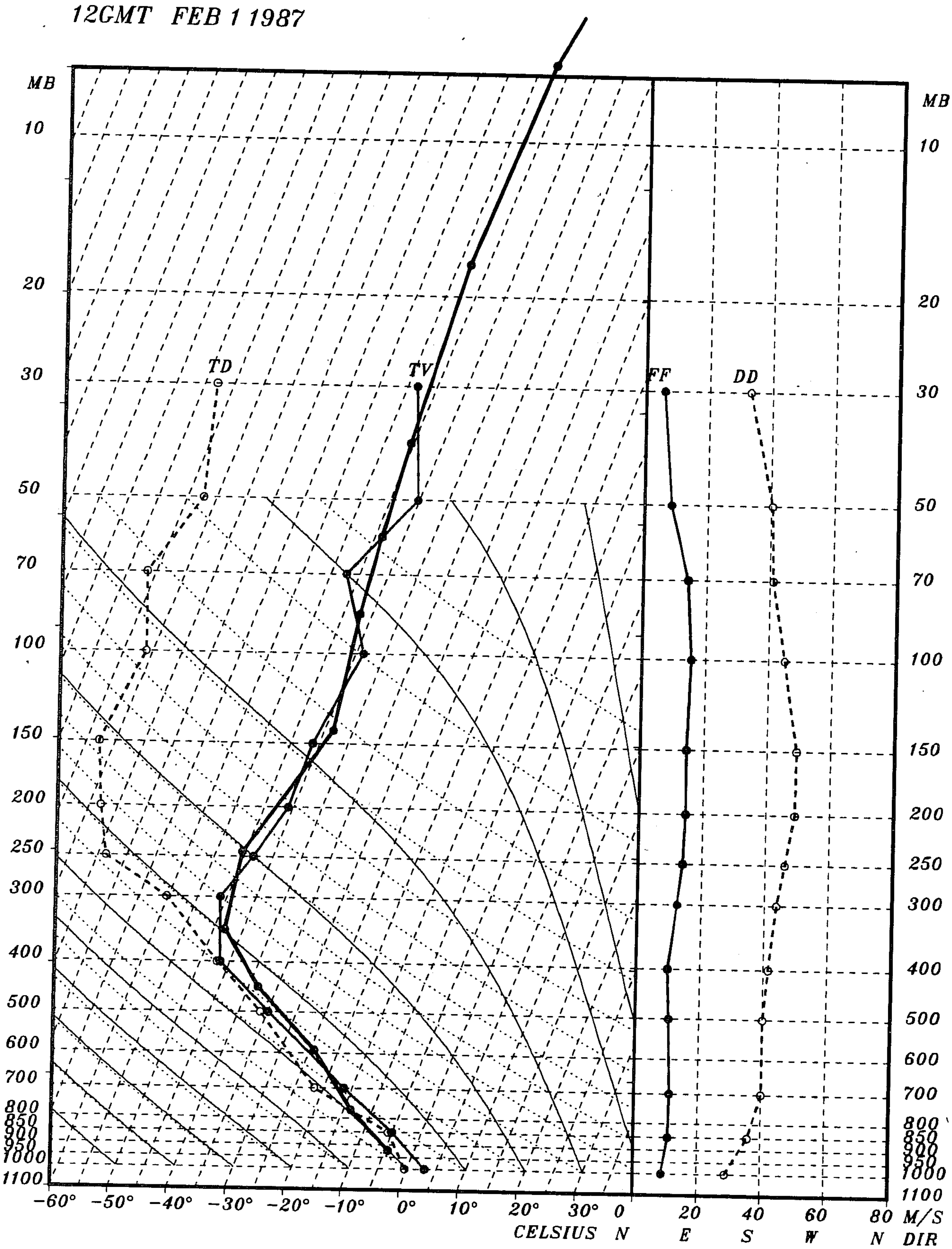
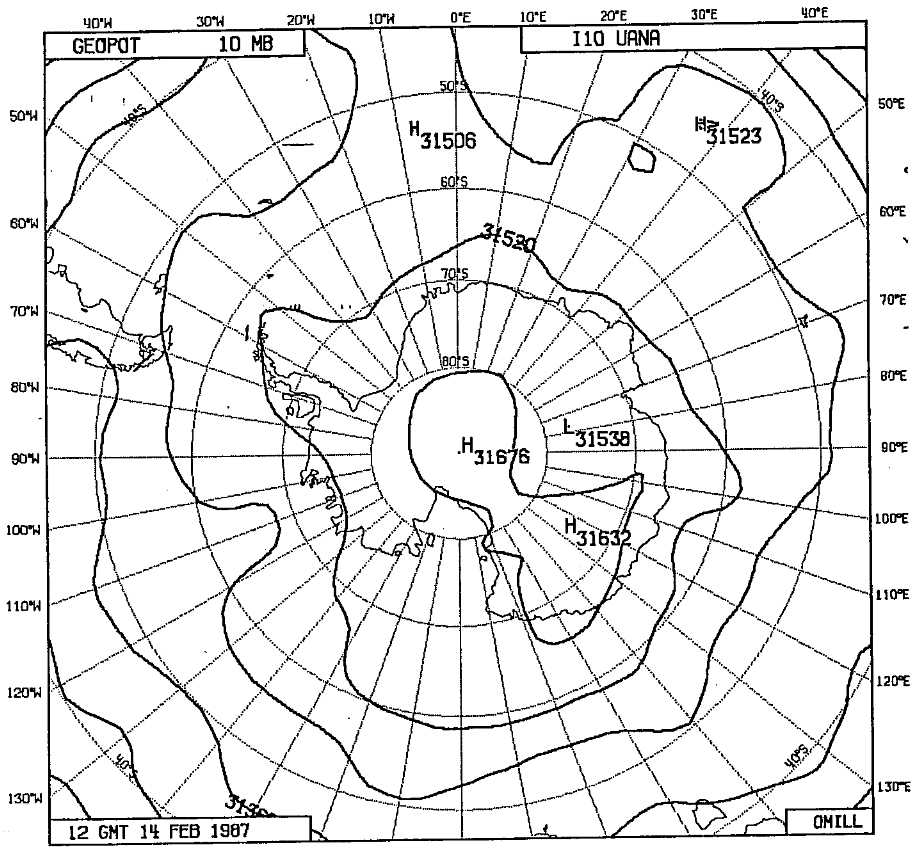


Fig. 12 Vertical temperature profiles of a SATEM and analysis on 1 February 1987, 12Z. Thick line: SATEM profile. All the other curves