N84 27295

THE IMPACT OF SATELLITE TEMPERATURE SOUNDINGS ON THE FORECASTS OF A SMALL NATIONAL METEOROLOGICAL SERVICE

Noah Wolfson[1], Albert Thomasell, Zvi Alperson[2], Harold Brodrick, Jy-Tai Chang[3], Arnold Gruber, and George Ohring
National Environmental Satellite, Data and Information Service, NOAA
Washington, D.C. 20233

ABSTRACT

The impact of introducing satellite temperature sounding data on a numerical weather prediction model of a small national weather service is evaluated. The Israel Meteorological Service model, a dry, 5-level, primitive equation model covering most of the Northern Hemisphere, is used for these experiments. Series of parallel forecast runs out to 48 hours are made with three different sets of initial conditions: 1) NOSAT runs, in which only conventional surface and upper air observations are used, 2) SAT runs, in which satellite soundings are added to the conventional data over oceanic regions and North Africa, and 3) ALLSAT runs, in which the conventional upper air observations are replaced by satellite soundings over the entire model domain. The impact on the forecasts is evaluated using three verification methods: 1) RMS errors in sea level pressure forecasts, 2) systematic errors in sea level pressure forecasts, and 3) errors in subjective forecasts of significant weather elements for a selected portion of the model domain. For the relatively short range ($\leq$48 hours) of the present forecasts, the major beneficial impacts on the sea level pressure forecasts are found precisely in those areas where the satellite sounding are inserted and where conventional upper air observations are sparse - i.e., the oceanic areas. Both the RMS and systematic errors are reduced in these regions. The subjective forecasts of significant weather elements are also improved with the use of the satellite data. Of special interest is the result that the ALLSAT forecasts are of a quality comparable to the SAR forecasts.

[1]National Research Council Research Associate

[2]Israel Meteorological Service, Bet-Dagan Israel

[3]Systems and Applied Sciences Corporation

ORIGINAL PAGE IS
OF POOR QUALITY

ABSTRACT

The impact of introducing satellite temperature sounding data on a numerical weather prediction model of a small national weather service is evaluated. The Israel Meteorological Service model, a dry, 5-level, primitive equation model covering most of the Northern Hemisphere, is used for these experiments. Series of parallel forecast runs out to 48 hours are made with three different sets of initial conditions: 1) NOSAT runs, in which only conventional surface and upper air observations are used, 2) SAT runs, in which satellite soundings are added to the conventional data over oceanic regions and North Africa, and 3) ALLSAT runs, in which the conventional upper air observations are replaced by satellite soundings over the entire model domain. The impact on the forecasts is evaluated using three verification methods: 1) RMS errors in sea level pressure forecasts, 2) systematic errors in sea level pressure forecasts, and 3) errors in subjective forecasts of significant weather elements for a selected portion of the model domain. For the relatively short range ($\leq$48 hrs) of the present forecasts, the major beneficial impacts on the sea level pressure forecasts are found precisely in those areas where the satellite soundings are inserted and where conventional upper air observations are sparse - i.e., the oceanic areas. Both the RMS and systematic errors are reduced in these regions. The subjective forecasts of significant weather elements are also improved with the use of the satellite data. Of special interest is the result that the ALLSAT forecasts are of a quality comparable to the SAT forecasts.

I. INTRODUCTION

The concept of radiometric sounding of atmospheric temperature profiles was first demonstrated with data gathered by infrared spectrometers on the Nimbus-3 satellite in 1969. Operational satellite soundings over oceanic areas were introduced by the VTPR (Vertical Temperature Profile Radiometer) instrument on the NOAA-2 satellite in 1972. Since 1979 (Smith et al 1979) the operational sounding system has been the TOVS (TIROS Operational Vertical Sounder) on the polar orbiting TIROS-N series of satellites. Early evaluations of the impact of satellite temperature soundings on numerical weather predictions showed only marginal improvements, except for the Southern Hemisphere, where, because of the sparsity of conventional observations, significant improvements were noted. More recent studies are somewhat more encouraging, Bengtssen et al., (1983) Halem et al., (1983).

