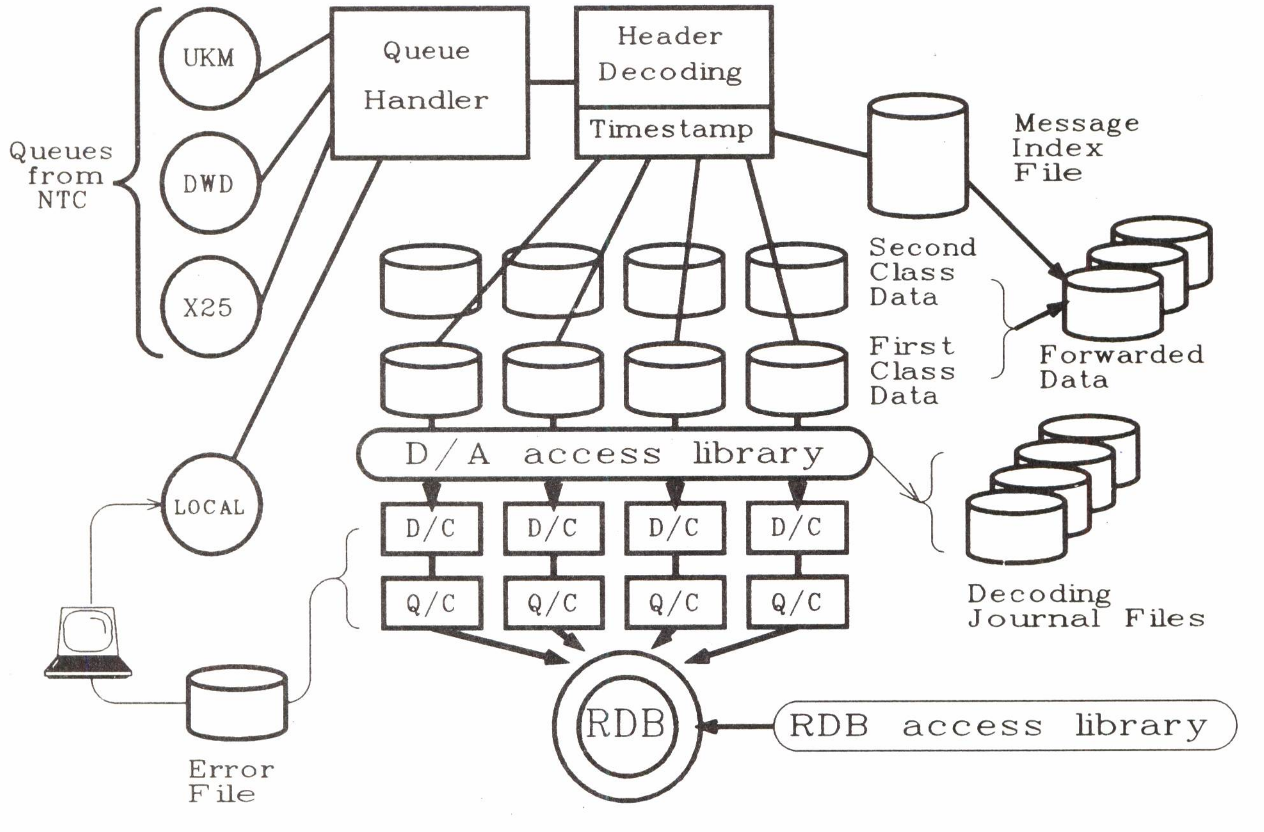
European Centre for Medium Range Weather Forecasts

ECMWF NEWSLETTER

Shinfield Park, Reading, Berkshire RG2 9AX, England. Tel: U.K. (0734) 876000, Int. (44 734) 876000, Telex: 847908

Number 46 - June 1989

The New Data Acquisition and Preprocessing Systems



IN THIS ISSUE

Page

METEOROLOGICAL

Changes to the operational forecasting system 2

Revisions to the model physics 3

Operational impact of the new model physics 11

Kalman filtering - a new approach to adaptive statistical interpretation of numerical meteorological forecasts 16

COMPUTING

The new data acquisition system 21

COMPUTER USER INFORMATION

Current Cray service 26

Still valid news sheets 28

GENERAL

Second workshop on meteorological operational systems 29

ECMWF calendar 30

ECMWF publications 30

Index of still valid Newsletter articles 31

COVER: The new data acquisition system. Please see page 21 for a full description.

This Newsletter is edited and produced by User Support.

The next issue will appear in September 1989.

The first two articles (page 2 and page 3) of this edition of the Newsletter give a detailed description of changes which were introduced recently into the operational model and an initial evaluation of their effects upon the forecast results.

The Computing Section (page 21) contains a description of the new data acquisition system based on a VAX 6210, which was to have appeared in the last issue of the Newsletter.

Users of the ECMWF computing facilities are reminded that the NOS/BE operating system will be withdrawn from service on 31 December this year. Information on replacement facilities to which users should transfer was given in ECMWF Newsletter No. 44 (December 1988).

It is now possible to dial through directly to most members of staff at ECMWF. Details of the new numbering system are given on the back page of this Newsletter.

* * * * * * * * * * * *

CHANGES TO THE OPERATIONAL FORECASTING SYSTEM

Recent changes

Articles describing the changes to the forecast model introduced on 2 May 1989 are to be found on pages 3 and 11.

Planned changes

The surface analysis code will be replaced by a new one. It is mainly a technical development, the analysis of surface variables is now performed inside the context of the main analysis program (rewritten in September 1986) rather than in a separate step. Little meteorological impact is expected on the SST and snow analysis (the only surface variables currently analysed). However, the SST computations will use as input the 2 degree mesh SST analysis from the NMC (instead of the 5 degree mesh). Thus more details are expected on the initial SST field used to run the ECMWF operational model.

- Bernard Strauss

* * * * * * * * * * * * *

REVISIONS TO THE MODEL PHYSICS
(Implemented on 2 May 1989)

Introduction

A set of three important modifications to the model's physics was implemented in the operational forecasting system on 2 May 1989:

a) A new parametrization scheme for radiative fluxes and the representation of cloud optical properties.

b) A reformulation of cumulus parametrization using the mass flux approach.

c) A revision of the gravity wave drag formulation.

These changes were the outcome of an extensive research and experimentation programme followed by a series of 13 parallel assimilations and 10-day forecasts during the period 19 April - 1 May 1989. This article describes the model changes and their main physical impact as deduced from the experimentation programme and confirmed during the parallel runs. A second article covers the effect of those changes on the operational performance of the model.

Description of the model changes

Radiation

After a thorough validation of the previous operational scheme and other available schemes against detailed line-by-line, narrow-band models and against satellite data, it was decided to adapt a code developed earlier at the University of Lille to the ECMWF model. The main differences from the old scheme in the new scheme are:

(i) smaller shortwave H_2O absorptivity, which reduces the clear-sky shortwave heating and increases the downward solar radiation at the surface;

(ii) a correct temperature and pressure dependence of the longwave absorption, which increases the longwave cooling in mid-troposphere and the stratosphere;

(iii) the presence of water vapour continuum absorption, which cools the tropical boundary layer;

(iv) cloud optical properties derived from a more realistic model cloud, and a diagnostic formulation of the cloud liquid water content independent of the model's vertical grid. These features both contribute to more radiatively active clouds and thus to the better representation of the radiation fields at the top of the atmosphere.

Convection

The Kuo scheme for cumulus convection which has been used in the ECMWF model up to now is replaced by a "mass flux" scheme. In this approach, subgrid vertical fluxes of mass, heat, water vapour and momentum are computed at each model level with the help of a simple cloud model interacting with its environment. The mass flux concept is supported by theoretical as well as observational studies and offers good prospects for further developments. The new scheme is applied to penetrative convection, shallow convection and mid-level convection and considers the effects of cumulus updrafts, saturated downdrafts and cumulus-induced subsidence in the environmental air. It also considers vertical transports of momentum by convective scale circulations. Updrafts and downdrafts are modelled as one-dimensional entraining plumes with simple cloud physics. The scheme is based on a moisture convergence hypothesis and is thus comparable to the previous Kuo scheme; however, it differs by these additional features:

(i) heat and moisture transports by cumulus-scale circulations, including downdrafts: this accounts for most of the differences concerning the vertical profile of heating and the spin-up of the hydrological cycle.

(ii) Momentum transport: this produces down-gradient momentum fluxes in the tropics and thus acts to decelerate the large-scale zonal flow in the upper troposphere.

(iii) Mid-level convection: this stabilizes the air above the boundary layer in the presence of conditional instability and large-scale ascent (extra-tropical fronts).

(iv) Entrainment of environmental air for calculating cloud ascents: this is important for producing realistic cloud profiles, particularly in the case of shallow convection.

Gravity wave drag

The current gravity wave drag (GWD) parametrization in its first version has been shown to provide excessive upper level drag and does not take into account some boundary layer dissipation processes which occur in nature. A revision to the GWD scheme has therefore been introduced to

(i) increase the surface momentum flux and introduce an additional low-level drag over orographic features, decreasing the low-level wind over mountainous areas;

(ii) modify the vertical distribution of the GWD stress, resulting in a reduction of the upper level drag; this increases the stratospheric flow over and downstream of mountain ranges, and leads to a better forecast of the jet.

Experimentation programme and main effects of the model changes in terms of physical quantities

The development and experimentation programme has been conducted over approximately two years with the above changes being tested individually and in combination, both in forecast and climate mode. The changes significantly affect several key features of the model dynamics and thermodynamics, as summarized below. Whenever possible, the results presented here come from the comparison of the 13 forecasts made during the parallel runs.

a) Energy and hydrological cycle

The new radiation scheme produces more realistic flux divergences, which were previously underestimated. This, in addition to a better formulation of cloud properties, leads to a cooling and thermal destabilization of the troposphere. In association with a more active convection scheme, the overall effect is an increase in the various terms of the energy and hydrological cycles. They are intensified by 25% and 20% respectively, as is evident from the global mean values of net radiative cooling and net heating by convection, large-scale condensation and surface heat fluxes (Fig.1) and from the values of precipitation and surface evaporation. In fact, whereas the energy balance after the spin-up in the previous operational model was maintained at too low values, the new physics now produces too high values as compared to climate estimates, with similar results for the hydrological cycle.

The spin-up of the hydrological cycle in the early stages of the forecasts is generally smaller with the new physics, which seems to indicate that the new convection scheme is more compatible with the assimilated data. However, during the first two days there is still a large imbalance between the moisture supply by surface evaporation and the loss due to precipitation, with the result that the model dries during the forecast, although less so than previously.

The forecast precipitation is now more concentrated along the ITCZ and more intense over the tropical continents.

b) Radiative fluxes and surface temperatures

A direct impact of the revised shortwave H_2O absorptivity and cloud optical properties is to decrease by about 10% the global shortwave atmospheric heating. Therefore, more solar radiation is available at the surface. This enhancement of the radiative energy at the surface contributes to the enhancement of convection over tropical continents and to warmer surface temperatures at higher latitudes, where the present operational model is often too cold. This temperature difference is somewhat reduced if one compares 2m or 30m (lowest model level) values instead of surface temperatures.

