GENERAL DESCRIPTION AND ORGANISATION OF ECMWF'S PLANNED FORECASTING SYSTEM

BY
R. NEWSON
EUROPEAN CENTRE FOR MEDIUM RANGE WEATHER FORECASTS

1. Introduction

The objective of the European Centre for Medium Range Weather Forecasts (ECMWF) to which topmost priority has been given is the development of numerical models of the atmosphere for the preparation of medium-range weather forecasts, the collection of data necessary to prepare such forecasts on a daily basis, and the distribution of resulting analyses and forecasts to Member States. Accordingly all the effort in the research field at the Centre has been concentrated towards the development of the appropriate numerical models; all the effort in the operations section of the Centre has been devoted towards the implementation of the necessary computing system and the programmes for operationally collecting and handling the input data to and output data from the numerical models. As a result of this concentrated effort, ECMWF expects to be able to begin its operational forecasting activity in the second half of 1979.

The time scale implied by the term 'medium-range forecasting' is taken to be 4-10 days. However, the data flow in ECMWF's system and the logical functions to be fulfilled are very similar to those used in short-range numerical prediction (i.e. days 1-3). Thus the design of ECMWF's overall forecasting system and the flow of meteorological information is like that conventionally used in any other numerically oriented national centre. The sort of operational system required is summarised diagrammatically in Fig. 1.

Although the design of the system planned to be used operationally at ECMWF is much like that of any other meteorological centre, there obviously are constraints imposed on the system by the time scale of the forecasts that ECMWF plans to carry out. These result from the

strong relationship between scale and predictability; for discussion see page 370 of "The Presentation of ECMWF Numerical Products to Member States" (hereafter referred to as B), concerned with the availability and presentation of forecast results to users in Member States.
Basically, the longer the period of the forecast then the more important becomes the use of sophisticated numerical models, and the larger the area of influence, and the domain of the analysis and forecast become.
(e.g. Bengtsson (1976)). Thus in general terms the guidelines used in formulating the ECMWF forecast system are:

(i) data acquisition - obtain maximum input from WWW network over the entire global domain as far as possible, troposphere and stratosphere-cut-off time 8-12 hours from observation time.

(ii) pre-processing and analysis - refined data control, detailed, global analysis, consistent in three dimensions;

(iii) forecast - sophisticated primitive equation model with high vertical and horizontal resolution over the entire global domain; careful parameterization of sub-grid scale processes including radiation, sea surface conditions,etc., and possible air-sea interaction;

(iv) post-processing - transformation from model co-ordinate system to standard form;

(v) dissemination - transmission of forecasts in coded digital form to users in Member States; graphical presentation of results for local use.

Another general feature of ECMWF's forecasting system is that it is designed basically to run as an automatic system and to produce results in all circumstances. However, there will be a human monitor on duty during the operational cycle; his principal duties will be to check quality and coverage of input data, and the realism (in the most general terms) of ECMWF's analyses and forecasts. The impact of the human monitor on the overall design and organisation of the forecasting system is indicated as appropriate in the description of the various parts of the system. A more detailed description of the various steps making up the complete forecasting system is given below and how these are being built around the computing facilities available within ECMWF.

2. The computing facilities available within ECMWF

The dominating factor in the choice of the computer is the forecast cycle itself, in which the same computations have to be made at every level and every grid point of the model at every time step; in ECMWF's first model this is likely to be of the order of nearly half a million computations at every time step (i.e. 15 levels, 30,000 grid points).

During the examination of the possible computer systems, it quickly became apparent to ECMWF that one single computer could not meet all the requirements, and therefore a system comprising two computers that could be linked was sought. This was because of the radically different nature of the elements making up the complete operational process. For example, whereas the forecast process requires intensive computing power, and only rudimentary file-handling processes, such aspects as pre- and post-processing require computer systems with well-developed and reliable file-handling techniques and software. This led to the conception of a hierarchical computer system - a main computer which would deal with the great burden of the forecast computation, front-ended by another computer dealing with general file handling, (including the interface to telecommunications) and the control of the whole cycle. In this system, the front-end computer would have to be of proven reliability to sustain data integrity. However, provided an ample link for transferring data between the two systems existed, the software of the main computer need not necessarily be as well developed as that of the front-end.

