INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

SMR/534-18

ICTP/WMO WORKSHOP ON EXTRA-TROPICAL AND TROPICAL
LIMITED AREA MODELLING
22 October - 3 November 1990

"General Problems of Data Assimilation: I"

P. UNDEN
ECMRWF
Reading, U.K.

OBSERVATIONS - AVAILABILITY AND QUALITY

By P.Unden, European Centre for Medium Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, England.

1. INTRODUCTION

The weather forecasting problem is largely an initial-value problem. In order to solve it by means of e.g. a numerical prediction model the initial state of the atmosphere has to be defined. This is done by analysing meteorological observations. There are many problems in defining the initial state which are due to the limitations in the geographical coverage and of the quality of the observations. This note will discuss these problems in relation to the task of numerical weather prediction (NWP). The material relies heavily on a lecture note prepared by Arpe (1985) which in turn makes use of notes by Bengtsson (1976).

2. REQUIREMENTS FOR ANALYSING METEOROLOGICAL FIELDS

2.1 Background

The global observing system has been built up gradually during the last 200 years but was mainly confined to surface observations until the second world war. These became synoptic ,i.e. carried out simultaneously a little over 100 years ago. The synoptic observations of pressure and wind enabled mapping and tracking of depressions, but not in real time until telecommunications became established. Synoptic meteorology developed based on air masses and frontal theories. The aerological network set up in the late thirties ,mainly for aeronautical purposes, enabled great advances in meteorology. For the first time the 3-dimensional flow could be studied in the northern hemisphere extra-tropics and analysis tecniques, diagnostics, research in the dynamics of the atmosphere and eventually numerical weather prediction methods were developed.

The obsevational network is very incoherent in its global coverage. There are large ocean areas almost void of manual observations and there are also large land areas which are mainly uninhabited. To set up and operate observation stations in such areas (or at sea) is very expensive and can only be achieved to a limited extent. Moreover is the above mentioned aerelogical network expensive to run due to manning and instrumentation costs. Some observing systems are also set up for more local use for short range forecasting. The highest density of observations is normally found in countries with well developed weather services.

There is also an increasing number of observations from moving platforms like buoys, ships, aircrafts and satellites. For the ships and aircrafts the geographical coverage is again limited due to the actual routes chosen.

This global network is coordinated and supported by the World Weather Watch of the World Meteorolical Organisation (WMO). One of their tasks is to try to increase the availability of observations where there is severe lack of them. It is then important to assess the requirements of observations for weather prediction and in particular for numerical weather prediction, which is the most important task on a regional or global scale.

The questions to be answered are basically as follows:

a) What are the most relevant variables to observe?

b) On which space and time scales do we need to sample?

c) What accuracy is needed for the measurements and transmissions?

These questions are also interrelated since some variables have large variations on smaller scales than others. Highly accurate measurents of say only one variable or only at some locations may be of little benefit for an analysis system.

The most basic considerations here are the space and time scales of the main meteorolical systems observed. In the extra-tropics these are rarely of wavelengths shorter than 1000 km and their periods are of the order of days. In order to map the synoptical systems one can conclude that the minimum sampling requirement would be 500 km in the horizontal and 12 hours in time. This would enable us to describe at least the large scale patterns of wind and pressure.

There are however significant weather phenomenea on smaller scales like fronts, cloud clusters and tropical cyclones. They would require a much higher sampling density. Most global prediction models, however, have effective resolutions not much higher than 200 km and it would not be desirable or even possible to try to introduce features of smaller scales than this.

In the vertical one can consider generalised baroclinic waves without trying to resolve fronts. These tilt with height and at least a 200 hPa resolution seems necessary to describe its broad features. This also agrees with experience from quasi-geostrophic models.

The relevant variables from observations are the ones which correspond to the model variables i.e. temperatures, geopotentials, pressures, wind components and humidities. The horizontal wind can be divided into a rotational and a divergent part and the rotational part is more accurately described than the divergent ones from the observations. The divergent part is however important for the spin-up in the model's physics. Humidity observations are also important for defining the precipitating processes but the available humidity observations are not capable of doing this very well. They represent scales much smaller than the typical model scale of 200 km. The model will in due course generate its own consistent humidity field after a couple of days, with or without humidity observations. Some surface parameters like snow depth and sea surface temperatures are also used but their impact on medium range weather forecasts is small.

Fig. 1a-1f shows the global data coverage for one 6 hour period as received at ECMWF. It contains the coverage of satellite wind data (SATEM 1000-2000 observations during a normal 6-hour period), aircraft reports (AIREP 400-800), drifting buoys (DRIBU 200-400), synoptical surface and ship observations (SYNOP and SHIP 7000-8500), satellite temperatures (SATEM and TOVS 4000-8500), balloon wind measurements (PILOT 150-300) and radiosondes (TEMP 600-680 at 00 or 12 UT).

2.3 Required accuracy

The ECMWF data assimilation system can cope with any accuracy of observational data as long as their accuracy is known and as long as the data have no systematic bias. In principle a large number of randomly inaccurate data can be used to obtain analyses with a higher accuracy than any of the data involved. Sample averages of observed data must, however, be more accurate than the first guesses used for the analysis. Observational data, which have a systematic bias have the most detrimental effect on the quality of the analyses. Also isolated inaccurate observed data are difficult to use.

The present operational observational network is very expensive to maintain and it would become even more expensive to increase the accuracies of the measurements. Attempts have been made to estimate the effect of the accuracy of initial data on forecast quality. The observational data errors are not the same as analysis errors, but the analysis errors will be highly influenced by the accuracy of observed data and therefore we can use such investigations to see if there is a limit of accuracy beyond which it would not be worth going. Such a study was carried out by Lorenz (1969) who investigated the error growth of a barotropic model.