Most previous evaluations of the impact of satellite soundings on weather forecasts have been based upon forecast models used by the larger meteorological services of the world. In the present study we examine the impact of satellite temperature soundings on the numerical predictions of a rather small national weather service - the Israel Meteorological Service (IMS). In addition to an evaluation of the impact on the forecast atmospheric flow field, we also evaluate the impact on subjective forecasts of actual surface weather elements. Preliminary results were reported by Thomasell et al., (1983) and Wolfson et al., (1983).

ORIGINAL PAGE IS
OF POOR QUALITY

2. ANALYSIS PROCEDURE

2.1 Experimental Design

The experiment was performed with the Israel Meteorological Service numerical weather prediction model, which is a dry, 5-level, primitive equation model. It operates at 5 sigma levels defined on an octagonal polar stereographic grid of 1844 points covering a large part of the northern hemisphere. The grid is aligned with 5°W longitude, is offset from the North Pole toward the Mediterranean Sea, and has a grid mesh of 360 km, true at 60°N. The integration domain is shown in figure 1. The model uses a 10-minute time step. The forecast variables are wind, absolute temperature and surface pressure. Geopotential is computed through temperature. Initial conditions, produced by a Cressman-type analysis (1959) scheme, comprise height and temperature at ten mandatory pressure levels from 1000 mb to 100 mb, plus surface pressure. Balanced winds are computed from the heights. For input to the model, the analyses are interpolated to the 5 sigma surfaces, and for display, the information is interpolated back to pressure coordinates. In these experiments, the analysis procedure used a 12-hour old analysis for a first guess. Conventional data for this model comprise radiosonde observations and standard surface and ship observations. Parallel forecast runs were made with the model with three different sets of upper air data providing initial conditions. For the NOSAT experiment the upper air data comprise conventional rawinsonde data, for the SAT experiment satellite sounding data are added, and for the ALLSAT experiment the conventional rawinsonde data are excluded.

For the SAT experiment satellite sounding data were introduced over the oceanic areas and over the North-Arican continent, where they supplemented sparse conventional data. For the ALLSAT experiment, satellite data were used over all of the area covered by the IMS model. All runs included sea level pressure as observed by the synoptic stations and ships. No aircraft or satellite-derived winds were used. As a result of this configuration of the operational system, the initial sea level pressure maps are identical for all experiments.

The model was run in the form indicated in Table 1 for up to 48 hours. The forecast fields were available for 12, 24, 36, and 48-hour projections. Data sets were available for three different time periods and the details for each experiment are given in Table 1. In each experiment the NOSAT forecast errors provided the reference against which SAT or ALLSAT forecast errors were compared.