are from the 11 layer analysis.



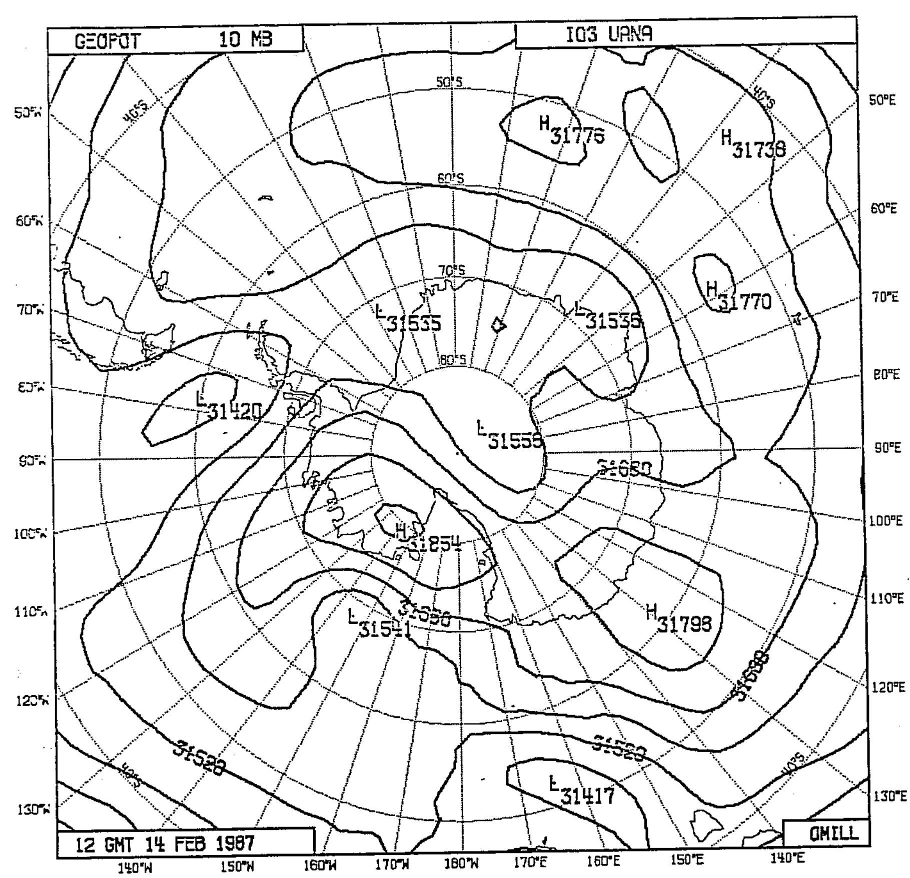


Fig. 13 10 hPa geopotential height analysis from 7 layer assimilation (top) and NO SATEM assimilation (bottom) on 14 February 1987, 12Z.