This model has an energy spectrum as indicated by the heavy curve in Fig. 2. After disturbing the initial values in the smallest scales the rerun forecasts take a while before the larger scales are influenced by these disturbances due to non-linear interactions. The thin curves in Fig. 2 indicate how much of the spectrum is affected by the initial disturbances after 15 minutes, 1 hour, 5 hours, 1 day and 5 days. The error energy is seen to double very quickly while it is

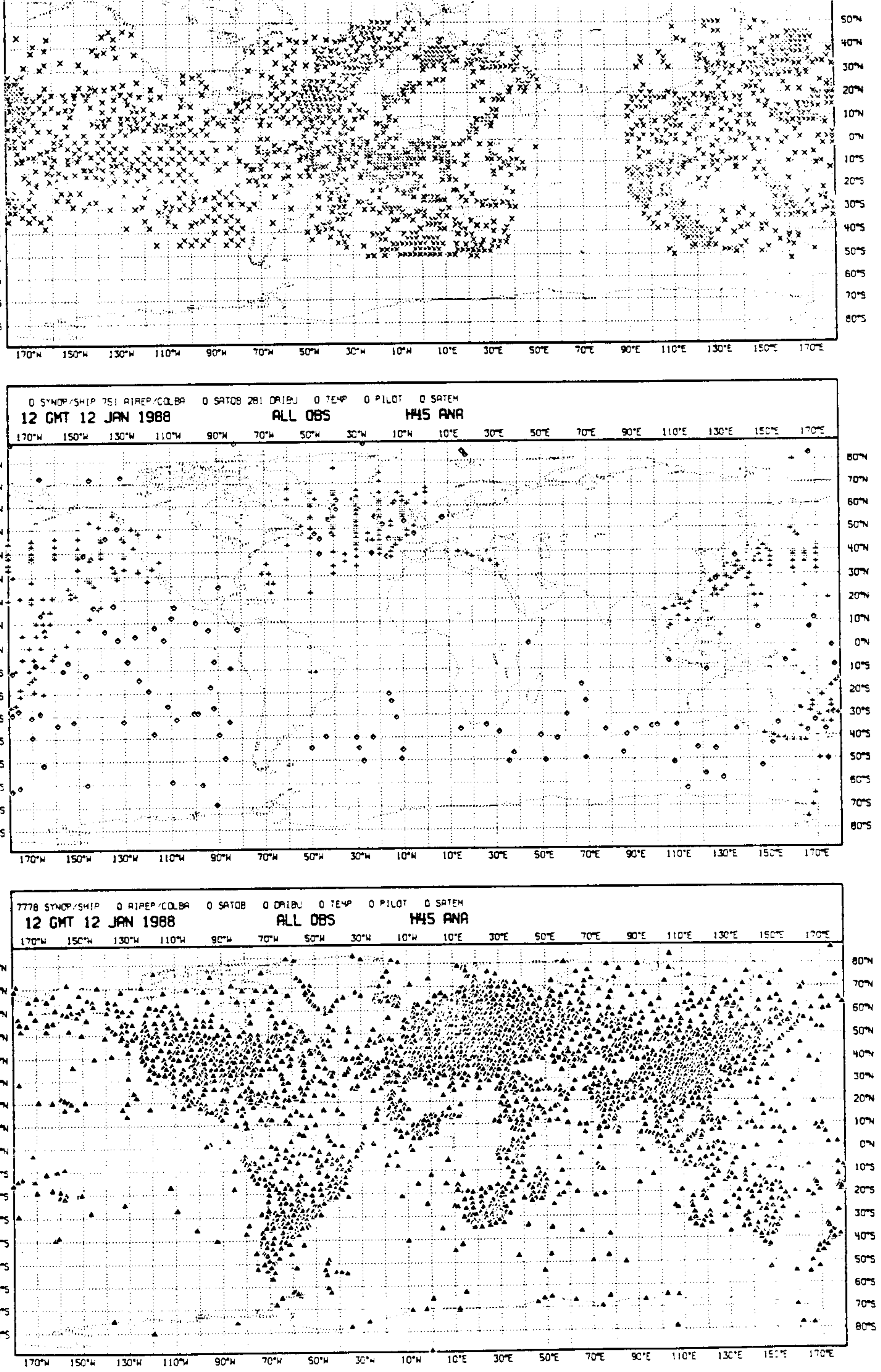
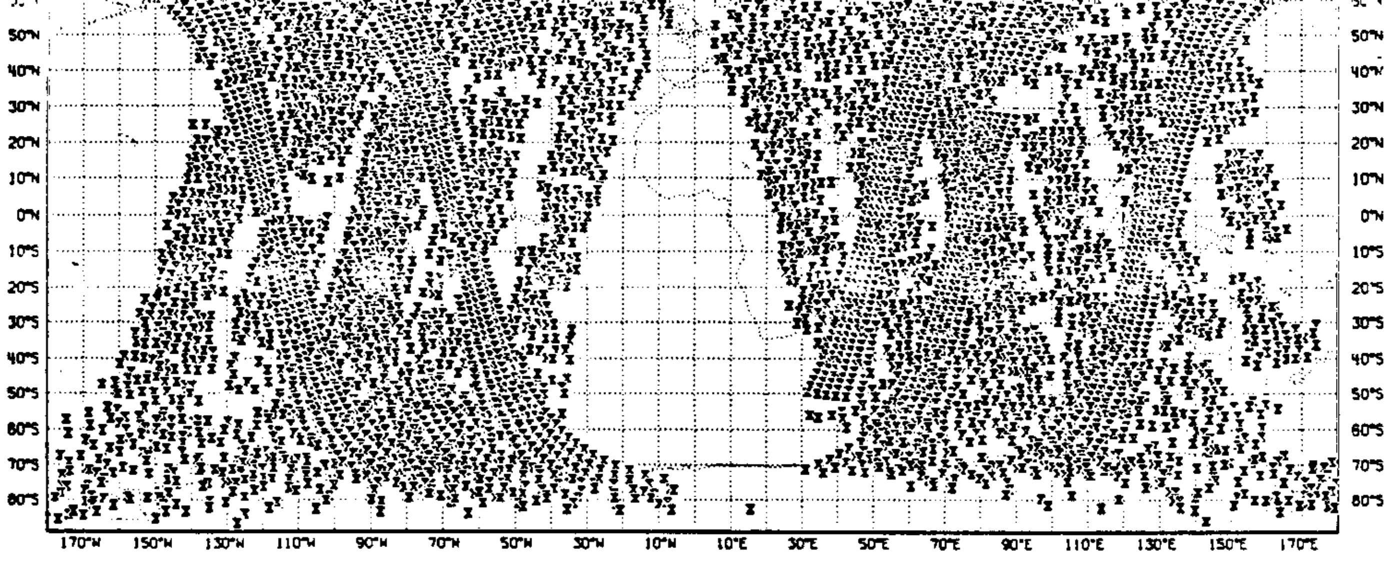
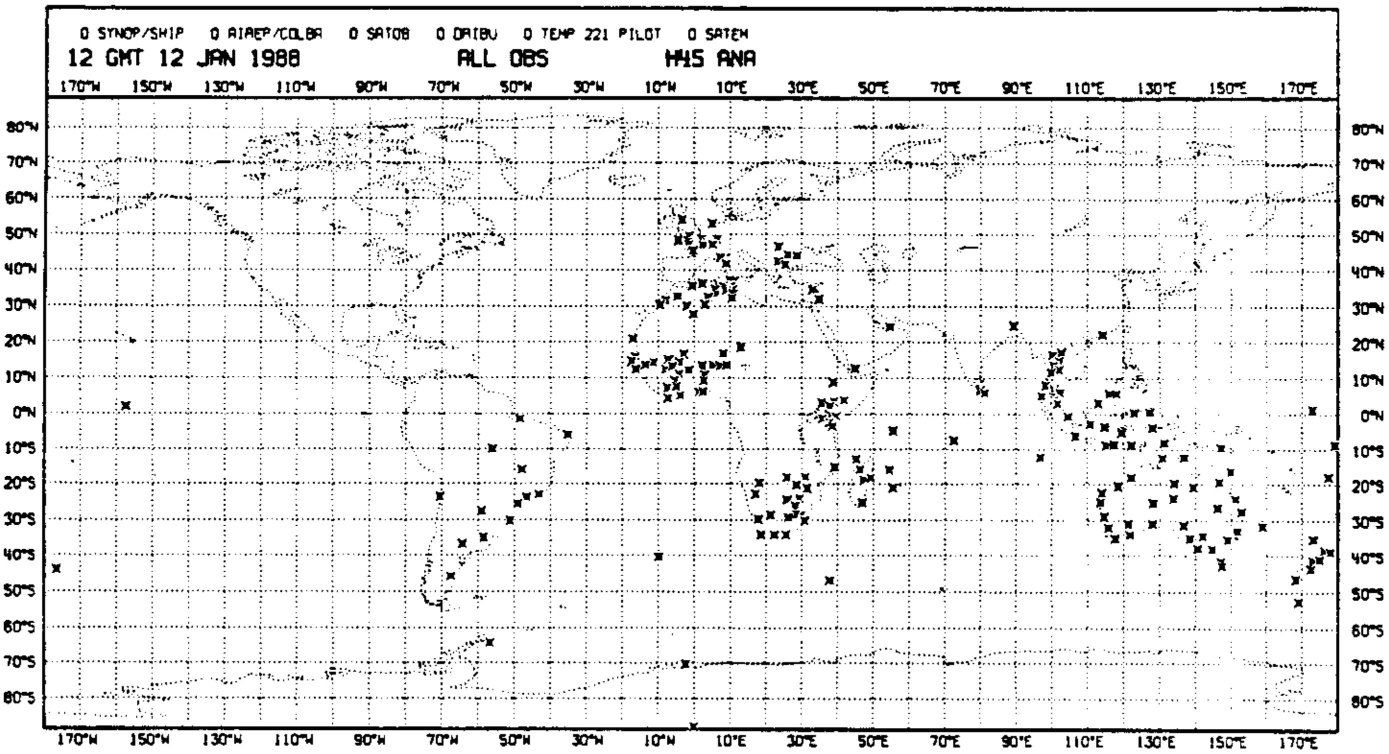
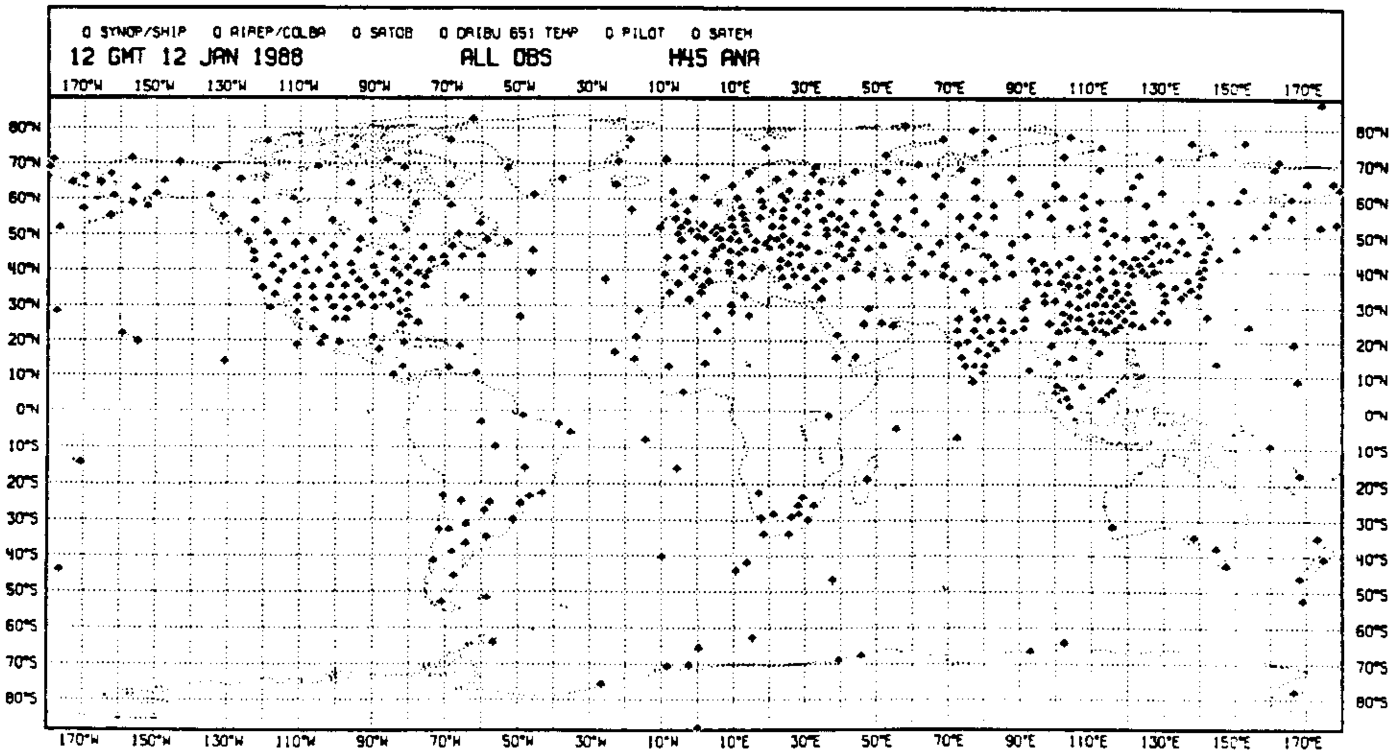