ORIGINAL PAGE IS
OF POOR QUALITY

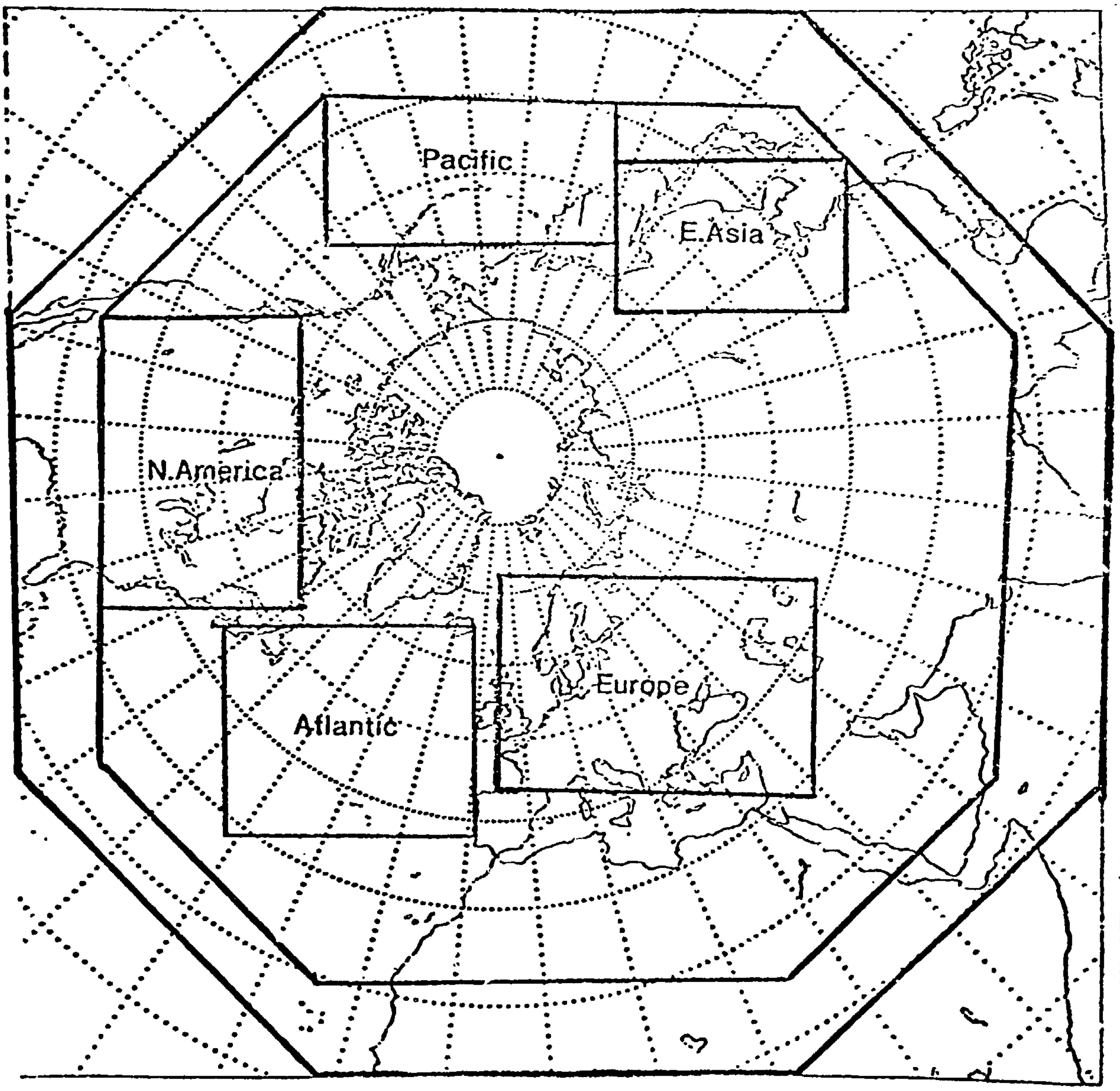

Fig. 1.

Areas covered by Israel Meteorological Service numerical weather prediction model (outer octagon) and areas (rectangles) used for verification.

ORIGINAL PAGE IS
OF POOR QUALITY

Table 1: Characteristics of the different experiments

Experiment No.	Period	Number of Satellites	Time Window for Satellite Data Insertion (HR)	Forecast Initial Time(GMT)	Experiment Type
1	1-12 Jan 1980	2	± 3	00, 12	SAT
2	15-23 Jan 1979	1	-3, + 6	12	SAT
3	20-25 Dec 1979	2	± 3	12	SAT
4	1-12 Jan 1980	2	±4.5	12	ALLSAT
5	20-25 Dec 1979	2	±4.5	12	ALLSAT

The time window (around analysis time) for the introduction of satellite data in the January 1980 and December 1979 data sets, when two operational satellites were available, was set to ±3 hours. This time range was considered to be a practical one for an operational system. It resulted in good data coverage for the western Atlantic for all the runs; however, the eastern Atlantic was only partially covered by satellite data at 12 GMT, and was not covered at all at 00GMT. Since, in regions of no observations the 12-hour old analysis defines the current analysis, a 12 hr discontinuity in the data was created along a line roughly north-south in the middle Atlantic. The time window for including satellite data in the January 1979 (experiment 2) data set, -3 to + 6 hours, was chosen to get enough data from the one available satellite to provide adequate coverage for the eastern Atlantic ocean. The time window for the ALLSAT runs (experiments 4 and 5) was set to ±4.5 hours. This resulted in nearly complete coverage of the model domain by the two-satellite system.