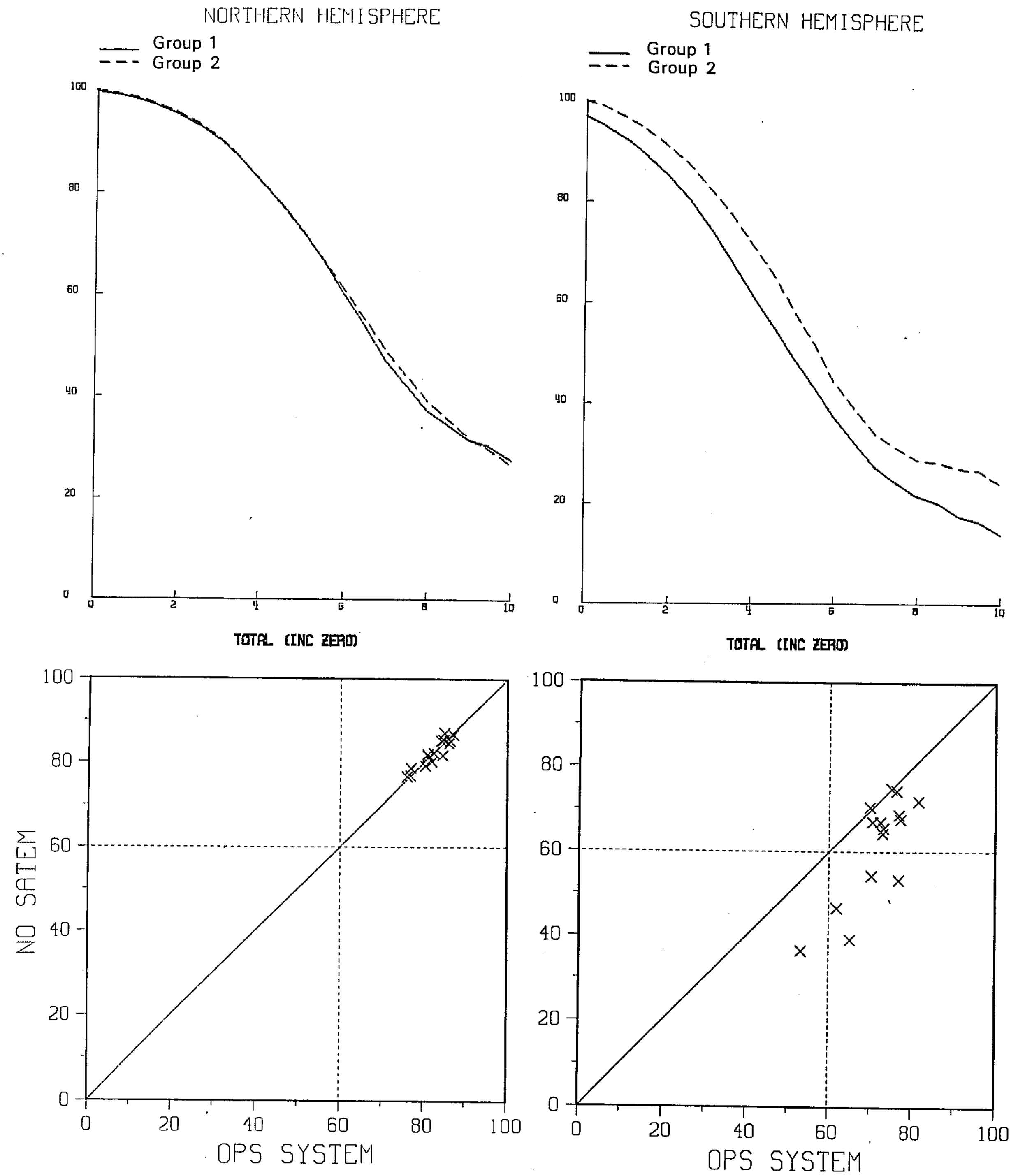


Fig. 14 Accumulated forecast scores on 15 cases comparing NO SATEM (full line) and (dashed line) (top).