Figure 1. Data covarage at ECMWF 1988-01-12 0901-1500 for a) synoptical surface reports (SYNOP and SHIP) b) aircraft reports (AIREP) and drifting buoys (DRIBU) c) satellite winds(SATOB)







d) radiosondes (TEMP and ASAP) e) profiles of wind (PILOT) f) satellite temperatures and thickresses (SATEM and confined to the smallest scales.

The same model was run with different initial perturbations to estimate the gain of predictive skill in the forecasts by halving the initial error. Table 1 gives the range of predictability as function of wavelength and initial error. Even with very large initial errors of 9 m/s, the synoptic scale circulation retains some predictive skill for one day, while some planetary scales retain some predictability for a week. If the initial error is halved, the range of predictability is nearly doubled for a large range of wavelenghts. This is also nearly true when halving the initial error another time, but it is obvious that further reduction of initial errors give less and less gain in predictive scill. For the synoptic scale of motion (L >1250 km) a reduction of the error below 1 m/s seems unnecessary.

If these results can be applied to the real atmosphere, there is no need to observe (or to have initial analyses) with higher accuracies than 1m/s as far as the prediction of synoptic planetary scales is concerned. Using the geostrophic aproximation we can estimate from this an optimal accuracy for the height field. More realistic experiments are, however, needed to confirm these results although they agree in general terms with what has been observed for e.g. the influence of one bad piece of data. Uncertainties in the analyses are not restricted to small scales as in the experiments discussed above and therefore in realistic experiments one would also find the error growth in long wawes due to analyses errors from the beginning of the forecasts. Also the the errors of the forecast model itself play an important role in the growth of forecast errors.

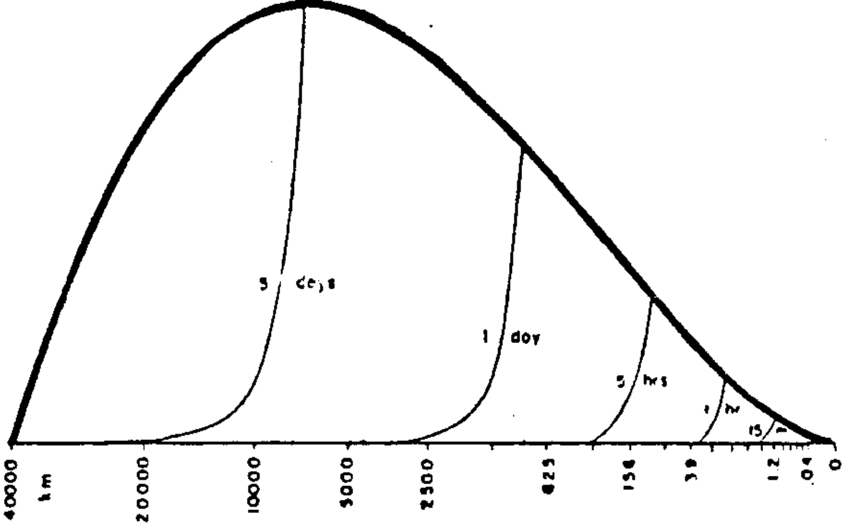


Fig. 2: Basic energy spectrum (heavy curve), and error-energy spectra (thin curves) at 15 minutes, 1 hour, 5 hours, 1 day, and 5 days, as derived from a numerical experiment. Thin curves coincide with heavy curve to the right of their intersections with heavy curve. The horizontal coordinate is the fourth root of the wavelength, labelled according to wavelength. Resolution intervals are separated by vertical marks at base of diagram. Vertical coordinate is energy per unit logarithm of wavelength, divided by the fourth root of wavelength. Areas are proportional to energy (from Lorenz, 1969)

Table 1: Predictability as a function of wavelength and initial error (d = days, h = hours) (after Lorenz, 1969).

wave length (km)	Initial Error (m/sec)					
	9	4.5	2.2	1.1	0.6	0.3
20000	5.7d	8.0d	9.2d	9.8d	10.0d	10.1d
10000	2.5	4.0	4.8	5.3	5.5	5.6
5000	1.1	2.0	2.6	2.9	3.1	3.1
2800	12.8h	23.8h	1.4	1.6	1.8	1.9
1250	6.1	11.9	17.8h	22.3h	1.0	1.1
625	3.0	6.0	9.6	12.7	14.4h	15.2h
312	1.5	3.1	5.2	7.2	8.5	9.1
156	0.7	1.5	2.8	4.1	5.0	5.5

3. HOW OPERATIONAL DATA MEET THESE REQUIREMENTS

3.1 Distribution of observational data

In previous section Fig. 1 shows the typical data coverage for a 6-hour period centred around a main hour (00 or 12 UT). As mentioned before the most important source of 3-dimensional data is from radiosondes. If they were distributed evenly with the 500 km resolution mentioned earlier they would describe the planetary synoptic pattern in the horizontal and vertical reasonably well. Fig. 3 shows schematically the density of observational data in a zonal mean required to have this resolution together with the available radiosondes at ECMWF for the same latitudinal bands during April 1984. It is obvious that south of 30 N the density of radiosonde stations is far too low.

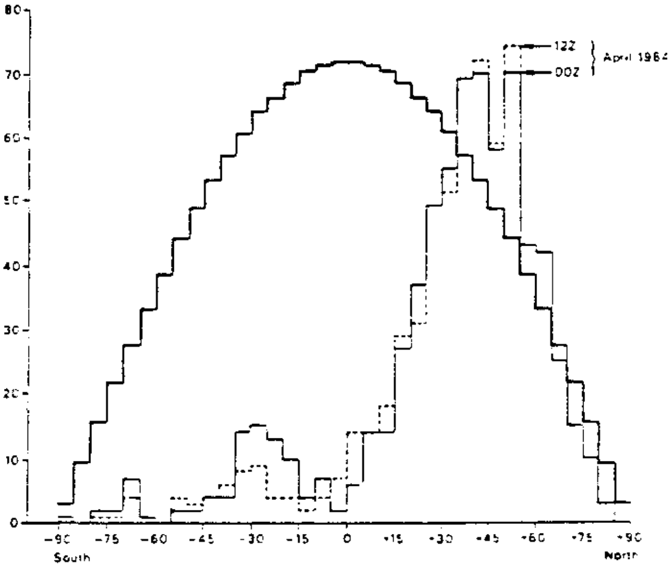


Fig. 3: Number of radiosondes in zonal belts of 5°. Thick solid line: requirement to meet 500 km distance between stations. Thin solid line: average number of radiosondes received during April 1984 at ECMWF for 00 GMT. Dashed line: same as above for 12 GMT.