In the event, a system comprising a CRAY-1 main computer, front-ended by a CDC CYBER 175 was selected as most nearly meeting ECMWF's requirements.

The final planned configuration of the complete ECMWF computer system is illustrated in Fig. 2. The CRAY-1 will have one million 64-bit words of memory with SECDED (single error correction, double error detection), 24 I/O channels, 4 disc control units, and 8 disc storage units. It has a floating point computation rate of 80 million additions per second and 25 million multiplications per second and vector programming is also available. The CDC CYBER 175 will have a central memory of 196K 60-bit words with 20 peripheral processors and 20 I/O channels, 4 disc controllers, 8 magnetic tape units, 2 card readers, 2 line printers and 1 card punch. A vital part of the complete system is the CRAY/CYBER link, and a major effort is being made to develop both the necessary hardware and software. The link between the central processors should be capable of sustaining a transfer rate of 5.10^6 bits/second.

The CYBER 175 will itself in turn be front-ended by and linked to sub-systems handling the telecommunications network with Member States, the local network of terminals, and graphics.

The overall demand on computing resources by the complete operational suite basically determined the configuration, and computer power of different elements of the computer system. However, as indicated above, the system is not tailor-made for the planned ECMWF operational forecasting process. Thus, although they determined the basic computer system, the sub-processes in the forecasting system must be built around the computer facilities and resources actually available.

3. Detailed description of the various steps in the overall forecasting system

As indicated in 1, the forecasting system can be broken down into a number of logical sub-processes, viz. data acquisition, pre-processing, analysis, forecast, post-processing and dissemination of results. Additional to this there are such aspects as the real time operational supervision, control and scheduling of all the programmes in the system, and the graphical display of results. Basically, these sub-processes have been divided into two categories for implementation in the ECMWF forecasting system. The first of these categories is made up of the analysis and forecast which need to take advantage of the computing power of the CRAY-1. The second category includes the remaining sub-processes (data acquisition, pre-processing, post-processing, graphics, overall control of the system) are more involved with data manipulation, file handling, and maintaining data integrity. These processes are naturally, therefore, programmed to run on the CDC CYBER 175 system to take advantage of the much more proven reliability and sophistication of the software on this system. The specifications for the link will allow the appropriate data transfer between the CYBER and CRAY for input to the analysis on the CRAY, and also, more critically in terms of data volume, the transfer of the products generated during the forecast period from the forecast model on the CRAY-1 to the CYBER for post-processing and dissemination. These are the general considerations and basic design strategy. The remaining sections of this paragraph describe in more detail the various sub-processes (especially the analysis and forecast) concentrating particularly on those meteorological considerations which are likely to have an impact on ECMWF's forecast results, and of which potential users should be aware.

3.1 Data acquisition

A supply of meteorological data is the basis of operational forecasting, and the Centre will obtain the data which is transmitted on the Global Telecommunications System (GTS) in the WMO coded format, via a link from the Bracknell Regional Telecommunication Hub. Basically this implies that the Centre should receive as input for its global analysis, all the data that is listed for global exchange over the GTS. The Centre will additionally have available a good deal of Region VI (European) data that is regionally exchanged. Also some regional information from the Eastern part of Region VI (North America), and from the Northern littoral of Region I (Africa) are available in Western Europe. However, this relatively limited availability of data is a cause of some concern. The present WWW and FGGE monitoring activities show that the achieved data coverage is very unsatisfactory and the amount of southern hemisphere data available is very limited, and would probably not give a satisfactory basis for a southern hemisphere analysis. However, tests with sets of data assembled in non-real time (e.g. the Data Systems Test (DST) Sets * collected by NMC Washington) have shown that satellite data can give satisfactory analyses in the southern hemisphere.