Scatter diagrams for NO SATEM versus operations at day 4 (bottom).
Left: Northern Hemisphere. Right: Southern Hemisphere.

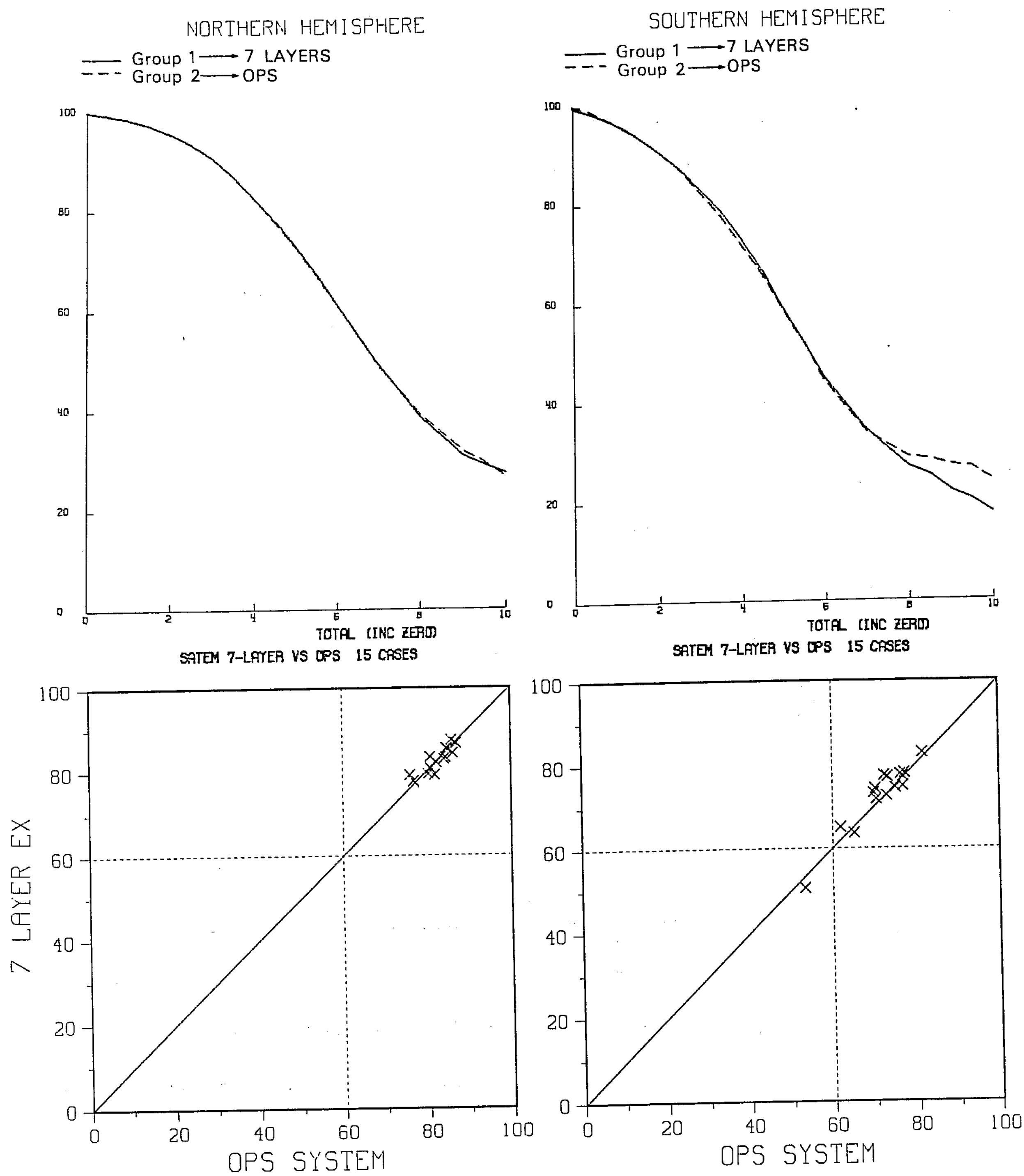


Fig. 15 Accumulated forecast scores on 15 cases comparing 7 layer (full line) and 11 layer (dashed line) (top).