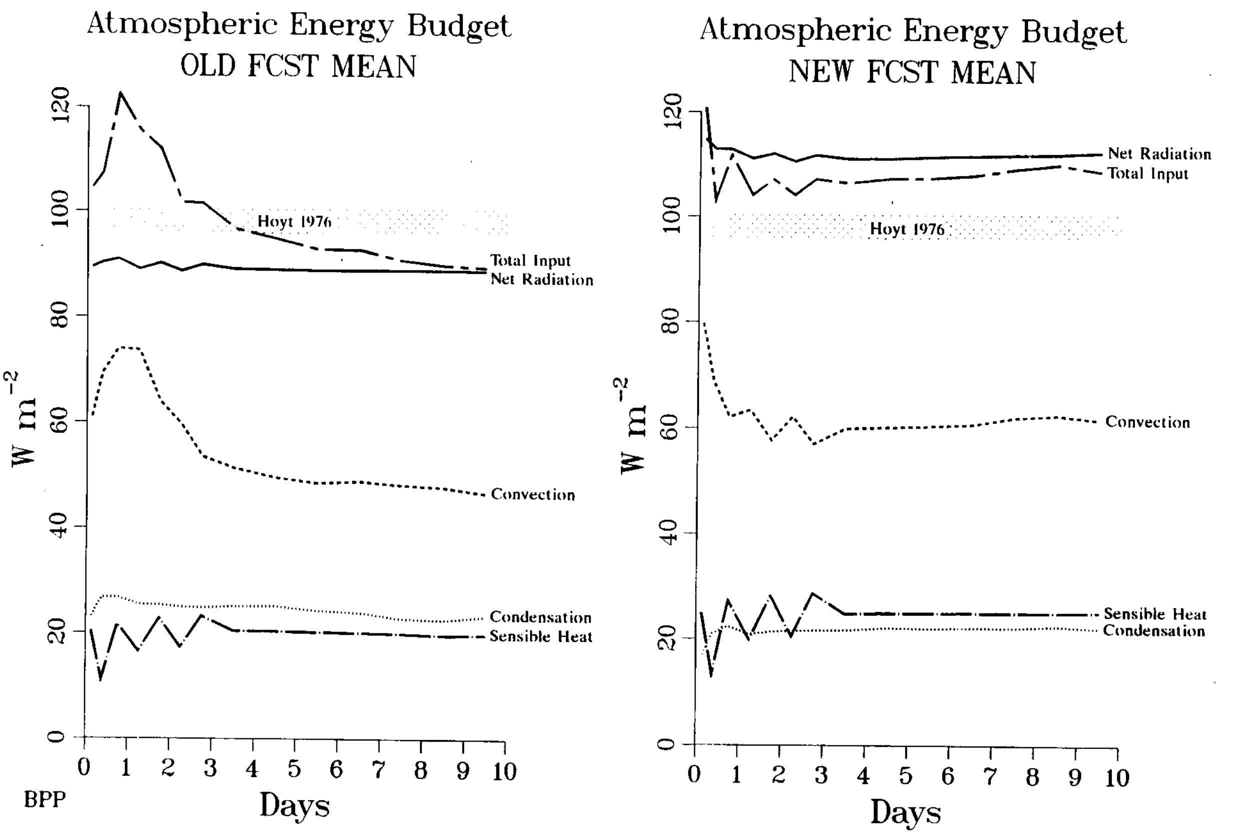


Fig. 1: Components of the global atmospheric heat budget averaged for the 13 parallel forecasts a) with previous operational physics b) with new operational physics.

A better temperature and pressure dependence of the longwave absorption corrects the underestimation of the clear-sky longwave fluxes. The modified diagnostic formulation of cloud liquid water content and revised longwave optical properties make the clouds more radiatively active. This leads to increased contrast in radiation fields at the top of the atmosphere, with marked minima over convective areas and maxima over clear-sky or low-cloud areas, in agreement with satellite observations (Fig. 2). This is an important improvement to the previous operational scheme, which failed to reproduce these features.

c) Tropical analysis and forecast

The revised physics influences the analysis and forecast of both mass and wind fields. The new scheme produces a warmer tropical lower and mid-troposphere and a colder upper troposphere whereby the temperature bias in short- and medium-range forecasts is greatly reduced (Fig. 3).

The intensification of the diabatic forcing leads to a stronger and more realistic Hadley circulation, which is seen already in the short-range wind forecasts. In the majority of cases, the mean zonal wind error is also reduced in the medium-range. 30-day simulations indicate a slight reduction of the error in winter and a larger one in summer, especially over the western Pacific, Atlantic and South America. This is essentially the result of cumulus momentum transport since the errors are increased to the same level as with the operational scheme, when momentum transport is switched off.

Tropical forecasts typically lose skill after a few days because of errors in the analysis and rapid error growth during the early stages of the forecasts. Since the analysis depends heavily on the parametrization through the first-guess, verification of short-range forecasts is very difficult and is best done by means of case studies. The study of Hurricane Gilbert (Fig. 4) has shown that the new physics has a marked positive influence on the forecast of the storm track and on the intensity of the vortex, which is considerably stronger than with the operational physics. The vertical structure, for moisture in particular, also appears more realistic with the new physics. An earlier study over tropical Australia (AMEX region) indicated that the new convection scheme intensifies the strength of tropical disturbances and tends to improve their analysis and forecast.

The effect of the stronger convective heating on the large-scale flow is most pronounced in the divergent part. The divergent circulation is intensified compared to the operational physics and maintained at a better level throughout 10-day or 30-day forecasts. In particular, the collapse of the circulation over the Indonesian area and the West Pacific, which was typical of the previous operational model, disappears and there is now a better agreement with the analyzed flow.

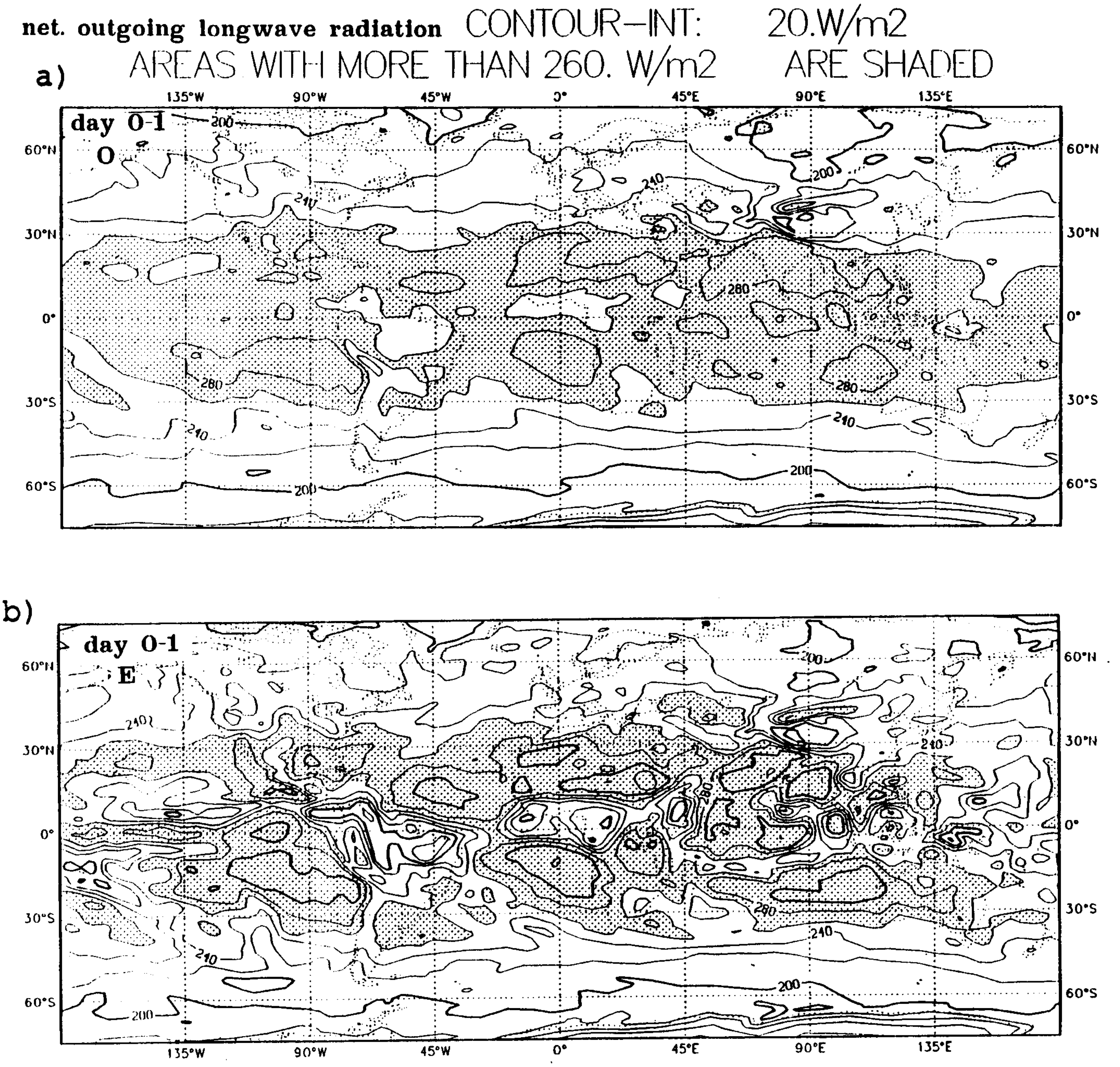


Fig. 2: Top outgoing longwave radiation at day 3 averaged for 13 parallel forecasts a) with previous operational physics b) with new operational physics.

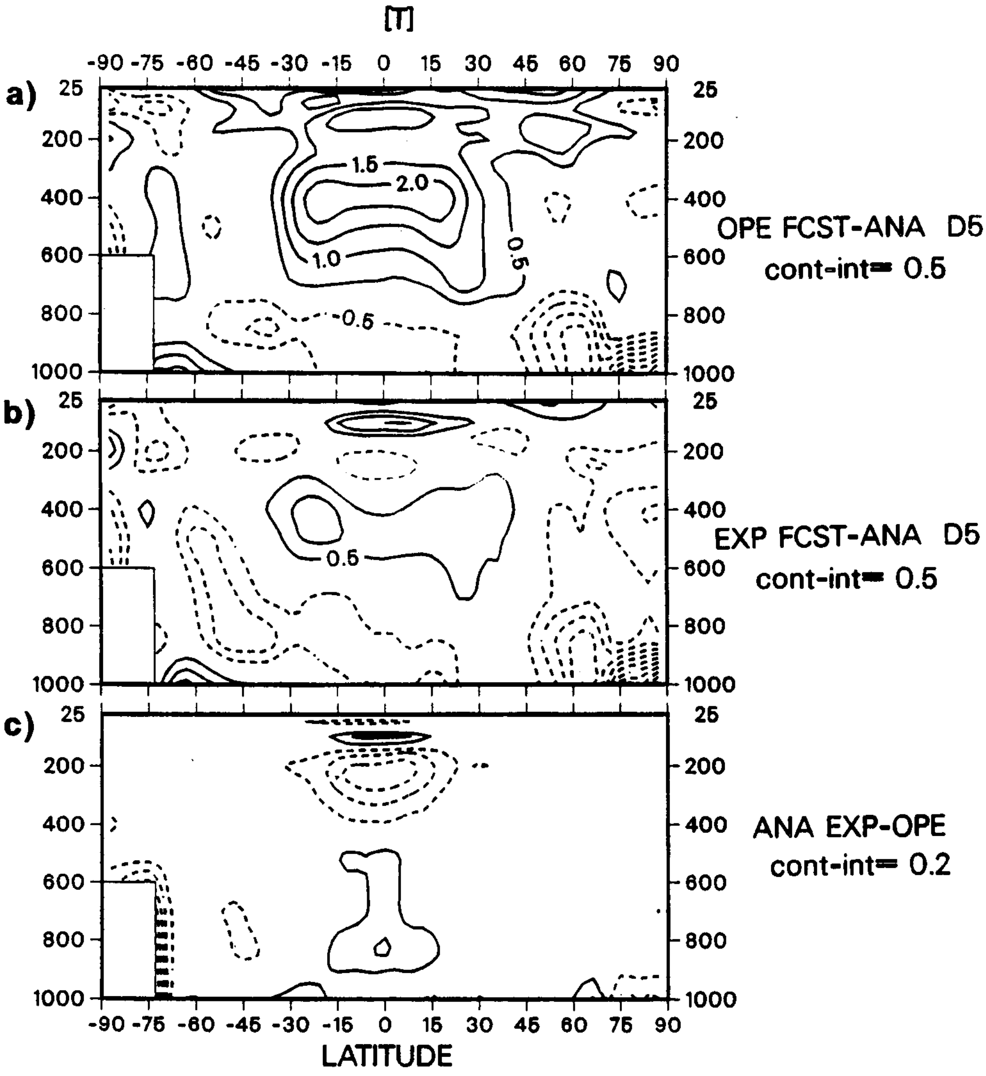


Fig. 3: Zonally averaged temperature error at day 5 (mean of 13 parallel forecasts) a) with previous operational physics b) with new operational physics c) Difference between the two sets of zonally averaged temperature fields.

d) <u>Extra-tropical forecasts</u>

The impact on anomaly correlation scores of the combined new radiation plus mass flux is positive in the average after day 5 for the northern and southern hemispheres, as deduced from 12 cases spread over the year. The corresponding scatter diagrams, however, indicate that this improvement is not systematic. This is probably due to the increase in eddy activity, which tends to spread the forecast skill measured by correlation-based scores. The effect of the GWD modification, tested on 5 northern hemisphere winter cases, shows a moderate but systematic improvement in northern hemisphere scores, which is additional to the effect of the radiation and convection changes. The effect of the GWD change is smaller in the 10-day range in the southern hemisphere.