*These sets basically comprise all available observations transmitted over the GTS to NMC Washington combined with a considerable amount of non-operationally available satellite and other data.

3.2 Preprocessing

The preprocessing subsystem includes the decoding, checking and quality control of the meteorological data received, the establishment of a data base and subsequent extraction of data for the analysis scheme. As the meteorological data is received in WMO bulletin format, it has to be carefully checked and decoded.
The basic methods used in checking the data are given below:

1. Surface (SYNOP,SHIP,SHRED,AUTO-SYNOP)
   (i) Climatological limit check for pressure, wind direction, wind speed air temperature, dew point, sea-surface temperature.

   (ii) Internal consistency check involving wind direction, wind speed, air temperature, dewpoint, pressure tendency, visibility, present weather, and all cloud data.

2. Air (CODAR,AIREP,COLBA)

   (i) Climatological limit check for pressure, height and wind direction.

   (ii) Climatological limit check for temperature and wind speed, the limits depending on pressure or height.

3. SATOB.

   (i) Climatological limit check for surface temperature, cloud height, pressure, wind direction.

   (ii) Climatological limit check for temperature and wind speed, the limits depending on pressure.

4. Sea (DRIBU,BATHY,TESAC)

(i) Climatological limit check for surface pressure, wind direction, air temperature, sea-surface temperature.

5. TEMP, TEMP-SHIP, TEMPDROP.

(i) Climatological limit check for surface pressure, surface temperature, surface wind speed.

(ii) Climatological limit check for upper air temperature and wind speed, the limits depending on pressure or height.

(iii) Lapse rate check of standard level and significant level temperatures.

(iv) Standard level data are re-computed from significant level data and compared with the standard level data received. Substitutions provided for missing or erroneous data.

(v) Hydrostatic check on standard level data. Substitutions provided for missing or erroneous data.

(vi) Climatological limit check for wind speed shear and directional shear between standard levels, the limits depending on pressure.

6. PILOT, PILOT-SHIP

(i) Climatological limit check for upper air wind speed, the limits depending on pressure or height.

(ii) Climatological limit check for wind speed shear and directional shear between standard levels, the limits depending on pressure.
The method used is similar to the same tests for TEMP.

7. SATEM

(i) Climatological limit check for surface temperature, pressure of cloud top, cloud cover, tropopause pressure, base pressure, reference pressure.

(ii) Climatological limit check for tropopause temperature, the limits depending on tropopause pressure.

(iii) Climatological limit check for mean temperature and precipitable water between 2 pressure levels, the limits depending on base and reference pressure.

(iv) Climatological limit check for thickness between 2 standard levels, the limits depending on base and reference pressure.

In addition to these checks, each observation undergoes checks for valid time and if applicable a test for the latitude/longitude/quadrant/Marsden square consistency.

As a result of the quality control tests, various levels of estimates of quality (on a scale from 1 - 4) will be ascribed to an observation or report as a whole, and the individual parameters making up the observation. These estimates of quality will be passed with the general observational input data to the analysis scheme, which will take account of the indicated quality in its selection and use of data. The estimates will also be used by the human monitor on duty during the operational cycle at ECMWF in examination of the general quality of the data, and to see if manual corrections of the data should be attempted. After decoding, checking and quality control the observational data, with its quality flags, is placed into a "reports data base".