Fig. 4 exhibits the geographical distribution of radiosonde data; clearly there appears to be a deficiency also in northern hemisphere mid-latitudes over oceans. There are three quasi-stationary weather ships in the Atlantic and there are also some ships of oppurtunity in both oceans reporting fairly regularly along their routes. These ships are definately too few to fulfil our requirement of a 500 km spacing. This picture has not changed much in recent times except that a few more automated ship soundings (ASAP) have become available.

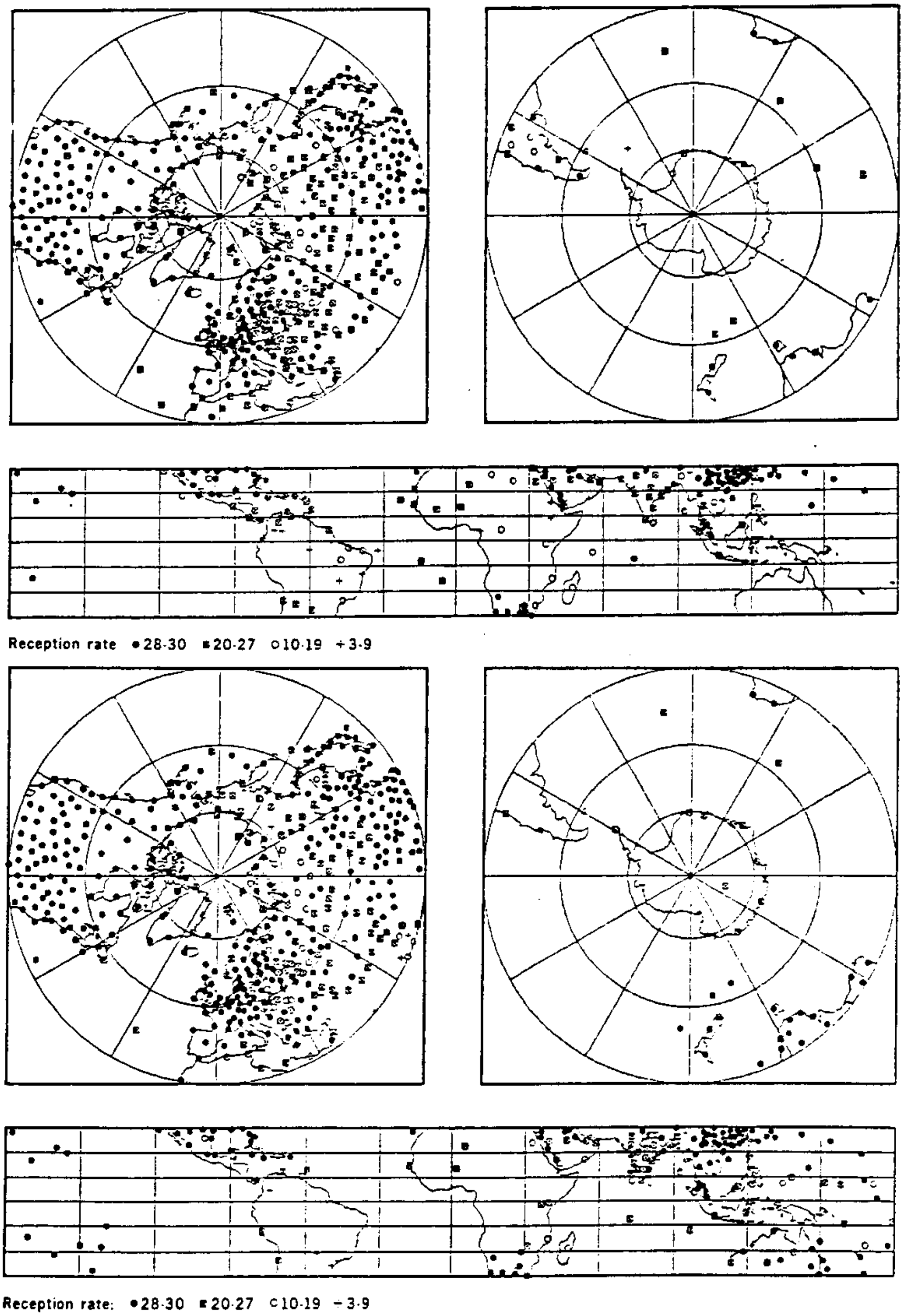


Fig. 4: Distribution and reception rate of radiosonde ascents from land stations at ECMWF during April 1984. Upper panels 12 GMT; lower panel 00 GMT.

The TEMP observations are augmented by PILOT data in the tropics and Australia. These contain only winds and many only in the lower part of the troposphere if made with theodolite. Others are made with RADAR or radio track finding equipment from a radiosonde station and these ones contain the full verical wind profile. In the tropics the wind field is however the most relevant parameter to observe.

3.2 Accuracy and representativeness of meteorological data

The barometer is relatively the most accurate meteorological instrument. The instrument error is as low as 0.1 hPa which should be compared with an atmospheric variability of around 100 hPa. There are however other effects which reduce the validity of the appearent accuracy.

The air flow around a building or ship may create dynamical pressure changes of the order of 1 hPa and is difficult to compensate for. Then the atmosphere has tidal waves dominated by the semi-diurnal tide and they appear as a pressure oscillation with amplitude up to 2 hPa in the tropics. The tides are difficult to represent very well in forecast models and since the phase of the tides follows local time a 6-hour forecast may be quite out of phase. Since the tides constitute the dominant part of the variance of surface pressure in the tropics it may be quite difficult to extract the "meteorological" signal.

Another problem is the use of high elevation pressure observations. These are usually reduced to sea-level or nearest standard pressure level using the hydrostatic formula. This entails some assumptions about the temperature below (or sometimes above) the station. Apart from often being unrealistic such practices vary from region to region. State-of-the-art data assimilation system prefer to use the station level pressure to avoid excessive extrapolation. Still some may be necessary since the model's orography usually doesn't coincide with the one of the station. Also in some cases the real altitude of the station may be incorrect.