The combined effect of the physics changes in the operational environment is illustrated in the following companion article.

- G. Sommeria, M. Tiedtke, M. Miller, J.-J. Morcrette

* * * * * * * * * * * * *

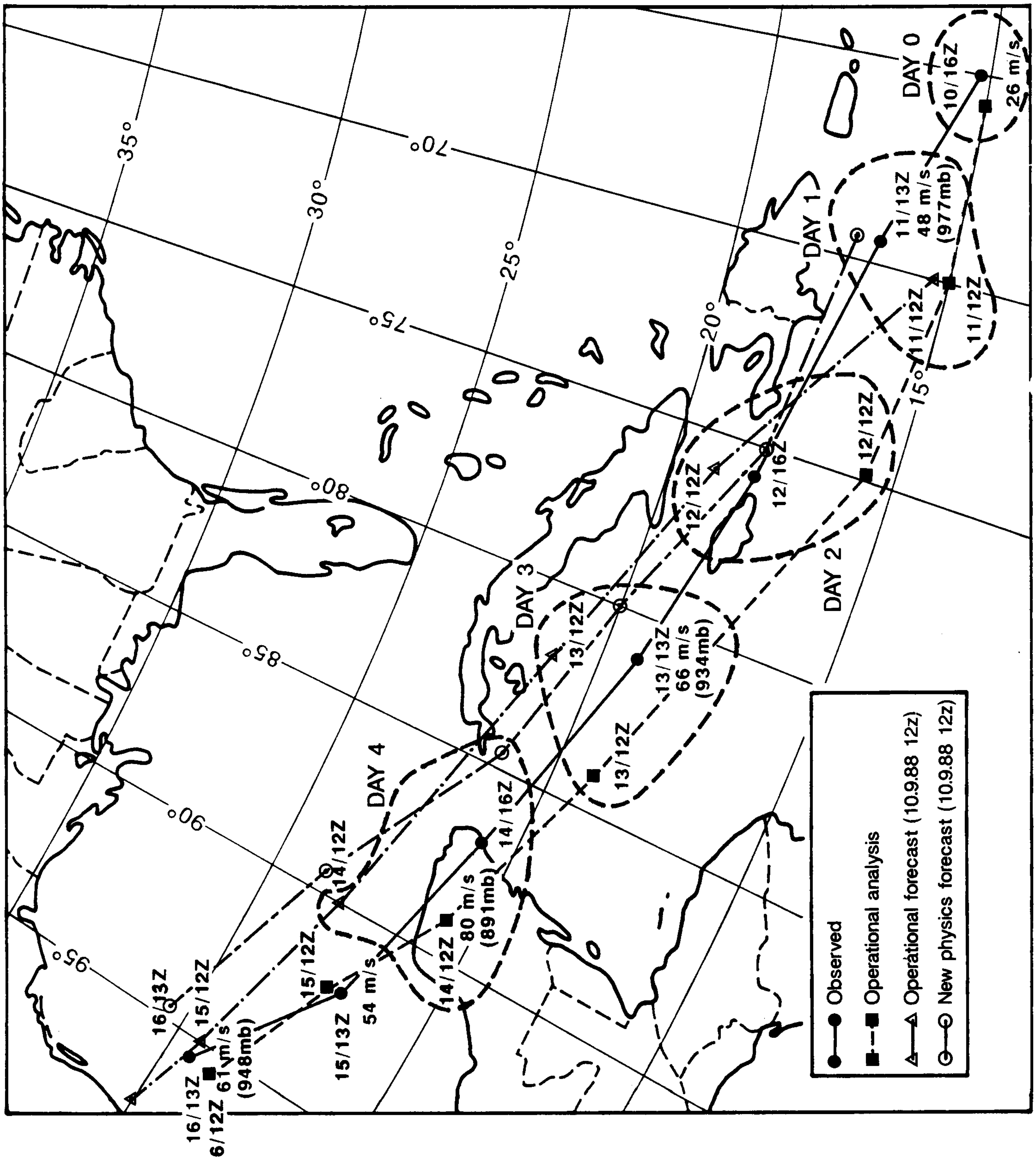


Fig. 4: Hurricane Gilbert

OPERATIONAL IMPACT OF THE NEW MODEL PHYSICS

From 19 April to 1 May 1989 a version of the forecast model incorporating the new physics parametrizations was run in parallel with the operational model. This article summarizes the main differences which have been found between the two series of forecasts, highlighting the features more relevant to the work of the operational forecaster.

2m temperature, precipitation and cloud

A significant impact was expected on cloud and precipitation and, consequently, on 2m temperature. The verifications of both 2m temperature and total cloud against European SYNOP observations over the 13 days of the parallel run do indeed show a significant reduction in the negative bias which existed in the old model. This is particularly noticeable for 2m daytime temperatures.

As for precipitation, the division between large-scale and convective precipitation has changed, so that more convective rain is produced at the expense of large-scale rain. This does not always seem to be realistic and some problems may be experienced with direct model output of convective precipitation. Some intense localised areas of large-scale precipitation have also been noticed in some tropical regions in the early stages of the forecast (first 24 hours). These seem to occur every day and are often in the same places, such as northern Brazil and eastern Africa. They are associated with a very unstable atmosphere and intense upward motion. On the other hand, convective precipitation in the tropics is more organised than in the old model. This is particularly evident in the ITCZ which is much more sharply defined.

Synoptic patterns

Noticeable differences in the forecast of the 500 hPa height field in mid-latitudes were found on several occasions, but no systematic trend was established. For surface fields, it was noticed that in most cases the test model predicted deeper surface lows; the average difference is approximately 4 hPa for forecast ranges from 3 to 5 days.

The impact of the new model in the Tropics is better assessed by looking at mean differences rather than daily charts (cf. previous article). However, Fig. 1 (a, b, c, d, e) presents the case of a tropical cyclone where the new model performed clearly better than the old one; it is the forecast of the tropical cyclone 'Orson' to the north east of Australia. The 72 hour forecast of the new model predicted a central pressure of 984 hPa and a maximum wind speed of 37 m/s, against 1000 hPa and 21 m/s according to the old model. There was one ship reporting a pressure of 948 hPa near the centre (due to lack of suitable reports it was not possible to verify the wind speed against observations).