The general plan of operation with regard to the data acquisition and pre-processing is that the observational data will be received continuously throughout the 24 hour period, decoded, checked, quality controlled and added in like manner continuously to the data base. However, the analysis (especially for the initial data to be used for the current forecast) will be run at the latest possible moment so that as much data as possible will have arrived. To facilitate this, the structure of the data base has been designed to permit construction of the input files for the analysis in as short a time as possible. The data base structure (and access routines) will also allow ready retrieval of all observations of a particular code type, over a particular area etc., as well as observations with particular estimates of quality as set by the quality control routines which is of importance in the manual monitoring of data. The present plan is to start the forecast at around 21-22Z, using basically 12Z as initial data modified to some extent by the rather limited 18Z analysis that can be carried out at 21 or 22Z.

3.3 Analysis

The term 'analysis' used so far really encompasses data assimilation, analysis of wind, geopotential and humidity fields, and initialisation.

The data assimilation scheme developed for ECMWF's planned operational system is an intermittent scheme where data will be assimilated in 6 hour intervals; observations performed within three hours either side of each six hours (i.e. 00, 06, 12, 18Z) will be assimilated and used for the analysis of that main hour.

Analysis of wind and geopotential will be performed on the basis of observed minus predicted differences of wind and absolute and relative geopotential using a multivariate three-dimensional method. Humidity will be analysed by a three-dimensional optimum interpolation procedure and sea surface temperature by a similar two-dimensional scheme. Initialisation will be carried out using a non-linear normal mode initialisation in order to eliminate gravity waves.

The data assimilation cycle is completed by making use of the prediction model (described in 3.4) to make a six-hour forecast to the next analysis time, when the whole process is repeated. In this way observations will be assimilated for the whole day for the operational prediction scheme; for a schematic illustration see Figure 3.

3.3.1 Analysis method

Basically the analysis method is an extension of optimum interpolation to a multivariate three-dimensional form. The technique enables consistent use to be made of observations with complex error characteristics (e.g. satellite soundings) as well as consistent use of observations which have large variations in their spatial distribution (e.g. air reports). The cross covariances are not obtained from empirical studies due to insufficient data material, but are calculated from height-height covariances (based on simplified empirical structure functions) by the aid of the geostrophic and the hydrostatic relationship. Because of the various assumptions made in the modelling of the error covariances, the interpolation is not truly optimal and the name "Statistical Interpolation" is used within ECMWF and in summary consists basically of using a three-dimensional, multivariate, statistical interpolation of normalised deviations from a predicted field, where the normalisation factors are the estimated root mean square errors of the predicted values. For full details of the planned ECMWF analysis and data assimilation techniques, see Bengtsson (1976), Bengtsson (1978), Larsen et al (1977), Rutherford (1976).

For practical application in ECMWF's forecasting system, the analysis area will be global and it is planned to exploit all input data (including synops, radio sondes, pilots, air reports, satellite wind and temperature soundings, etc.) available operationally that are passed from the reports data base by the pre-processing system. Data checking will be performed on the same basis as the analysis scheme, and an interpolation error calculated (in three dimensions) at the point of observation, by using surrounding observation data. At this stage, the estimates of quality set in the pre-processing phase are also taken into account. When the deviation at this point is larger than, say, three times the interpolated error, the data is not considered by the analysis scheme, and flagged for examination - these flags are at a later stage added to the version of the report in the reports data base.

Statistical interpolation is very demanding of computer resources, and this is why the analysis process will be run on the CRAY-1. The method requires more computations than many other interpolation methods, since the solution of a set of simultaneous equations is necessary for each value interpolated. The order of the system is equal to the number of observational data used. Most schemes try to reduce the number of equations if possible by a careful selection of influencing data. The present scheme takes advantage of the fact that vector-processing computers such as the CRAY-1 can solve larger systems without greatly increasing the computation time. Thus less time in the program need be spent selecting observations since most observations near the area being analysed are used. Such a simple selection scheme has the added advantage that the choice is much less dependent on the precise position being analysed, and the same selection of observations may be used for all analysis points within an area.