Scatter diagrams for 7 layer versus 11 layer at day 4 (bottom).
Left: Northern Hemisphere. Right: Southern Hemisphere.

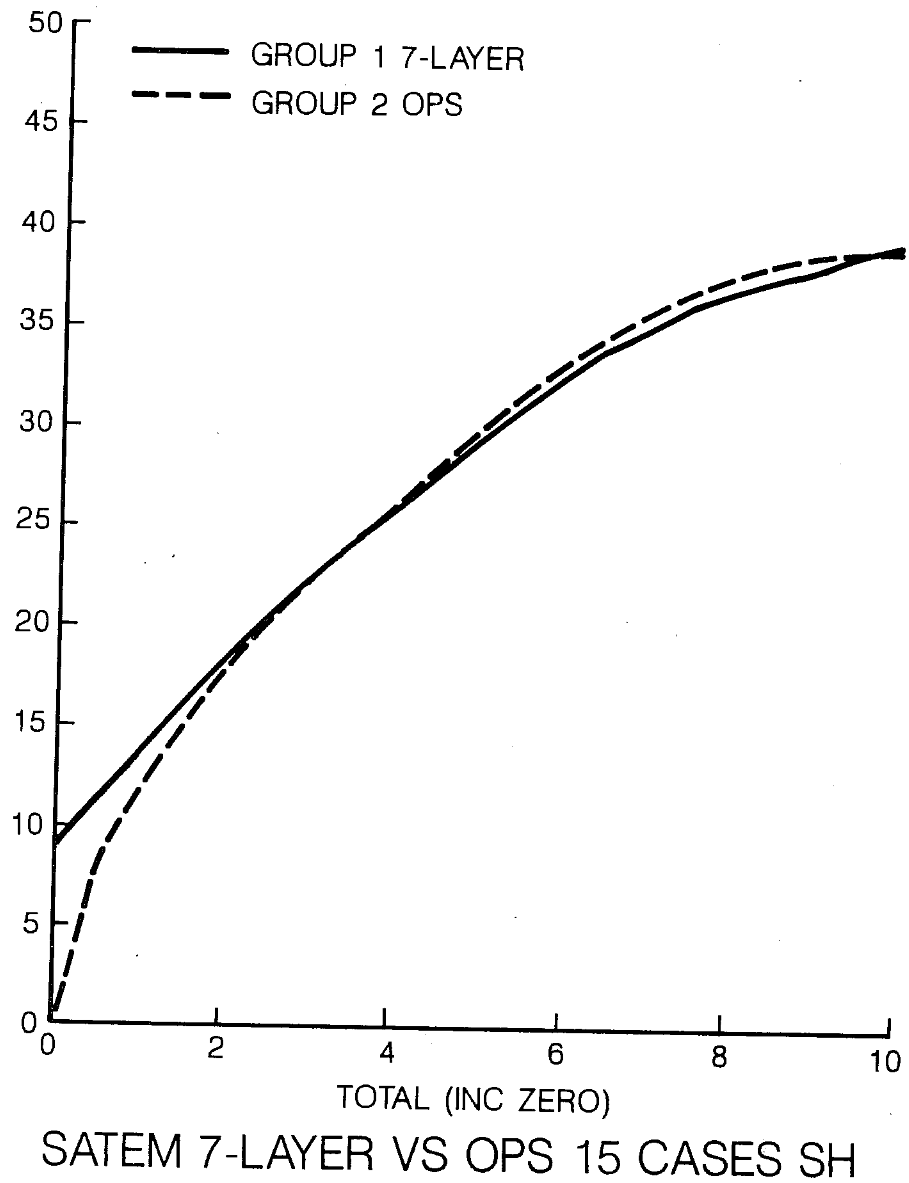
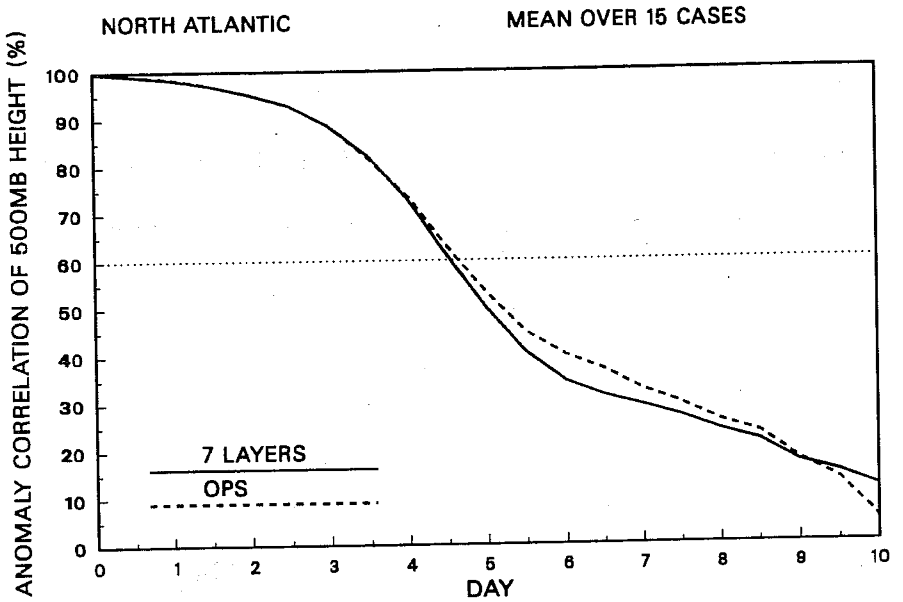