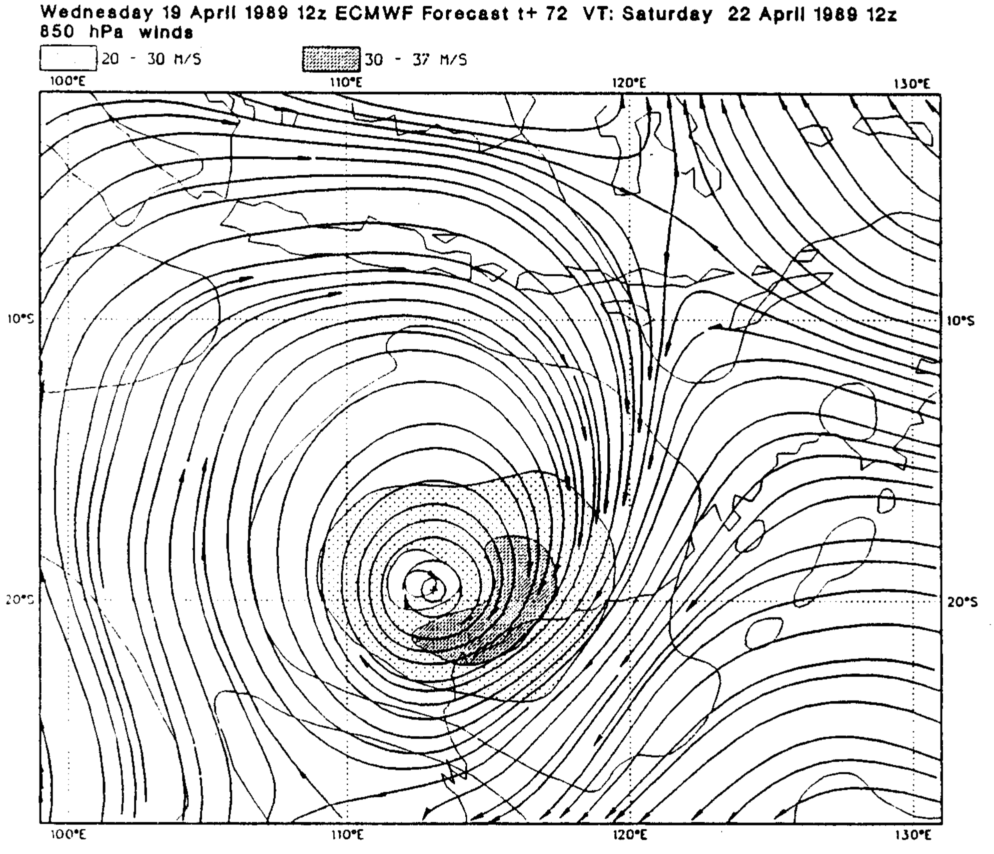


Fig. 1 a): The new model

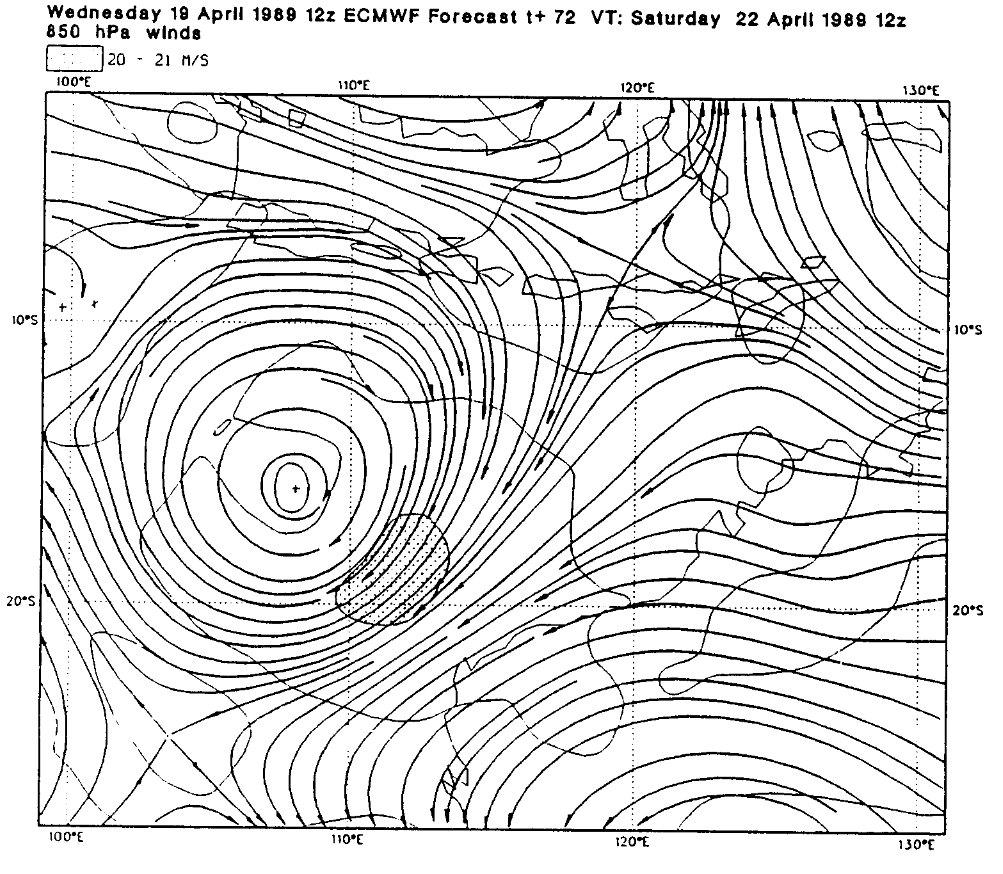


Fig. 1 b): The old model

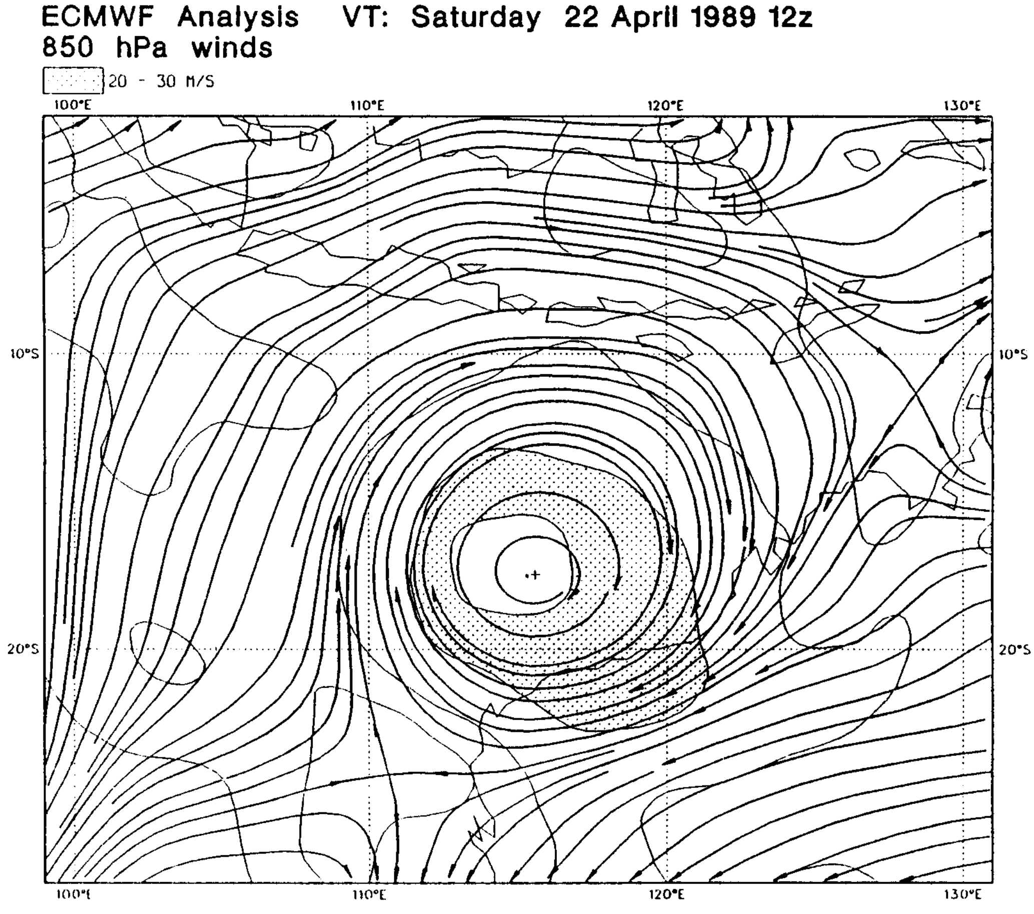


Fig. 1 c): The analysis corresponding to the forecast in Figs. 1 a) and 1 b)

Scores

Objective scores calculated on a daily basis and then averaged over the 13-day period give some idea of the statistical impact of the change as a function of days into the forecast; the largest and probably most significant difference is found for the 500 hPa anomaly correlation coefficient over Europe (Fig. 2 a, overleaf), with a gain in predictability of around 6 hours for the new model over the old one. Another noticeable difference concerns the RMS vector wind errors at 200 hPa for both hemispheres, which grow slightly faster in the new model (Fig. 2b , overleaf). In general terms, all the RMS scores averaged over the period saturate at a higher value in the new model (but note that at the same time the mean anomaly correlation coefficient is worse in the old model). In the Tropics, the absolute correlation of wind vector at 200 hPa shows no difference up to day 3, then deteriorates, but at 850 hPa there is a marked improvement throughout the forecast.

- Bernard Strauss

* * * * * * * * * * * * *

Wednesday 19 April 1989 12z ECMWF Forecast t+ 72 VT: Saturday 22 April 1989 12z
[OD Archive Version 11] SURFACE: MSL

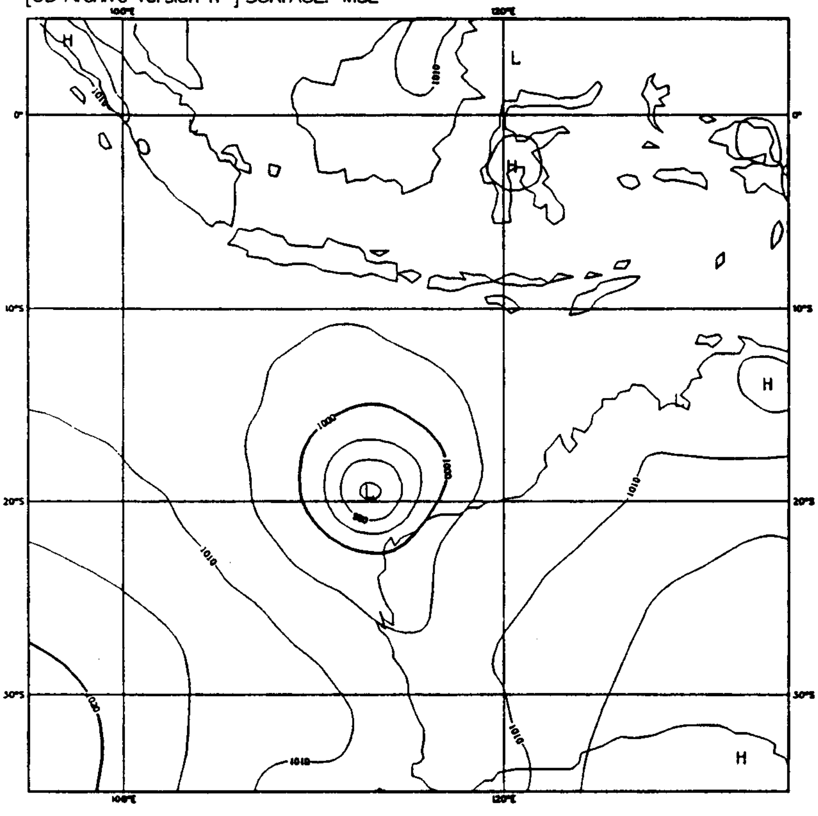


Fig. 1 d): The new model

Wednesday 19 April 1989 12z ECMWF Forecast t+ 72 VT: Saturday 22 April 1989 12z
SURFACE: MSL

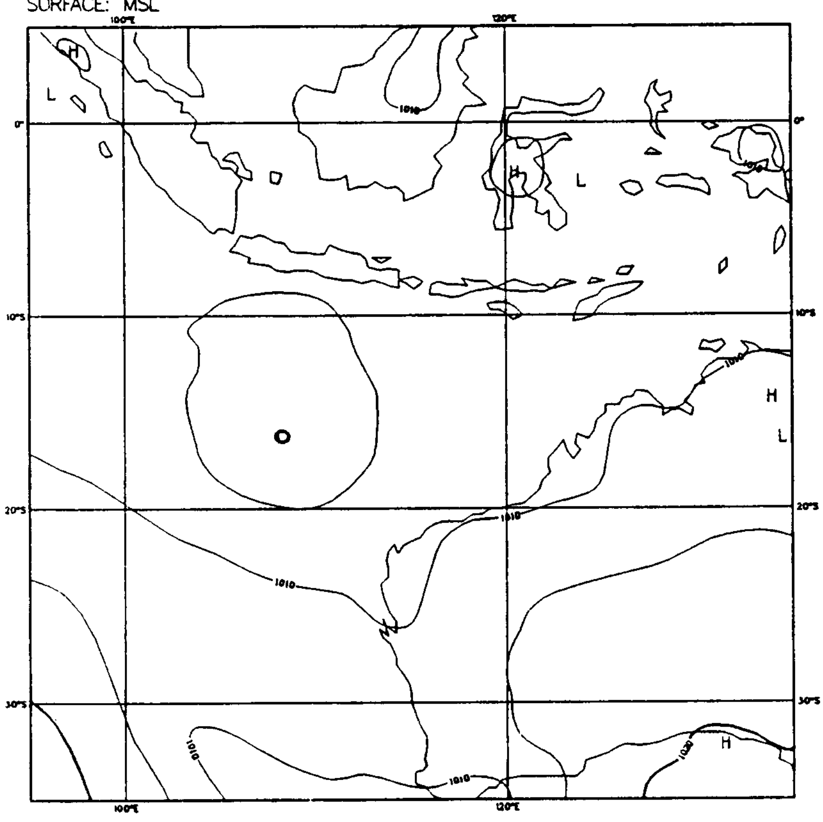


Fig. 1 e): The old model

Fig. 2 a)

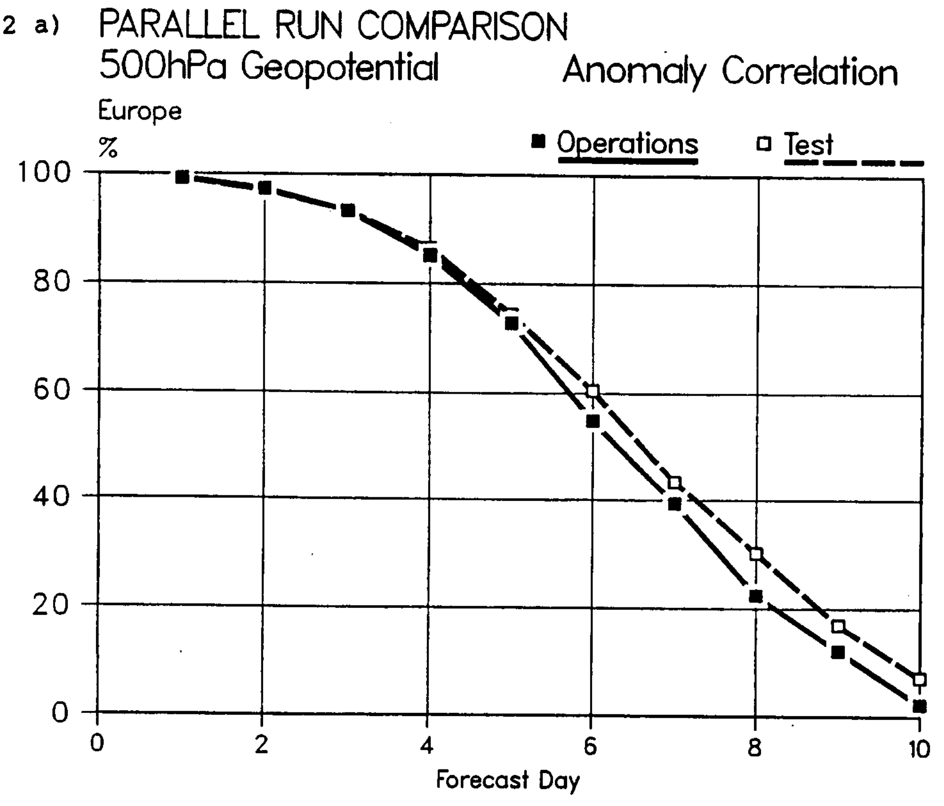
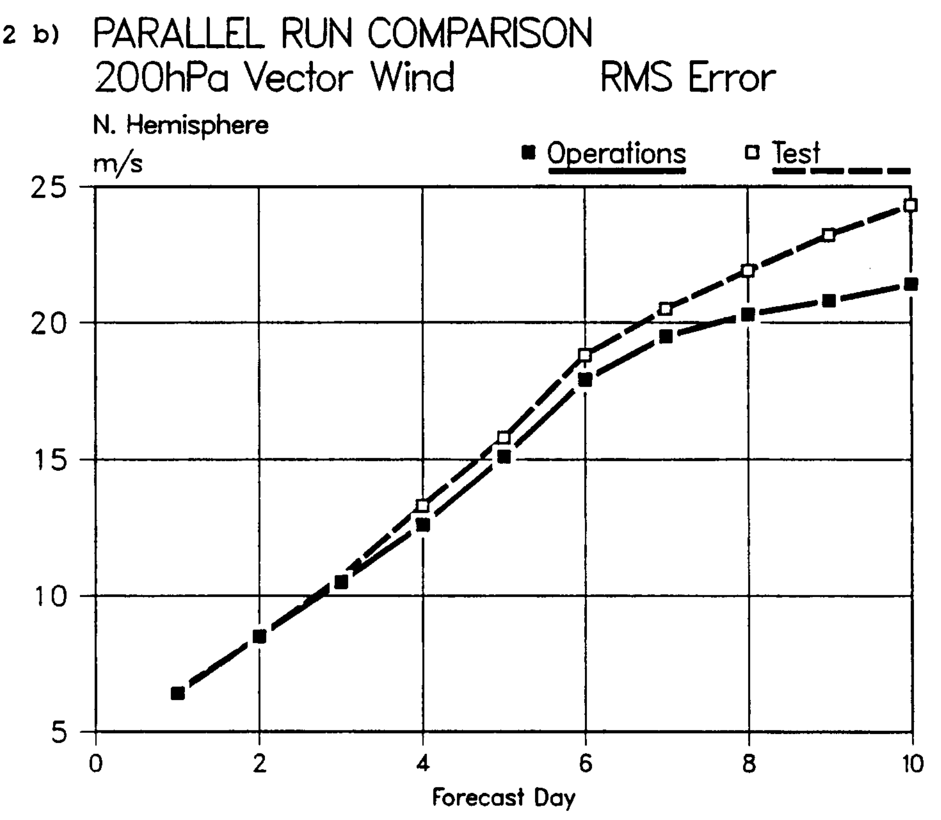