Temperature can in principle be measured down to an accuracy of 0.2-0.1 K at screen height (2 m above ground). The turbulent fluctuation during a day with convection may however be of the order of 1 K and local variations during a cloudless night may be much larger. The representativeness is in other words much less for temperature than pressure under such conditions. In the free atmosphere the temperature is measured with a calibrated thermistor in a light disposable instrument (radiosonde). The can nominally have an accuracy down to 0.2 K in ideal condition. In flight they are affected by the efficacy of the airflow, reaction time of the instrument and, most importantly, radiation errors. These are due to solar heating during the day and radiational cooling to space during night. The effect is mostly pronounced in the stratosphere and may cause discrepancies of up to 3 K between different instrument types.

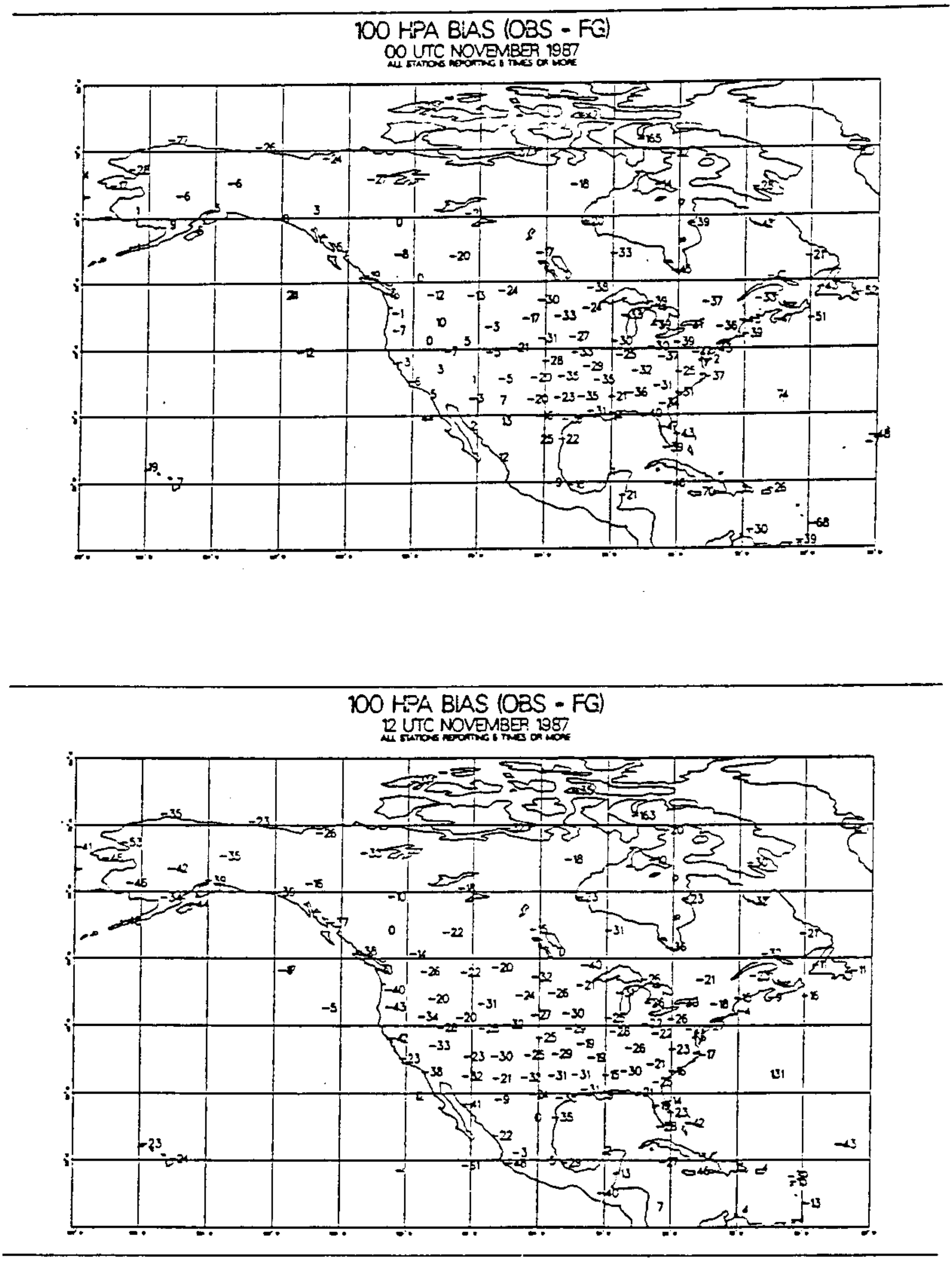
Wind measurements are mainly affected by the turbulent motions which cause rapid fluctuations in the wind speed. It is a problem of measuring on the right time scale which at the surface may be achieved by time averaging. In the free atmosphere this is not possible since the balloon rises with a speed of about 300 m/minute and is thus affected by local turbulence. Time averaging here means vertical averaging and this is done when tracking the path of the balloon. Other errors are introduced by the wind tracking equipment itself and this increases with very low elevation angles (=long distance and strong wind) with ground-based systems. Comparisons between different systems indicate differences of 2-4 m/s at various levels.

Geopotential (or sometimes referred to as just height) observations are computed by integrating the hydrostatic equation and are thus affected by the errors in temperature. A systematic bias in the temperature has an accumalative effect on the geopotential and this can be quite serious at high levels. Geopotential biases for individual stations are therefore common and the conceived observation error increase with height. Examples of such biases are shown for North America in Fig. 5 (from Radford (1987)). At 00 UT the Pacific area is in daylight and much less bias against the first guess is observed than over the eastern states. At 12 UT a more uniform bias is evident since the whole continental area is in darknes. These pictures indicate a significant day-night problem with the geopotentials.

The polar orbiting satellites are the only potentially global observing systems and have a useful role to bridge the gaps where conventional data do not exist. They are however contaminated by clouds and this leads to less accurate estimations in cloudy areas. Also ice, snow and certain ground properties present problems and this renders the lowest levels in the retrieved vertical profiles in such areas almost useless. The retrieval of temperatures is done by inversion of the radiative transfer equation and this is not a uniquely detrmined problem, i.e. there are several temperature profiles that would give rise to the same radiation. Statistical methods are employed to relate radiation to radiosonde measurements and to keep the errors within reasonable limits most of the time. Comparisons between co-located radiosondes and satellite temperatures show RMS-differences of 2-3 K.

Measurements from geostationary satellites consist of wind retrievals and for some satellites humidity and temperature profiles. The wind data is the most commonly used product. The winds are produced by tracking cloud patterns and this suffers from some problems. Clouds do evolve and do not necessarily follow the actual wind (c.f. orographic clouds which are avoided in this context). Also the cloud top heights are not well known. Overall this produces errors of about of 5-7 m/s in the RMS sense.