This paper will present and discuss results only of the impact on the sea level pressure forecasts. It is felt that, because of sparse radiosonde coverage over the oceans and the short range of the forecast, most of the impact should be found over the oceanic areas. At sea level, pressure is fairly well observed over the oceans; this allows the forecast fields to be verified with confidence there.

3. RESULTS

3.1 RMS analysis

The forecast sea level pressure fields were compared with their corresponding verifying analysis fields and RMS differences were calculated for each of the five verification regions for each of the five experiments, for forecast periods of 24 and 48 hours. These areas are presented in figure 1. The RMS differences for 48 hr are presented in Table 2, where a negative sign means that satellite data result in a beneficial impact. To help determine which numbers are significant, Student's t test was applied. We have adjusted the numbers of degrees of freedom to account for the statistical dependence of consecutive days.

Table 2: Average difference (SAT or ALLSAT -Minus NOSAT) RMS errors of 48 hour sea level pressure forecasts. Negative values indicate beneficial impact of satellite data. Statistically significant values at the 95% confidence level are underlined.

Exp.	Exp. period	Exp. type	E. Asia	Europe	N. America	Pacific O.	Atlantic O.
1	1-12 Jan 1980	SAT	-0.5	-0.1	0.1	-1.7	-1.1
2	15-23 Jan 1979	SAT	-1.4	-1.0	0.1	-2.7	-0.8
3	20-25 Dec 1979	SAT	-0.1	-0.6	-0.2	-1.9	-1.1
4	1-12 Jan 1980	ALLSAT	0.2	0.3	0.8	-1.2	-1.0
5	20-25 Dec 1979	ALLSAT	-2.5	-0.5	-0.2	-5.7	-1.0

3.1.1 Experiments 1, 2, and 3,-SAT experiments

Table 2 clearly shows that consistently beneficial impacts are found over the oceanic regions. For the Pacific Ocean the average reduction in RMS error is about 2 mb; for the Atlantic region the reduction is about 1 mb. The impacts over the continental verification areas are smaller but still beneficial, except for North America where there is no significant impact. Western North America is very close to the model boundary and accurate representation of synoptic systems and especially jet-streams from westerly and southerly directions is difficult. Of the first three experiments, only experiment 2 had satellite data in the eastern Atlantic; note the improved impact over Europe in this experiment. The results indicate that for these relatively short forecasts the largest impact is located over the oceanic regions, where the satellite data are inserted and where conventional observations are scarce.

3.1.2 Experiments 4 and 5 - ALLSAT experiments

Experiments 4 and 5 are all satellite data experiments. They examine the possibility of utilizing only satellite soundings for upper air observations, i.e. no radiosonde observations are used; however surface data are required to provide a reference level for the satellite soundings. The results are presented in Table 2.

ORIGINAL PAGE IS
OF POOR QUALITY

Table 2 shows that over the oceanic areas the ALLSAT shows lower RMS errors then NOSAT. The improvement is more apparent over the Pacific Ocean than over the Atlantic Ocean. The results over the continental areas are mixed. For January 1980 they show up to a 1 mb increase in the RMS versus NOSAT and SAT, but for December 1979 an improvement is observed over East-Asia Europe and North America.

3.2 Systematic errors

This section deals with the geographic distribution of the systematic error of the numerical forecast model used in these experiments, and the impact of satellite data on the systematic error. The mean or systematic forecast error at a gridpoint p is given by

$$\bar{E}_x(p) = \frac{1}{N}\sum_{n=1}^{N}\left(X_f(p,n) - X_a(p,n)\right) \qquad (1)$$

where X_f and X_a are the forecast and the verifying analyzed values, respectively, and N is the sample size or number of maps over which the average is calculated. It is easily seen that

$$E_x(p) = \bar{X}_f(p) - \bar{X}_a(p), \qquad (2)$$

and that a field of systematic errors may be computed as the difference between the average forecast field and the average verification analysis. The importance of analyzing systematic errors lies in the fact that the sources for this error can usually be identified and even treated.