PARALLEL RUN COMPARISON
500hPa Geopotential
Anomaly Correlation
Europe
%
Operations
Test
100
80
60
40
20
0
0
2
4
6
8
10
Forecast Day

Fig. 2 b)


PARALLEL RUN COMPARISON
200hPa Vector Wind
RMS Error
N. Hemisphere
m/s
Operations
Test
25
20
15
10
5
0
2
4
6
8
10
Forecast Day

KALMAN FILTERING - A NEW APPROACH TO ADAPTIVE STATISTICAL INTERPRETATION OF NUMERICAL METEOROLOGICAL FORECASTS

The Kalman filter was developed around 1960 to be used for navigational purposes for the American space programme. From more or less accurate measurements of the positions and/or velocities of the spacecraft, the Kalman filter made an optimum estimation, in real-time, of the true position and/or velocity for the future time intervals. The use of Kalman filter in meteorology was suggested by M. Ghil (ECMWF Seminar 15-19 September 1980) as a way to facilitate the optimization of weights in numerical analysis schemes.

A fruitful application in weather forecasting, suggested in 1987 to the author by Carsten Simonsen, DMI, Copenhagen, would be to Kalman filter the output from the ECMWF model, especially the near surface parameters (2m temperature, 10m wind, dew point etc.). By Kalman filtering it would be possible not only to correct for systematic errors but also adjust the ECMWF forecast to specific weather conditions on a specific location (islands, mountains etc.).

Kalman filter - a brief description

To illustrate the basic ideas behind the Kalman filter we shall use the problem of correcting systematic errors in the forecast 2m temperatures from a numerical model. Let TFC(t) be the forecast value for a certain timestep and TOBS(t) the verifying observation at a certain location. The difference, the observed error, is then

$$TFC(t) - TOBS(t) = ERR(t) \qquad (1)$$

ERR(t) may stand for any difference between forecast and observed values, also between a forecast valid at the model (envelope) surface and an observation at some other altitude.

Due to inadequacies in the model resolution, the boundary layer parametrization and description of the state of the ground, the 2m temperature forecasts often suffer from underforecasting of temperature. This is well known among forecasters and they keep themselves informed about the actual, seasonally dependent or weather regime dependent biases for various locations within their area of responsibility. During the cold season in Scandinavia, this bias can during daytime amount to minus 4-8 degrees, during summertime only one or two degrees.

In a sense, the forecasters have conceptually formed a "statistical model" of the error in the form

$$Y(t) = \text{expected error}(t) = X1(t) \qquad (2)$$

where X1(t) is assumed to be a regionally slowly varying, seasonally and/or weather type dependent correction coefficient. All observed errors ERR(t) are, according to the theory from which the basic Kalman filters are developed, assumed to be normally distributed around X1(t).

A more elaborate model might be

$$Y(t) = \text{expected error}(t) = X1(t) + X2(t)*TFC(t) \qquad (3)$$

where we assume that the systematic errors are dependent on the temperature, i.e. the colder the forecast, the larger the (negative) error. Also, in this case, the observed errors are assumed to be normally distributed around the line defined by (3).

Our problem is that we do not know the coefficients X1 and X2, they have to be estimated. If X1 and X2 were constants in time, not necessarily geographically constant, they could be computed with good accuracy by linear regression techniques on historical data, the accuracy depending on the amount of historical data. They could also be computed recursively, by use of recursive linear regression analysis. Starting by assuming X1 = X2 = 0, we will gradually approach better estimates of X1 and X2.

X1 and X2 are not natural constants, but vary in time due to season, weather type and changes in the atmospheric model. Values of X1(t) and X2(t) typical for one period may not be representative for another period, only a few weeks later. Thus we have to update our estimations of X1(t) and X2(t) continually.

One solution would be to save data from the last 2-3 weeks and then every day perform a linear regression analysis. The values obtained for X1(t) and X2(t) valid for some weeks back are then assumed still to be valid for some days ahead. However, in keeping the estimations up to date we lose some accuracy in the estimations, owing to the reduced data volumes available.

In this context, recursive methods are better. By weighting new data somewhat higher than for normal recursive regression, they have a greater impact on the estimations of X1(t) and X2(t), which will more truly reflect the present conditions (last 3-5 weeks, with emphasis on the last 1-2 weeks).

By changing the weighting on new data, this recursive linear regression can be made more or less sensitive to new information, and so the adaptation can be varied. The weightings can either be optimised by being tested on historical data or changed on a daily basis by the system itself, according to some chosen set of rules. One can assume that during the transition seasons, spring and autumn, the system works better if it can adapt more easily to new circumstances than during winter and summer, when the conditions are fairly similar for several months.

Such a scheme can be developed using fairly simple recursive regression techniques. The advantage with the Kalman filter is that it is specifically designed to handle such problems. In a very elegant way, economizing the computer resources, it provides the user with a very efficient tool. Problems such as adjusting the adaptivity by giving weighting to data of different qualities or temporary lack of data etc. are rather easily dealt with. The main problem, formulation of a suitable model, however, is left to the meteorologist who develops the system. Models other than (1) and (2) probably better describe the systematic errors, at least when it comes to dewpoint temperature and wind speed.

It is important to understand that the filter "remembers" past data only in the form of actual coefficients X1 and X2 and covariances between errors of these co-efficients. What is intuitively felt as "memory" is the degree of adaptivity. The only past data that has to be stored in an operational system using Kalman filter are yet to be verified forecasts. As soon as they have been verified, and the coefficients and covariances correspondingly modified, these data are not required.

Through external interference, normally automatically, the adaptivity can be changed, e.g. when external conditions have altered fundamentally. When there is a transition from snowcover/bare ground (both in the ECMWF model and in reality) the coefficients have to adjust to new types of systematic errors.

During spring 1988 a typical correction for 12 UTC temperatures in Stockholm during March was according to an operational Kalman filter:

$$Y(t) = 0.5 - 0.4*FCT(t)$$

during the snowmelting period 28 March - 8 April, the Kalman filter slowly changed this model to

$$Y(t) = 1.9 + 0.1*FCT(t)$$

which then was almost unchanged during the whole of April and beginning of May.

Results

Fig. 1 shows an example of early morning temperature forecasts for Stockholm based on ECMWF forecasts during the second half of January 1989. The forecast range is 66 hours and the graphs compare the direct model output with the prediction adapted by Kalman filtering and the final product issued by the meteorologist. The verifying observations are shown as individual dots. The statistical adaptation clearly reduces the bias, providing the forecaster with better guidance. Whereas the subjective forecasts on 6 occasions (out of 17) suffered errors exceeding two degrees, this was never the case with those based on the Kalman filter. The dampening of the variability of the forecast temperature (compared to the observed and subjectively forecast) might be interpreted as a weakness, but is in complete agreement with the objective of minimizing the error of a forecast valid at a specific time. If extremes are the focus of interest, the predictand should be changed to the maximum or minimum temperature during a certain time interval.

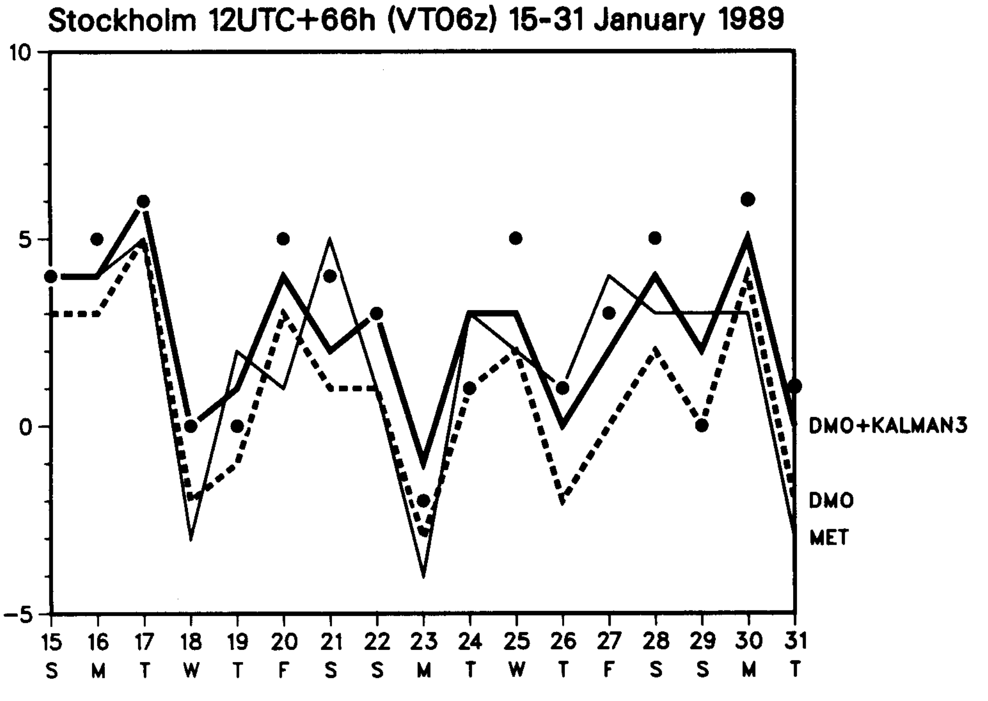


Fig. 1: Forecasting the morning temperature in Stockholm two days ahead. Original ECMWF 2m temperature (dashed thick line), Kalman filtered (solid thick line) and (of the Kalman filtering) independently made subjective forecast by the duty meteorologist (thin solid line) made in the morning 48 hours earlier.

Conclusions

By using a Kalman filter a new approach to correcting direct model output has been developed and evaluated. The approach described above avoids the problems of large historical data files and changes in the atmospheric models. Whereas conventional MOS or PPM methods generally use predictors from the free atmosphere and leave it to the statistical schemes to work out the conditions in the boundary layer, the philosophy behind this application of the Kalman filter is to let the atmospheric model do the main "thinking", leaving it to the statistical scheme only to adjust for seasonally dependent, systematic biases, due to e.g. deficiencies in the model's boundary layer parametrization.