Therefore, since the number of data chosen for one analysis matrix can be quite large, and since the same choice of data is to be used for both height and wind analyses over quite a large analysis volume, the selection algorithm used does not evaluate and select observations individually. Instead the globe is divided conceptually into boxes of such a size that when analysing for a box in a typical area of relatively high observation density only data from the nearest neighbouring boxes are needed. This box size has tentatively been set to (6^{o}x6^{o}), modified at high latitudes.

A pre-analysis program will read the observations box by box, and detect situations of high observational redundancy. Such groups of close data values (taking account of the confidence flags set by the quality control) are then compressed into "super observations" of improved accuracy. The pre-analysis programme then selects sufficient "primary" observations to represent the box, when analysing each of its neighbouring boxes. The selection is thus dependent on the amount of types of data available, on proximity of the neighbouring box, and on the geographical position of the analysis area. In this way it is possible to select and flag observations in a neighbouring box, effectively setting up a minimum influence distance within which all data are used. This ensures the necessary overlapping of data to avoid discontinuities at box boundaries. Any remaining "secondary" observations are used only to define the small scales when analysing the box itself. The analysis program treats the analysis points within one box at a time, using its primary and secondary observations plus the relevant primary observations from its neighbouring boxes. If this is not sufficient then observations from the neighbour's neighbours are used, and so on. In this way data selection is extended in data sparse directions when necessary. Overall each complete analysis cycle is expected to take about 20 minutes to half-an-hour on the CRAY-1.

3.3.2 Normal mode initialisation

Once an analysed field has been obtained, initialisation of the field is necessary to eliminate unwanted oscillations resulting from inconsistencies between the height (or temperature) and wind fields. The approach planned at ECMWF exploits the normal modes of the forecast model, and is based on the solution of the model equations linearised about a simple basic state. For full details see Temperton (1977). The initialisation step will take only a minute or two on the CRAY-1 computer.

3.4 Forecast

The development of several versions of possible forecast models to be used operationally at ECMWF has reached an advanced stage. The general design of the forecast model has been carried out in a very flexible way, having been separated logically into two parts - one part describing the adiabatic or "dynamical" processes, the other part the parameterization of sub-grid scale processes. The numerical method of solving the basic adiabatic equations can therefore be fundamentally changed very easily in the ECMWF system (e.g. grid point to spectral, explicit timestep schemes to semi-implicit) without affecting the parameterization part of the model. Likewise the parameterization of the sub-grid scale processes can be modified without changing the adiabatic part of the forecasting system. The horizontal and vertical resolution is also fully flexible. This flexibility of design is expected to be a great asset in the first year or two of operational forecasting at ECMWF, when quite fundamental changes may have to be introduced in the light of experience. For the first phase of operation, it is intended to use a horizontal resolution of 1.5^{o} latitude/longitude and a vertical resolution of 15 levels. Although the models have been designed for medium-range weather forecasting, they can also well be used for short-range weather forecasting, and, with some minor extensions and modifications, for simulation of climate and climate sensitivity studies. Specific details of possible interest in the formulation of the model are as follows :

(i) The vertical co-ordinate system is the now virtually standard sigma system (i.e. pressure normalised with respect to surface pressure).

(ii) The grid point version of the model uses latitude and longitude as the horizontal co-ordinates. The grid system is a regular latitude/longitude grid, and the variables are staggered in space. On this horizontal and vertical grid, the standard primitive equations (i.e. the equations of motion, continuity, thermodynamics and continuity for water vapour) are solved. Basically a second order centred spatial differencing scheme is used with the following properties (subject only to time truncation, and non-adiabatic effects):

(a) the total mass of the model is conserved.

(b) kinetic energy, moisture, total potential energy as well as mass weighted squares of humidity and moisture are individually conserved under advective processes.

(c) potential vorticity and potential absolute enstrophy are conserved by the horizontal advection terms.

(d) total energy is conserved with a correct conversion between kinetic and total potential energy.