The systematic errors and their differences were calculated for all five experiments. The Student t test at the 95% confidence level was used to determine significance. From an examination of the results from all our experiments (not shown), the following can be stated:

(a) Most of the impact is positive (i.e., beneficial)
(b) Impact was found in or very close to those areas where satellite data had been inserted.
(c) Most of the statistically significant impact areas are associated with the presence of well developed lows e.g., the Icelandic or Aleutian lows. The impact is usually found to the southeast of the centers of those average systems.
(d) The magnitide of the impact over the Atlantic Ocean is smaller than that over the Pacific.

The above results suggest that the statistically significant (95% confidence level) contribution of the satellite data in reducing the systematic error is realized in the area of active weather systems. This is physically reasonable, since we expect that the satellite data, although noisier than conventional rawinsonde observations (Gruber and Watkins, 1982) do have the capability of delineating the large horizontal temperature structure typically associated with active weather systems. Thus the signal to noise ratio is large enough to result in significant reduction of the systematic error.

It was also found that negative impact occurs in areas of time discontinuity in the data.

Such discontinuities can occur on the border lines:

1) between those areas where satellite data are inserted and those areas which are not updated by either satellite or conventional observations,
2) between neighboring swaths of observations taken by different satellites at different times.

4. SUBJECTIVE WEATHER FORECASTING

One disadvantage of the RMS measure of accuracy is that there is not a direct correspondence between reduction in RMS height or pressure errors and the quality of actual weather forecasts, which affect our everyday life. A more direct way to gauge the impact of the satellite data would be through the vehicle of subjective weather forecasts.

4.1 Subjective weather forecasting

To evaluate the impact of the satellite sounding data on the routine weather forecasts, subjective weather forecasts were issued for a network of stations presented in figure 2. The stations are located about 500 km apart from each other so that synoptic scale effects could be detected. The forecasted elements were: wind speed and direction, cloud amount and type, and precipitation. Each element was forecasted independently. The forecasts were given in categorias which are presented in Table 1 and were issued for 12, 24, and 48 hours.

Table 1: Forecast categories for weather elements.

Elements	category 1	category 2	category 3	category 4
wind speed (ff)	ff < 15 kt	ff ≥ 15 kt		
wind direction (dd)	0° < dd ≤ 90°	90° < dd ≤ 180°	180° < dd ≤ 270°	270° < dd ≤ 360°
cloud type	clear	low	middle	high
cloud amount (N, octas)	0 ≤ N ≤ 2	3 ≤ N ≤ 4	5 ≤ N ≤ 6	7 ≤ N ≤ 8
precipitation	yes	no		

4.2 Forecast procedure and verification

The procedure for issuing the subjective weather forecast was as follows. First the forecaster was presented with two sets of forecast maps as produced by the numerical prediction model with (SAT) and without (NOSAT) satellite data, but was not told which was which. He then issued two weather forecasts for each station using the two sets of numerical forecast maps. The number of maps available was limited because of practical reasons. The following maps have been used:

(1) Current operational sea level pressure as analyzed at the IMS.
(2) Forecasted sea level pressure and the 500-1000mb thickness.
(3) Current and forecasted 500mb height and vorticity.