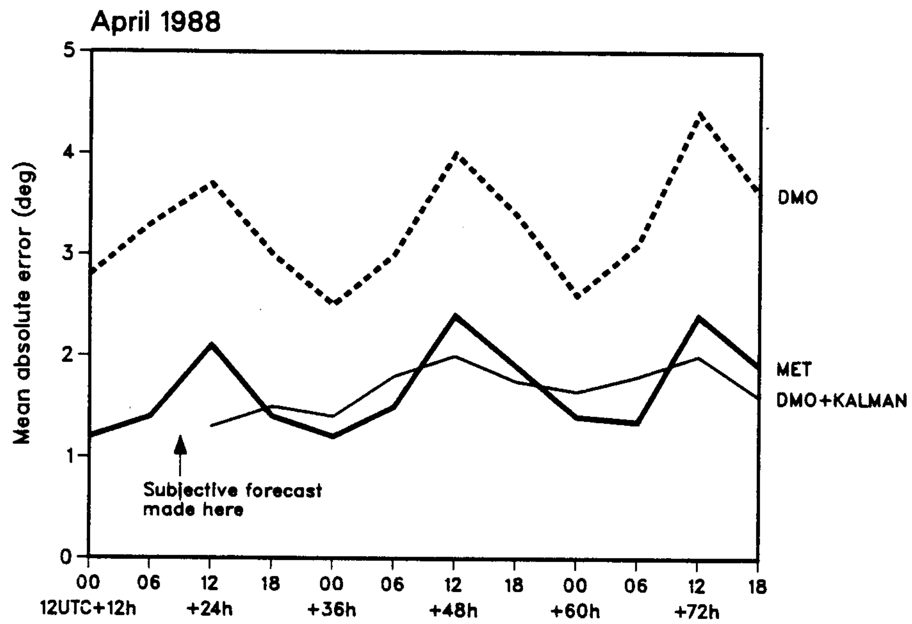
The limited historical memory is compensated for by a more flexible adjustment to extreme seasons (early winters, late springs with late snowcover). By externally changing the sensitivity of the Kalman filter or reducing its "memory" is it possible to facilitate the adjustment to changes in the systematic errors.

Since November 1988 daily forecasts three days ahead have been transmitted to an energy board in central Sweden providing Kalman filtered values of 2m temperature using the following model:

$$Y(t) = X1(t) + X2(t)*FCT + X3(t)*FCT850.$$

Instead of using the temperatures, forecast anomalies have been used, making the coordinate system "follow along" with the observations during the seasonal variations.

The forecasts have proved of high standard and have been produced at a lower cost than would have been the case if subjectively made by a forecaster. Verification results show that Kalman filtered ECMWF forecasts exhibit the same or higher skill than forecasts independently made by experienced meteorologists.



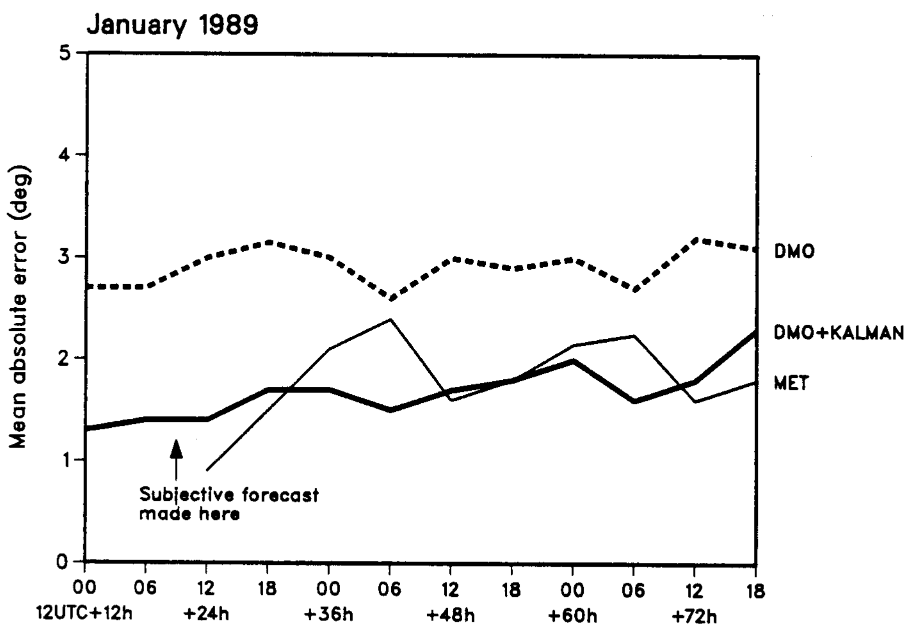


Fig. 2: Mean absolute errors for Kalman filtered and non-Kalman filtered ECMWF 2m temperature forecasts for Stockholm in April 1988 and January 1989, months characterised by large systematic errors. For comparison, the subjective scores are shown (thin line).

- Anders Persson,
SMHI, Norrköping

* * * * * * * * * * * *

THE NEW DATA ACQUISITION SYSTEM

In October 1988, the Centre took delivery of a new DEC VAX 6210 with 16 megabytes of memory and two 600 megabyte disks. This machine provides the motive power for the new data acquisition and preprocessing systems, which will start to take over from the old NOS/BE-based systems later this year. The new systems will include several technical advancements (see 'The New ECMWF Preprocessing Scheme', ECMWF Newsletter No. 43).

The data acquisition part of the overall project is concerned with receiving incoming observations, sorting them into streams for preprocessing, and filing all the messages into a message database for future access.

An experimental version of the data acquisition has been running on the VAX 6210 since the New Year, and work is proceeding to integrate it with the preprocessing system.

Data acquisition defined

In the ECMWF context, data acquisition consists of:

1. Keeping track of incoming files of observation messages received from the Telecommunications Centres (Deutscher Wetterdienst and United Kingdom Meteorological Office), and processing them on a first-come, first-served basis.

2. Extracting individual messages from the files and determining:

   a) validity - the message must have a valid WMO format header, which details the message type, the geographical area of the observation, and the time at which it was made;

   b) duplication - the same message may be received several times: both telecommunications centres intentionally send the same data and partial transmissions may cause additional repeats;

   c) stream assignment - a message with an intelligible header can be assigned to a preprocessing stream on the basis of its type.

3. Writing the messages to the appropriate stream of the message database for the preprocessing system to read. Duplicate messages are not passed to the preprocessing, but are kept in a 'shadow' database so that they can be checked manually if needed.

Further features

To complete a usable data acquisition system, a number of other features must be added to the bare bones outlined above. Some of the more important are:

Crash-proof operation

Inevitably, an operational system like the data acquisition can crash occasionally, as can the computer system supporting it. It has to be ensured that the data acquisition never 'drops' a file or a message as a result of such a crash.

Communication with preprocessing

The preprocessing system which is receiving the messages from data acquisition will consist of a number of separate jobs, each of which handles a particular stream and may be stopped and started at any time. Eventually, the preprocessing job will 'catch up' with the data acquisition: when that happens, it must wait for more data. The data acquisition then alerts the preprocessing job, when it has provided more data for it.

Archiving of the message database

It is not an absolute requirement that the messages must be saved in an archive in their original form, but the overall system provides a much better service if this is done.

Extraction from the message database

Either for internal use or for Member States' own processing, the message database must allow sets of data to be extracted for special purposes. In the case of Member States, the most useful service would allow a defined 'shopping list' of messages to be passed on as they are received.

Additional sources of data and test databases

For experiments, or for developing new preprocessing modules, it is useful to be able to introduce special data (either from tape or from other telecommunications centres) into the data acquisition. For the same reasons, it is necessary to be able to run a 'test' message database alongside the 'production' message database without any interaction between them. It is useful to be able to add special decoding modules to the data acquisition when running with new data sources.

The overall design

The design which has finally been adopted for the data acquisition allows for all the above features. So far as we can tell, it has never lost a message during the tests (though a message may occasionally be passed on twice as a result of a crash). There is a message database on the VAX, along with an access package for the preprocessing suite, which allows straightforward access to the messages and which automatically feeds the next unprocessed message to the preprocessing stream, even if the stream previously crashed or was cancelled. New sources of data can easily be connected to the system, and several separate data acquisition systems can be run in parallel, if necessary.

Fig. 1 shows the structure of the data acquisition, and its relationship with the preprocessing suite (rather simplified in this representation).

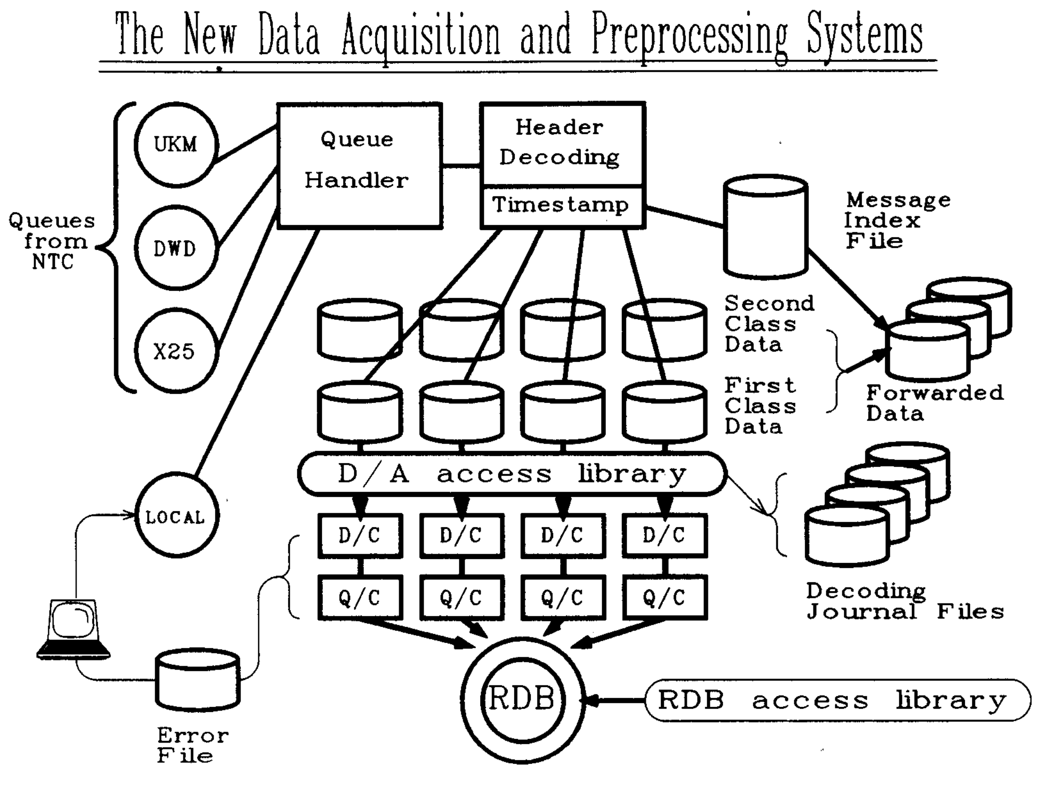


Fig. 1

The origin of all data in the system is a group of input queues, which are normally built and filled by the telecommunications system (NTC/NTS). A new file of messages arrives every ten minutes from United Kingdom Meteorological Office and every twenty minutes from Deutscher Wetterdienst. The data acquisition scans the queues and checks the files, looking for a time when they were placed there. Although data acquisition is normally expected to keep up with arriving data (it can handle about ten times the amount of data we are currently receiving), it always processes the oldest files first. The diagram also shows queues created by data arriving on X.25 packet switched connections, and a queue created by a manual correction process.

These queues are normal VAX queues, which can be examined at any time to check how big a backlog exists, what time the various files arrived, etc.

When the data acquisition finds that any of the queues contains unprocessed data, it takes the oldest unprocessed file from the queue and unpacks the messages contained in it. Each time a message is removed from the queue, a timestamp value is created from the VAX system clock and attached to the message. This timestamp will stay with the message throughout its life in the data acquisition system, and serves to identify it uniquely.

The message database consists of a number of separate files, each of which feeds a single preprocessing stream: these are known as 'stream files'. Stream 1, for example, receives all messages of types SI, SM or SN (surface temperature reports). To form these separate stream files into a single logical database, an additional index file which includes an entry for every message handled by the data acquisition, is also kept. For each incoming message the first task is to create a new index entry.

While the index entry is being created, it is a simple matter to check whether an identical message has already been received. A message is regarded as a duplicate if it has the same WMO header and is no longer than the earlier version (since it is not uncommon for a partial message to be received, followed by the whole message). Duplicate messages (unless they were created by human intervention) are regarded as 'second class data', and are written to a set of files which will not be archived or passed to preprocessing. New messages are 'first class data', and are written to the main message database files.