Table 2 displays the conceived observation errors as they are used in the ECMWF analysis scheme (1988). Note however that these errors also include the representativeness error with respect to the effective model resolution of about 200 km in the horizontal (and a variable one in the vertical as well of the order of 1 km increasing for higher levels). This error is significant and may be of the same order as the actual accuracy of the observation.



100 HPA BIAS (OBS - FG)
00 UTC NOVEMBER 1987
ALL STATIONS REPORTING 5 TIMES OR MORE

100 HPA BIAS (OBS - FG)
12 UTC NOVEMBER 1987
ALL STATIONS REPORTING 5 TIMES OR MORE

Fig. 5 Mean observed minus first-guess difference at 100 hPa for November 1987 (metres), 00 UTC data (above) and 12 UTC data (below)

Inst. Type	1000	850	700	500	400	300	250	200	150	100	70	50	30	20	10
SONDE/PILOT WIND	2.2	2.5	2.6	3.1	3.7	3.8	3.3	3.0	2.8	2.4	2.4	2.4	2.5	3.1	3.5
SONDE GEOP.	5.0	5.4	6.0	9.4	11.6	13.8	14.2	15.2	18.2	21.4	25.2	29.8	31.2	38.1	50.0
SATEMS															
Clear path	-	11.8	11.0	14.6	11.8	14.2	-	22.5	-	23.8	16.8	17.3	24.5	-	52.4
Partly cloudy	-	13.1	12.2	16.2	13.1	15.8	-	25.0	-	26.4	18.7	19.2	27.2	-	58.2
Microwave	-	14.4	13.4	17.9	14.4	17.4	-	27.5	-	29.0	20.6	21.2	29.9	-	64.0
SATOBS															
GOES															
METEOSAT	2.5	2.5	2.5	2.5	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
HIMAWARI															
AIDS/ASDAR	3.5	3.5	3.5	3.5	3.5	4.0	4.5	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
AIREP															
PAOBS	32.0														
SYNOP/SHIP WIND	3.6														
SYNOP/SHIP GEOP.	7.0														
DRIBU GEOP.	14.0														

Table 2 The rms observation errors (units ms^{-1} and m). The SATEM thickness errors are given for layers with pressure at the top of the layer given in the first row.

4. QUALITY CONTROL

A vital component of any automated system is quality control. Observations received at a meteorological centre are sometimes severely in error and if used directly they can produce absurd results. The errors may be due to instument faults, reading errors, calibration errors, coding errors or corruption in the transmission process. The methods available today for detecting these errors are the following:

a) Internal consistency. Some observation types contain related parameters like temperature and heights or temperatures and freezing precipitation.

b) Climatological limits. Observed values should not be allowed to exceed the known climatological extremes for a particular level in the atmosphere (and season and region).

c) First guess check. Use of a short range numerical forecast is a very powerful tool for discovering incorrect observations. Statistics of the distributions of observed-first guess values can be accumulated and probabilities of the event of a certain departure can be calculated. Some observations fall in the tail of the distribution (see Fig. 6) and the probability of them being correct is very small.

d) Independent analysis. A more expensive and slightly more subtle way of checking is to analyse at observed variable at its position using surrounding observations but not the observation to be checked. The estimate is then compared with the observed value and if it exceeds certain values it can be cosidered to be probably incorrect and rejected.

e) Time continuity. A history can be kept of observed values and quality flags for individual observation platforms. If the values suddenly change more than reasonable changes per time unit or if the quality flags consistently indicate a problem, then the observation is suspect and can be rejected.

f) Data monitoring. In practice the previous method is expensive in terms of data storage and computing and only works well for certain observation types and areas. It can then be augmented by manual or semi-manual day to day monitoring of observation departures for suspect observations, certain platforms and gathering of say monthly statistics. Then manual decisions have to be taken based on this material whether to exclude certain observations permanently (until they improve). Figure 7 shows an example of one radiosonde station which had a very distinctive stratospheric bias problem in August for both 00 and 12 UT.