ORIGINAL PAGE IS
OF POOR QUALITY

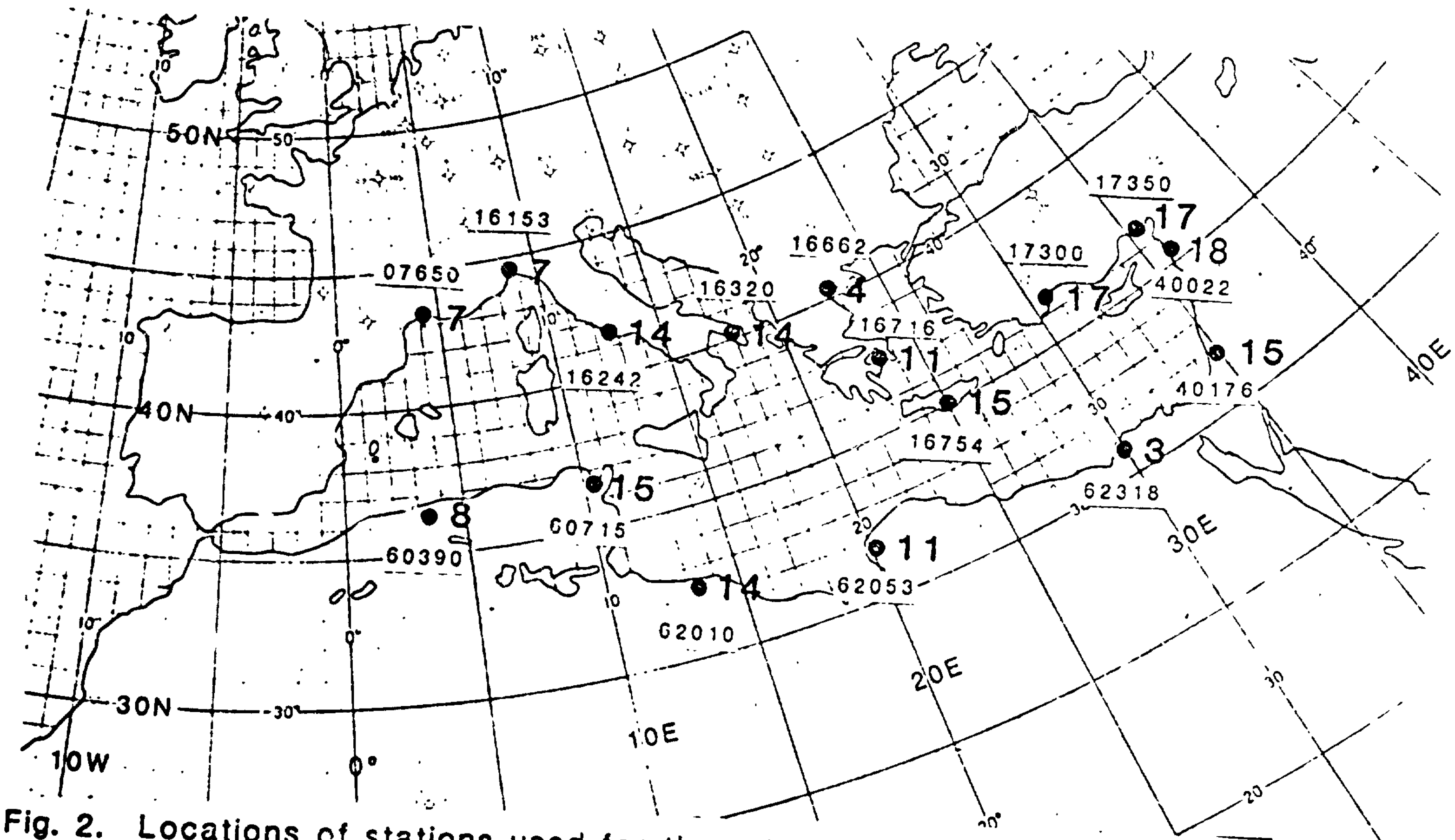

Fig. 2. Locations of stations used for the subjective weather forecasting experiment. Bold numbers indicate the percentage of the total forecasts for which there is a difference between the SAT and NOSAT predictions.

The Hansen-Kuippers (HK) Score which is the hit rate minus the false alarm rate (Mason 1982) was used for verification purposes. This score has two advantages:

(1) It discourages persistence or climatologically based forecasts by resulting in very low scores for this type of technique.
(2) The score is independent of the frequency of the forecasted events.

4.3 Results

Satellite impact on the subjective weather forecasts may be evaluated by addressing the following problems:

(1) What is the geographical distribution of the impact?
(2) At what forecast time was the impact realized?
(3) How was the quality of the forecasts changed?