Other beneficial features resulting from choice of the grid are the good dispersion properties for inertia-gravity waves giving an accurate simulation of the geostrophic adjustment process, and the avoidance of two grid length computational noise in the terms governing the motion of the pure gravity waves.

(iii) The present timestepping scheme is based on the leapfrog method for the dynamics and a forward timestep for the remaining terms. A linear time-filtering operator is also used to suppress growth of the spurious computational mode associated with the leapfrog scheme.

(iv) As a result of the convergence of the meridians towards the poles, the horizontal differencing schemes are modified in the polar regions, and there is a polar grid point. Further in order to overcome the severe stability restriction arising from the convergence of the meridians, a latitudinally dependent spatial filtering operator is applied either selectively or to all the terms in the finite difference equations (based on damping or chopping of Fourier modes).

(v) As well as the explicit timestep scheme mentioned in (iii) above, a semi-implicit method has also been developed for use in conjunction with the finite differencing scheme. This semi-implicit solution is highly advantageous in that the timesteps 6 - 9 times the length of the explicit can be taken.

(vi) An alternative spectral form of the model also exists. In the formulation of this, vorticity and divergence equations are used instead of the equations of motion in the original form. The vertical resolution is identical to the grid point model ; the horizontal resolution consists of a triangular truncation, but a rhomboidal truncation could be readily implemented. The time integration is semi-implicit where the linear gravity wave terms from both the vertical advection and conversion terms are treated implicitly. The vertical mean temperature profile can be chosen arbitrarily. With the programming techniques employed, a resolution of triangular trucnation 80 woud be possible on the CRAY-1.

(vii The present version of the Centre's parameterization scheme has only recently been extensively tested in practical predictions, and its behaviour is still being assessed; a number of minor modifications may be necessary. The model includes parameterizations of all the physical effects likely to be of importance in medium-range forecasting. In brief :

(a) The parameterization of radiation processes (solar and terrestial) is carried out with the emphasis on effects with short time scales. Cloud/aerosol effects (including scattering) are therefore given higher priority than gaseous absorption (the opposite is usually the case in climate simulation). The radiation program produces fluxes on the three dimensional grid

(a) continued

with a time frequency necessary to calculate the diurnal cycle (probably every third or fourth hour). The radiation programs, like the rest of the parameterization programs, have been designed in a mathematically simple way to allow maximum vectorisation to take advantage of the CRAY-1 power.

(b) The method of calculating the surface fluxes depends on the static stability of the atmosphere near the ground. It is based on the Monin-Obukov similarity theory where profiles of wind and temperature near the ground depend only on the height, the surface fluxes of momentum and heat, and the exchange coefficient.

(c) Above the surface layer the fluxes are calculated by means of the mixing length theory. The fluxes are calculated for momentum, moisture and heat and they extend through the whole atmosphere.

(d) For horizontal diffusion, experiments are presently being carried out with horizontal diffusion types of the form $k_o|\nabla^n\alpha|\nabla^m\alpha$ or $k_o|\nabla^2\zeta|\nabla^m\alpha$ where k_o is a constant, ζ is the relative vorticity, n and m integers, α the diffused variable. This kind of horizontal diffusion can mainly be regarded as a smoothing operator necessary to eliminate computational noise. The larger m and n, the smaller the part of the small scale spectrum which is influenced. 10-day integrations with a 2^o horizontal resolution suggests that a fourth order scheme is a good choice.

(e) For dry convection a dry adiabatic adjustment is done by the diffusion type scheme described in (c). Adjustment by the diffusion scheme also includes a corresponding adjustment of moisture and momentum.

(f) For moist convection, the scheme presently being used is the one designed by Kuo for the parameterization of deep convection in tropical regions. It has also been successfully tested in extra-tropical regions. In Kuo's scheme the convection processes are carried out by cumulus clouds which are forced by the mean low-level convergence and the processes only take place in regions of conditionally unstable stratification. Kuo's original scheme has been modified to consider not only convective clouds originated at upper levels where moisture convergence is observed. Additionally the added moisture is not considered for the whole grid column but only the contribution which takes place in the layer between the lifting level and the top of the cloud.