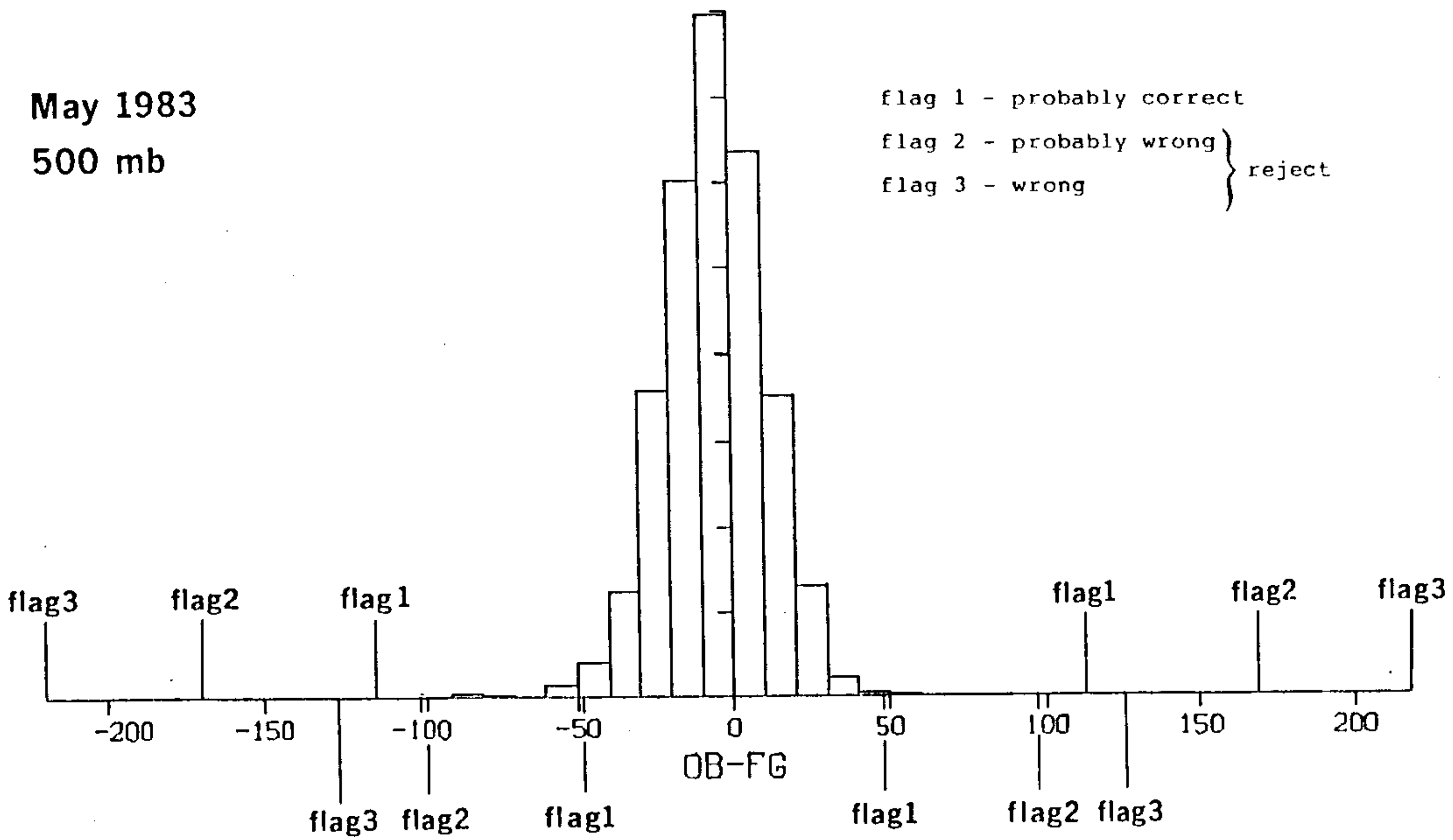


Fig. 6 Histogram of (observed - first guess) 500 mb height values for N. American TEMP reports for May 1983 (12 GMT datum time). Units : metres. The markings on the vertical scale are in units of 100 reports. The old flags are shown above the abscissa and the new ones below.

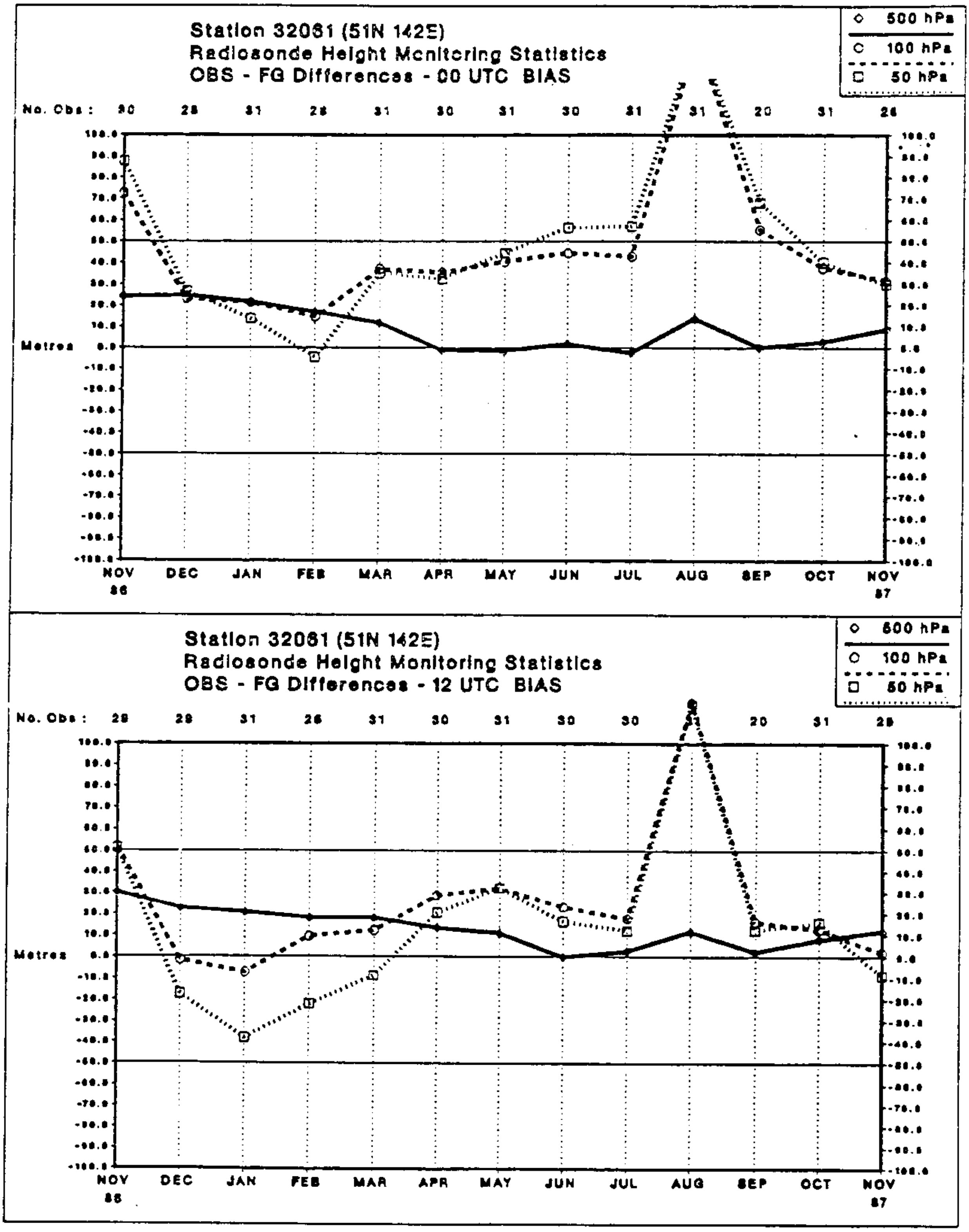


Figure 7. Time graphs of mean monthly differences between observations and first guess of geopotential height at station 32061 in meters, 00 UT data above and 12 UT below. From Radford (1987).

REFERENCES

Arpe,K. 1985. Meteorological data: Kind, distribution, accuracy and representativeness. ECMWF Lecture Note No. 2.1, 53pp, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, England.

Bengtsson,L., 1976. Initial data and some practical aspects of weather forecasting. Weather Forecasting and Weather Forecasts: Models, Systems and Users - Vol.1. NCAR/CQ-5 1976 - ASP.

Lorenz,E.N. 1969. The predictability of a flow which posesses many scales of motion. Tellus, 21, 289-307.

Radford,A.M. 1987. ECMWF radiosonde monitioring results. ECMWF/WMO Workshop on Radiosonde Data Quality and Monitoring, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, England.