The index to the message database contains all information about each incoming message except for the actual message contents. By reference to the index, it is possible to answer questions such as 'are messages received earlier, on average, from Deutscher Wetterdienst than from United Kingdom Meteorological Office, and by how much?', or 'what proportion of messages from United Kingdom Meteorological Office arrive in the hour 00 UTC to 01 UTC, etc. It is also possible to find which of the telecommunications centres provided any individual message, to find what time the message arrived at the Centre, and when the message was entered into the message database.

The preprocessing streams read their data from the message database files at the same time as the data acquisition is writing them. In order not to use up too much time in checking the files for new data when the preprocessing is waiting, the files are read through an access library which includes a signal mechanism. When a preprocessing stream has read all the data available, it will become inactive until data acquisition activates it. Data acquisition does this each time it finishes processing an incoming file, so that preprocessing will normally follow along one file behind data acquisition. Files arrive about every six minutes, so that the preprocessing batch jobs will not waste too much time in stopping and restarting frequently.

To allow the automatic restarting of a preprocessing stream from the same point at which it stopped, the access library also keeps a journal file for each preprocessing stream. Into the journal file are written records indicating the start and finish of each file received from the queues. When a preprocessing stream suffers an unscheduled stop (crash or cancel), the state of the journal file always indicates precisely which queue file the stream had been processing at the time of the stop. When the preprocessing stream starts up again, the message database access mechanism repositions the stream input file, so that the preprocessing stream continues to read messages from the same queue file. In the most common case, where the preprocessing stream was paused between files, messages are neither lost nor duplicated. If the system or the preprocessing crashes during the handling of messages from a queue file, messages are reprocessed from the start of the file.

At the same time as the preprocessing stream is being chosen for a message, data to be forwarded to Member States can be selected. It is possible also to extract data from the message database later, if the selection characteristics are awkward to apply at arrival time.

Finally, some time after the message files have all been handled by preprocessing and their contents entered onto the reports database, the stream files and the index file are archived. It is planned that these files will be simply copied into CFS, through the direct interface on the VAX. Most accesses to the message database archives will be for statistical purposes, and these can be satisfied by looking at the index file only. When actual messages must be fetched from the archives, the index file indicates which stream received the message in question, and so which stream file has to be fetched out of CFS. The 'second class data' files which received the duplicated messages are simply deleted from disk: they are kept initially so that a human operator can select an alternative copy of a message, if one exists.

Current status

The data acquisition system has been run on a regular basis since the New Year, and appears to work well. The VAX interface to CFS, needed for file archiving, is not yet working, so that a fully operational data acquisition cannot yet be constructed. However, work is going ahead to connect the preprocessing streams to the message database, and the target of achieving an operational role for the overall system later this year looks achievable.

- Dick Dixon

* * * * * * * * * * * *

CURRENT CRAY SERVICE

The Centre's present CRAY X-MP/48 went into service on 11 February 1986. Initially, it had plenty of capacity, so turnaround was quick. As time has progressed, so the system, like all machines providing a service to a large user community, has become more and more congested. Users write bigger programs requiring more memory, analyse more data, requiring heavier I/O traffic across the link and run more jobs, leading to lengthening job queues and worsening turnround. All this inevitably leads to more system activity too.

At present we have the following situation compared to 1986:

- operating system overheads have increased 55%
  from 6.2% to 9.6% of the CPU utilisation

- blocked I/O time (when no job wants the CPU but all are waiting for I/O) has increased 30%
  from 24% to 31% of the total utilisation

- link traffic has increased 18%
  from 0.88% to 1.04 Mbits/second average

- user CPU time has decreased 12%
  from 57% to 50% of the CPU utilisation.

Figs. 1 and 2 show the average daily turnround for two popular classes of job, NORMAL1 and NORMAL2. As can be seen, turnround day by day is erratic. However, the trend of the peaks is that they get higher and higher as time progresses.

The main bottleneck at this time is the size of main memory (8 Mwords); the machine is always full of jobs. As the average memory occupied by each job has increased, so jobs are rolled in and out of memory more frequently. On average, one job is now rolled every 1.6 seconds, which increases system overheads.

Paradoxically, the introduction of CFT77, although good for individual jobs, has worsened the overall user CPU utilisation. This compiler improves (i.e. reduces) the total CPU time a given job takes, and hence reduces the number of accounting units used, however, it does not affect the I/O time. Consequently, many of the major jobs in the Centre are now I/O bound, reducing the CPU utilisation. This means that jobs take roughly the same elapsed time, although the CPU does less work.

The effect of more data being transferred is, inevitably, that jobs wait longer for the I/O to complete. This, in turn, increases the chance that at any one time there are no jobs ready to use the CPU, hence blocked I/O time increases.

The overall picture, therefore, is of a machine fully loaded, often overloaded. A small increase in any part of that load can lead to a dramatic worsening of some measure of performance; that is, we are constantly on the edge of instability. Hence the erratic turnround which users have been experiencing recently.

\- Andrew Lea

* * * * * * * * * *

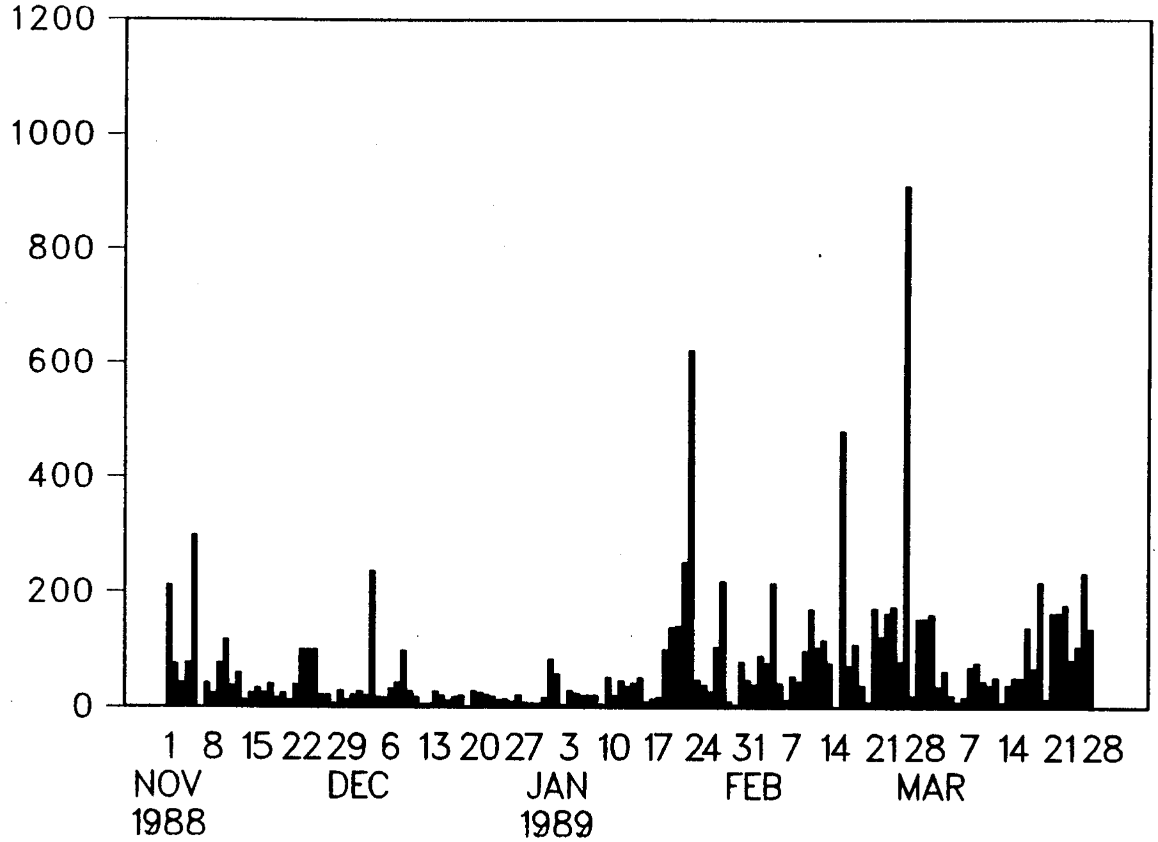


Fig. 1: Average turnaround (in minutes) of Cray jobs in NORMAL1 class.

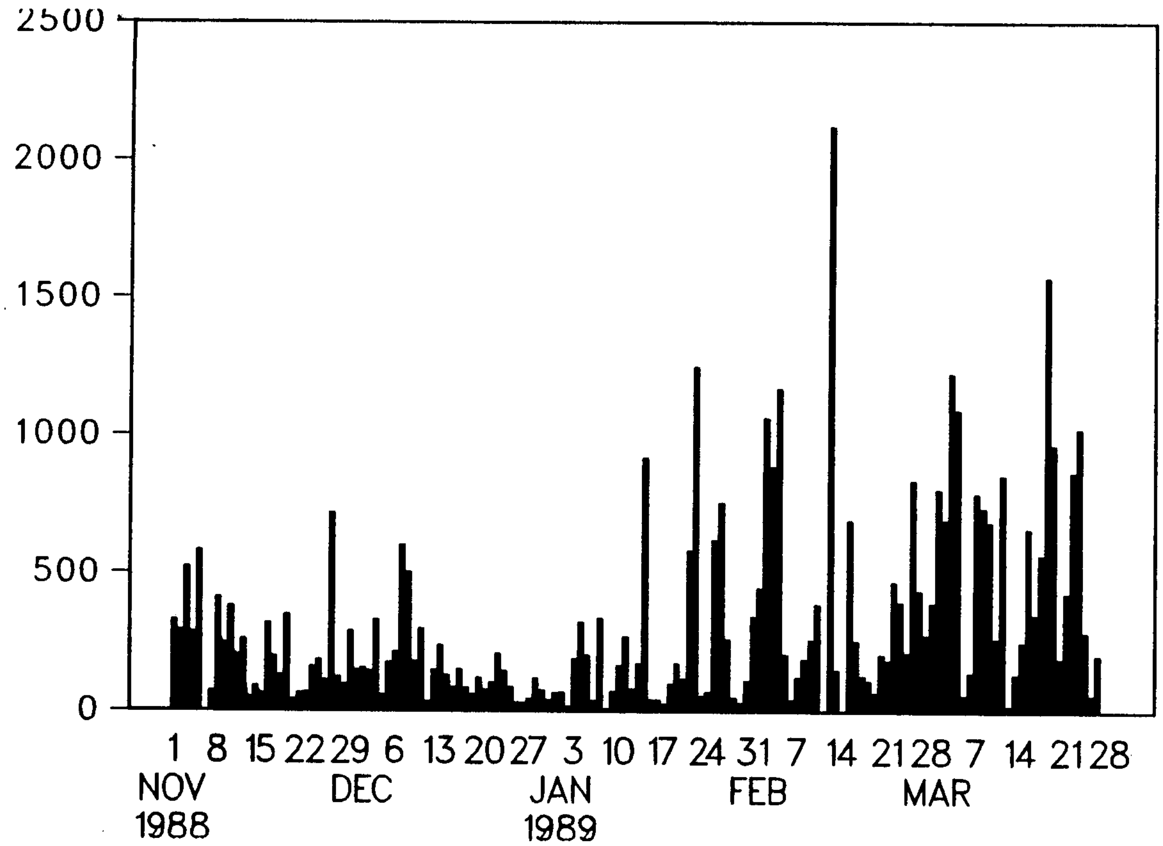


Fig. 2: Average turnaround (in minutes) of Cray jobs in NORMAL2 class.

STILL VALID NEWS SHEETS

Below is a list of News Sheets that still contain some valid information which has not been incorporated into the Bulletin set or republished in this Newsletter series (up to News Sheet 235). All other News Sheets are redundant and can be thrown away.