4.3.1 Spatial distribution of the impact

The cumulative combined 12, 24, and 48 hours spatial distribution of the percentage of different station precipitation forecasts is presented in figure 2. One may observe that there are preferred areas where the number of different forecasts maximizes. The maximum difference area stretches from the Tyrranean Sea through Crete to the northeastern corner of the Mediterranean Sea. This area coincides roughly with one of the frequent cyclonic tracks in the Mediterranean Sea.

4.3.2 Forecast Quality

The impact of the satellite sounding data expressed in HK units, in which a higher score means a better forecast, was calculated using the impact for each station as an observation point. The Student t test was used to assess the statistical significance. The results are shown in Table 4.

Table 4: Average HK scores for precipitation forecasts with and without satellite data.

(significant results at the 95% confidence level are underlined)

	January 1980			January 1979		
Time (hrs)	12	24	48	12	24	48
HK with satellite	.55	.43	.28	.35	.44	.27
HK without satellite	.49	.35	.29	.31	.35	.30
Impact	.06	<u>.08</u>	-.01	.04	<u>.09</u>	-.03

Most of the impact is positive but not always statistically significant.

5. CONCLUSIONS

Our results indicate that a small national weather service using a modest numerical weather prediction model can improve its numerical predictions of sea level pressure and, through subjective interpretation of the numerical predictions, its forecasts of significant weathar elements by including satellite temperature soundings in the initial conditions for the model. However, we believe that the degree of improvement to be expected will depend on the details of the model's analysis and forocast system. For example, the Israeli forecast model uses a 12-hour old analysis as the first guess field for the current analysis. In sparse data regions, satellite soundings, despite their inherent errors, improve the analysis over that based upon the 12-hour old analysis and a few current radiosondes. Models using a 12-hour forecast for a first guess field may show different impacts.

Our results indicate that forecasts based upon an observing system consisting of surface observations and satellite soundings may be better than forecasts based on the conventional surface and upper air observing networks, and are comparable to forecasts based on a hybrid satellite sounding and conventional data system. This result has important implications for the planning of future observing systems and should be verified in experiments with more sophisticated numerical forecast models.

ACKNOWLEDGEMENT

This research was supported by the Office of Climate and Atmospheric Research, NOAA. The authors also thank Ms. R. Woodard for typing the manuscript, and Mr. R. Ryan for preparing the figures.

REFERENCES

Bengtssen, L.M., Kanemitsu, P. Kallberg and S. Uppala, 1982: FGGE research activities at ECMWF. Bulletin American Meteorological Society, Vol. 63, pp. 227-303.

Cressman, G.P., 1959: An operational objective analysis system. Monthly Weather Review, 87, pp. 367-374.

Gruber, A., and C.D. Watkins, 1982: Statistical Assessment of the Quality of TIROS-N and NOAA-6 Satellite Soundings. Monthly Weather Review, 110, 867-876.

Halem, M., E. Kalnay, W.E. Baker and R. Atlas, 1982: An assessment of the FGGE satellite observing system during SOP I. Bulletin American Meteorological Society, 63, pp. 407-426.

Mason, I., 1982: On scores for yes/no forecasts. Preprints Ninth Conference Weather Forecasting and Analysis. June 28-July 1, 1982. Seattle, Washington, 483 pp.

Smith, W.L., H.M. Woolf, C.M. Hayden, D.Q. Wark, and L.M. McMillin, 1979: The TIROS-N Operational Vertical Sounder. Bulletin of the American Meteorological Society, 60, pp 1177-1187.

Thomasell, A., Jy. T. Chang, N. Wolfson, A.Gruber and G. Ohring, 1983: The impact of satellite soundings and their distribution on forecast accuracy. Preprints Sixth Conference on Numerical Weather Prediction. 6-9 June 1983, Omaha, Nebraska, 421 pp.

Wolfson, N., A. Thomasell, G. Ohring, A. Gruber, H. Brodrick and Jy. T Chang, 1983: The impact of satellite sounding data on the systematic error of a NWP. Preprints Sixth Conference on Numerical Weather Prediction. 6-9 June 1983, Omaha, Nebraska, 421 pp.

C-5