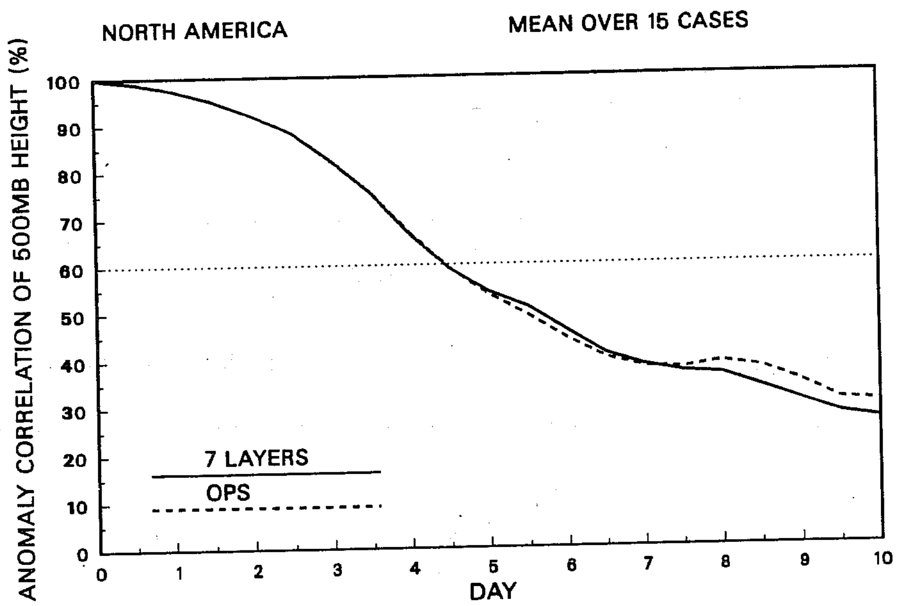


Fig. 16 Mean RMS temperature errors for 7 layer assimilation versus 11 layer assimilation for the Southern Hemisphere.