No.	Still Valid Article
16	Checkpointing and program termination
67	Attention Cyber BUFFER IN users
73	Minimum Cyber field length
89	Minimum field length for Cray jobs
93	Stranger tapes
120	Non-permanent ACQUIRE to the Cray
121	Cyber job class structure
135	Local print file size limitations
140	PURGE policy change
158	Reduction in maximum print size for AB and AC
176	Archival of Cyber permanent files onto IBM mass storage
178	TIDs on Cray include 2 chara. TID plus 3 chara. source computer ID.
	Caution with ACQUIRE on RERUN jobs
186	PROCLIB changes
187	Maximum memory size for Cray jobs
189	ROUTEDF
190	Using ROUTE to direct RJE output to the Centre
194	Preventive maintenance schedules
198	Using the MOHAWK printer
201	New Cray job classes
203	Magnetic tape problems and hints on avoiding them
204	VAX disk space control
205(8/7)	Mispositioned cursor under NOS/VE full screen editor
207	FORMAL changes under NOS/VE
	Job submission from within a Cray job, using LAUNCH
208	Restriction of Cray JCL statement length
212	MFICHE command from NOS/VE
214	NAG Fortran Library Mark 12
	News Sheets on-line
215	MARS - data retrievals and model changes
219	MARS-Retrieval of most recent fields extraction utility
223	Corrections to ECFILE bulletins B8.3/1 and B8.3/2
	Aborting programs under VAX VMS
224	CRAY deferred class
	Job information cards
226	CRAY Class X
227	Extension of NOS/VE SUBCJ.
228	NOS/VE level 1.4.1
229	ECFILE audit facility
230	Access to AB printer via NOS/VE CDCNET
	Replot facility for DISPLOT
231	METGRAM under NOS/VE
232	NOS/VE passwords - how to change
235	VAX public directory - how to create

* * * * * * * * * * * * *

SECOND WORKSHOP ON METEOROLOGICAL OPERATIONAL SYSTEMS

4-8 December 1989

Introduction

Since the first workshop on Meteorological Operational Systems, held at ECMWF in December 1987, much progress has been made in establishing binary data representation and codes as the approved formats for data exchange in the meteorological community. Future extensions of the codes aim at the provision of the binary equivalent code in a tabular form and at the integration of oceanographic data in the binary representation, for which co-operation between WMO and the IOC will be required.

Standard data interfaces provide the means to develop effective operational systems to generate, preprocess, postprocess, monitor and display the data which are either model generated or collected from the global observational network.

The workshop will examine the progress that has been made over the last two years, report on the development of state of the art meteorological systems and address future trends in Data Management in the context of the WMO long-term plans.

Use of binary data representation in meteorological systems

Following the approval of GRIB and BUFR by WMO/CBS and with the further development of BUFR in tabular form (BTAB) the data management tools have been, and will continue to be, created to develop standard meteorological systems for data acquisition, quality control, preprocessing and archiving which allow efficient exchange, processing and storage of data. Coupled with the continuing development of standards for programming languages and software transportability, many of the new meteorological applications systems will be sufficiently portable to be of general interest. The workshop will examine the present status of binary data representation standards, the systems for management of meteorological data in use or under development at various centres and the future trends in data management.

Observational meteorological data

Data requirements for global numerical forecasting will be discussed and compared with current data availability and quality. Lead centres for data monitoring appointed by WMO/CBS will present their initial experiences. Discussions will also include an investigation of the ways in which remedial action may efficiently be taken to overcome data deficiencies.

Visualisation of meteorological data

High resolution global forecast models generate output data which can only be evaluated and inspected by means of effective graphics tools. The increasing amount of satellite data requires additional graphics tools to enable the meteorologist to comprehend the data and compare them with the model results. State of the art graphics systems will be presented and demonstrated. The importance of 3-D graphics, animation and image processing on PCs and powerful workstations will be discussed with a view to working out suggestions for future developments.

Horst Böttger

* * * * * * * * * * * * *

ECMWF CALENDAR

17 April-16 June	Meteorological Training Courses
	Met 3 Use & interpretation of ECMWF products (5-16 June)
28 August	*ECMWF holiday*
4-8 September	Seminar - Ten years of medium-range forecasting
25-27 September	Scientific Advisory Committee - 17th session
27-29 September	Technical Advisory Committee - 14th session
3-5 October	Finance Committee - 44th session
29-30 November	Council - 31st session
4-8 December	Second workshop - Meteorological operational systems
25-27 December	*ECMWF holiday*

* * * * * * * * * * * *

ECMWF PUBLICATIONS

TECHNICAL REPORT No. 62: Atmospheric effective angular momentum functions for 1986-1987

* * * * * * * * * * * *

INDEX OF STILL VALID NEWSLETTER ARTICLES

This is an index of the major articles published in the ECMWF Newsletter plus those in the original ECMWF Technical Newsletter series. As one goes back in time, some points in these articles may have been superseded. When in doubt, contact the author or User Support.

		Newsletter	
	No.*	Date	Page
CRAY			
Bi-directional memory	25	Mar. 84	11
Buffer sizes for jobs doing much sequential I/O	14	Apr. 82	12
CFT 1.11 Subroutine/function calling sequence change	19	Feb. 83	13
CFT 77	36	Dec. 86	12
CFT 1.14	32	Dec. 85	22
COS 1.14	32	Dec. 85	22
Cray X-MP/48 - description of	30	June 85	15
Cray X-MP/22 - hints on using it	26	June 84	10
Dataset storage	13	Feb. 82	11
Multifile tapes - disposing of	17	Oct. 82	12
Multitasking ECMWF spectral model	29	Mar. 85	21
	33	Mar. 86	9
	& 37	Mar. 87	5
Public Libraries	T5	Oct. 79	6
CYBER			
Arithmetic instructions - comparative speeds of execution on the Cyber front ends	14	Apr. 82	17
Cyber front ends - execution time differences	15	June 82	9
Buffering or non-buffering on Cyber?	15	June 82	10
CMM-Fortran interface	10	Aug. 81	11
Cyber 855 - description of	21	June 83	18
Dynamic file buffers for standard formatted/ unformatted data	3	June 80	17
Formatted I/O - some efficiency hints	4	Aug. 80	9
FTN5 - effective programming	9	June 81	13
	& 10	Aug. 81	13
- optimisation techniques	14	Apr. 82	13
	& 15	June 82	10
Magnetic tapes - hints on use	T2	Apr. 79	17
- making back-up copies	1	Feb. 80	9
- stranger tapes: slot numbers	36	Dec. 86	15
Public libraries	T5	Oct. 79	6

GENERAL	No*	Newsletter Date	Page
Computer changes planned in 1988	40	Dec. 87	22
COMFILE	11	Sept.81	14
Data handling sub-system	22	Aug. 83	17
Data handling system, phase 2	39	Sept.87	12
ECFILE - the ECMWF permanent file system	42	June 88	14
ECMWF publications - range of	26	June 84	16
Electronic mail to ECMWF	44	Dec. 88	20
MAGICS - the ECMWF meteorological applications graphics integrated colour system	35	Sept.86	20
Magnetic tapes - various hints for use of	31	Sept.85	17
MARS - the ECMWF meteorological archival and	32	Dec. 85	15
retrieval system	& 33	Mar. 86	12
Member State TAC and Computing Representatives and Meteorological Contact Points	45	Mar. 89	14
Output files - controlling destination of, in Cray and Cyber jobs	14	Apr. 82	20
Resource allocation in 1989	45	Mar. 89	11
Resource distribution rules	18	Dec. 82	20
"Systems" booking times	27	Sept.84	
Telecommunications - description of new system	31	Sept.85	13
Telecommunications schedule	32	Dec. 85	19
Upper and lower case text files	11	Sept.81	15

METEOROLOGY			
ALPEX: the alpine experiment of the GARP mountain sub-programme	14	Apr. 82	2
Alpex data management and the international Alpex data centre	11	Sept.81	1
Bogus data - use of to improve analysis and tropical forecast	45	Mar. 89	3
Cloud Cover Scheme	29	Mar. 85	14
Divergent structure functions in the analysis	42	June 88	2
Diurnal radiation cycle - introduction of	26	June 84	1
Envelope orography - discussion of its effects	33	June 86	2
ECMWF Analysis and Data Assimilation System	T3	June 79	2
ECMWF Analysis System - new version	35	Sept.86	16
ECMWF Limited Area Model	16	Aug. 82	6
ECMWF Operational Schedule, Data and Dissemination	12	Dec. 81	1
ECMWF Preprocessing - new scheme	43	Sept.88	3
ECMWF Production Schedule	6	Dec. 80	5
Facilities to verify and diagnose forecasts provided by the Data & Diagnostics Section	8	Apr. 81	3
Forecast products of various centres decoded and plotted at ECMWF	9	June 81	3
Forecast model - T106 high resolution	29	Mar. 85	3
Global forecast experiment at T213 resolution	41	Mar. 88	3
GTS: ECMWF grid code product distribution	27	Sept.84	6
Operational analysis - revised use of satellite data	39	Sept.87	4
Operational Archive Access facilities	16	Aug. 82	14

METEOROLOGY (cont.)	No*	Newsletter Date	Page
Operational Forecast Suite (EMOS)			
- general description	T1	Feb. 79	6
- convection - parametrisation of	43	Sept.88	6
- data acquisition and decoding	T6	Dec. 79	1
- initialisation	T6	Dec. 79	4
- initial conditions - the spin-up problem	39	Sept.87	7
- quality control	1	Feb. 80	3
- bulletin corrections (CORBUL)	2	Apr. 80	1
- archiving	3	June 80	4
- post processing	4	Aug. 80	3
- significant change made	12	Dec. 81	3
Pseudo "satellite picture" presentation of model results	1	Feb. 80	2
Skill forecasting - experimental system	40	Dec. 87	7
Spectral model	7	Feb. 81	4
- development of	15	June 82	1
- as new operational model	20	Apr. 83	1
- Gaussian grid and land-sea mask used	21	June 83	8
- increased resolution - studies of	38	June 87	10
- parameterisation of gravity wave drag	35	Sept.86	10
- surface and sub-surface scheme revised	38	June 87	3
- T106 high resolution version	31	Sept.85	3
- vertical resolution increased from 16 to 19 levels	34	June 86	9
Systematic errors - investigation of, by relaxation experiments	31	Sept.85	9

* * * * * * * * * * * *

* T indicates the original Technical Newsletter series

USEFUL NAMES AND 'PHONE NUMBERS WITHIN ECMWF

			Room*	Ext.**
Director	-	Lennart Bengtsson	OB 202	2001
Acting Head of Operations Department	-	Geerd-R. Hoffmann	OB 010A	2050
ADVISORY OFFICE - Open 9-12, 14-17 daily			CB Hall	2801
Other methods of quick contact:	-	Telex (No. 847908) (Bleeper 139)		
	-	Telefax (No. 869450)		
	-	COMFILE (See Bulletin B1.5/1)		
	-	VMS MAIL adressed to ADVISORY		
Contact scientific and technical staff via VMS MAIL, addressed to surname				
REGISTRATION				
Project Identifiers	-	Pam Prior	OB 225	2384
User Identifiers	-	Tape Librarian	CB Hall	2315
COMPUTER OPERATIONS				
Console	-	Shift Leaders	CB Hall	3333
Reception Counter) Tape Requests)	-	Tape Librarian	CB Hall	2315
Terminal Queries	-	Norman Wiggins	CB 028	2308
Operations Section Head	-	Peter Gray	CB 023	2370
Telecoms Fault Reporting	-	Michael O'Brien	CB 035	2306
DOCUMENTATION	-	Pam Prior	OB 225	2384
Distribution	-	Els Kooij-Connally	OB 316	2752
LIBRARIES (ECLIB, NAG, CERN, etc.)	-	John Greenaway	OB 226	2385
METEOROLOGICAL DIVISION				
Division Head	-	Horst Böttger	OB 007	2060
Applications Section Head	-	Rex Gibson	OB 101	2400
Operations Section Head	-	Bernard Strauss	OB 004	2420
Meteorological Analysts	-	Taskin Tuna	OB 005	2424
	-	Alan Radford	OB 002	2421
	-	Alex Rubli	OB 003	2425
Meteorological Operations Room	-		CB Hall	2426/2427
COMPUTER DIVISION				
Division Head	-	Geerd Hoffmann	OB 009A	2050
Operating Systems Section Head	-	Claus Hilberg	CB 018	2350
User Support Section Head	-	Andrew Lea	OB 227	2380
GRAPHICS GROUP				
Group Leader	-	Jens Daabeck	OB 016	2375
RESEARCH DEPARTMENT				
Head of Research Department	-	David Burridge	OB 119A	2005
Computer Co-ordinator	-	David Dent	OB 123	2702

* CB - Computer Block
OB - Office Block

** The ECMWF telephone number is READING (0734) 876000, international +44 734 876000, or direct dial to (0734) 499 + last three digits of individual extension number.