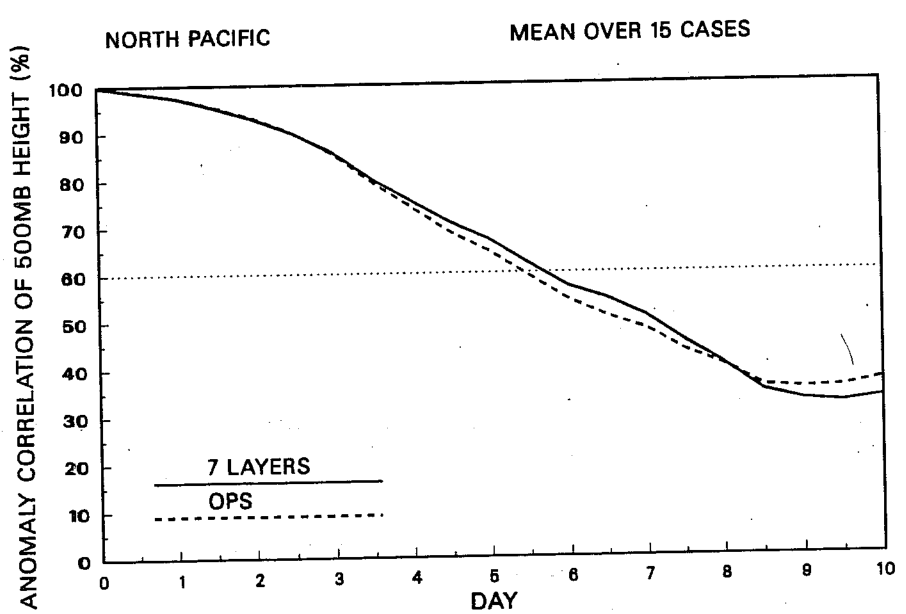
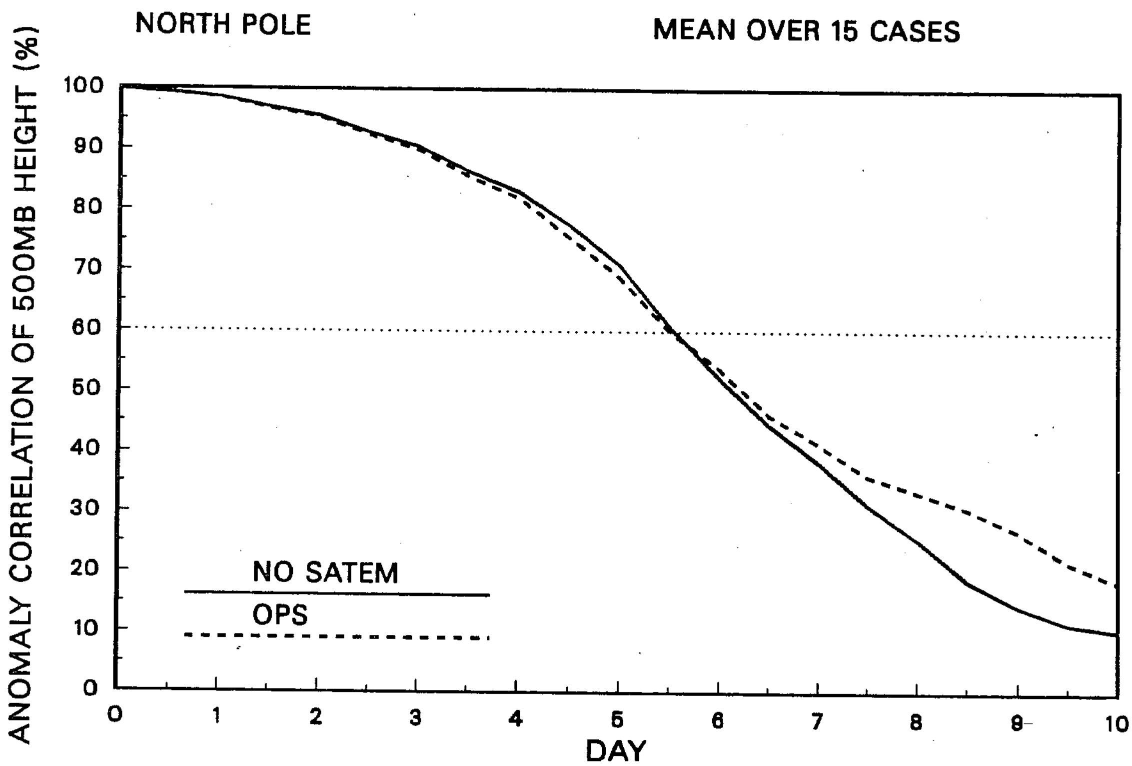


Fig. 17 Anomaly correlation of 500 hPa geopotential height, accumulated on 15 cases, comparing the 7 layer assimilation (full line) with the 11 layer assimilation (dashed line) on 3 different areas.



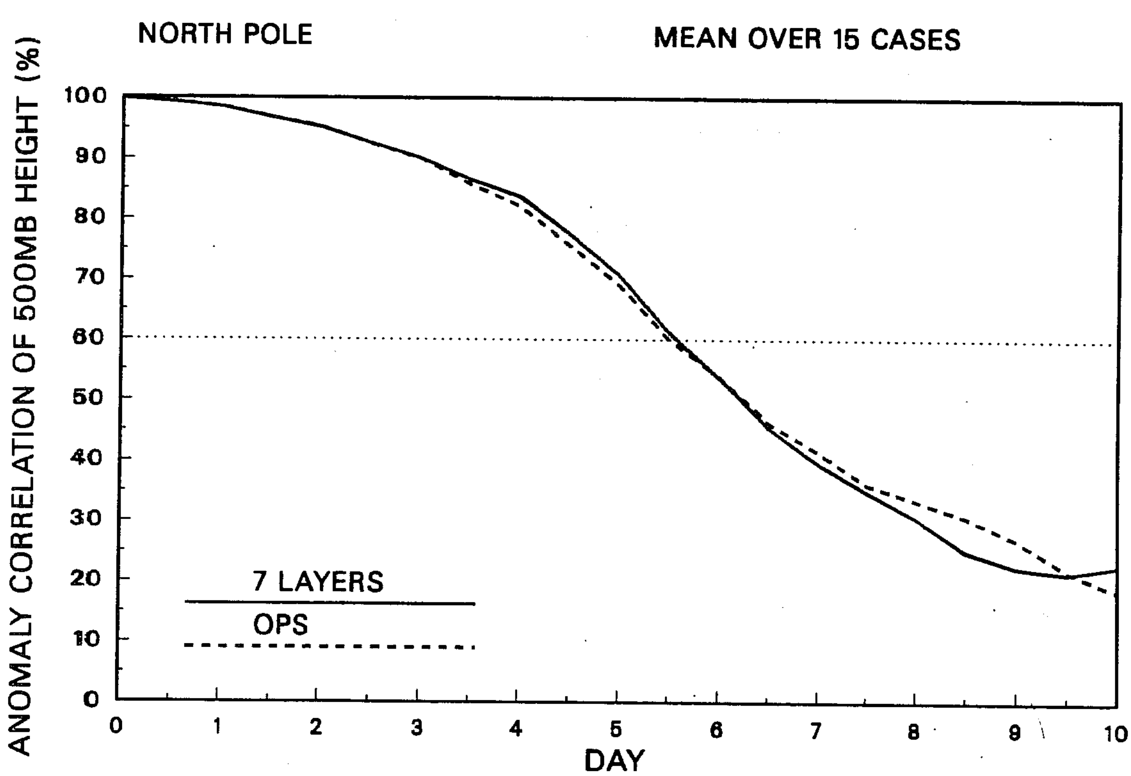


Fig. 18 Anomaly correlation of 500 hPa geopotential height, accumulated on 15 cases, on the North Pole (North of 70N).
NO SATEM vs 11 layers (OPS): top
7 layers vs 11 layers (OPS): bottom