(g) Large scale condensation and rainfall is predicted in the conventional way, but both large scale and convective rainfall may be evaporated within the model.

(h) The model also computes heat and moisture fluxes at the surface and these (together with the radiative fluxes) are used to predict the change of the land surface temperature. The hydrology of the land surface is also predicted.

The above only gives the main characteristics of the forecast model likely to be used by ECMWF in its operational forecast system in the most general and brief terms. All these features however will naturally have an effect on the quality and realism of the numerical results obtained by ECMWF, and it is important, therefore, to be aware of these characteristics when the results are used. For detailed descriptions of the numerical models, and the representation of adiabatic and physical processes, see Burridge and Haseler (1977), ECMWF Seminar proceedings (1977), and Bengtsson (1978).

So far several 10-day forecast experiments have been carried out, and in particular an extensive series of experiments is in progress on the CRAY-1, using various versions of the model and different parameterizations of the physical processes. Assessment of the results of these experiments will lead to a decision as to the version of the model and physical processes that should be used for the first phase of operations. From the computer point of view, the forecast will start from an analysed and initialised field, for 12 or 18Z for the current day and will then continue forward until the 10-day forecast is completed. Each 6 hours of forecast time, a file containing all the predicted products will be produced and passed to the post-processing phase(3.5). With the anticipated resolution of of $1\frac{1}{2}^{o}$ latitude/longitude, and 15 levels in the vertical for the first operational model, and using a semi-implicit timestep scheme, one forecast day will take 30 - 40 minutes of CRAY-1 time.

3.5 Post-processing and dissemination of forecast products

Post-processing is the part of operational process which takes the fields produced by the analysis and forecast, carries out any subsequent re-organisation or processing of the data, stores them in a data base, and makes the data available in the appropriate format to users in Member

States, to the human monitor locally and for longer term archiving. There is a considerable amount of data to be dealt with in this phase of the operation. For, basically, each day, operationally, there will be analyses for each of the four main hours (18Z on the previous day, 00Z,06Z, 12Z and possibly 18Z on the current day), and forecasts at six-hourly intervals up to 10 days. Each analysis or forecast history file will contain fields of up to 6 parameters on 15 levels in the vertical, and also 10 - 15 two-dimensional fields such as surface temperature, rainfalls, etc., thus in total, for each operational run, something like 4500 fields of data will be produced. It is stressed at this point that many of these fields, especially for periods later in the forecast, will be of little meteorological interest to practical forecasters, and the exact subset and range of these data fields that will be generally available has not yet been fully decided. This whole question is discussed in much greater detail in B, page 370.

With such a volume of data to be handled, and a variety of uses to be satisfied, the post-processing involves an intricate sub-structure of interlinked processing tasks. Bearing this in mind, and the need to be able to present a standard relatively constant interface to users in Member States, it seemed essential in the design that at the heart of the post-processing system should be a data base containing all the data fields in a relatively fixed format, and readily accessible for the different uses. Defining such a constant interface should keep this part of the system isolated from the changes of the model type, co-ordinate systems, vertical and horizontal resolutions, which may occur fairly frequently in the first phase of operation.

The post-processing system has thus been designed in the following way :

(i) Analysed or forecast "three-dimensional" data (e.g. dynamical parameters such as wind components, temperature, moisture),are transformed from sigma levels (on which they have been basically computed) to a number of standard pressure levels. "Two-dimensional" data (e.g. surface fields) are stored directly in post-processing data base.

(ii) Analysed or forecast "three-dimensional" data are converted from grid-point form on pressure surfaces to spectral form (probably triangular truncation up to 80). Subset of spectral array (probably T40) stored in post-processing data base for eventual addition to archive of processed data (see also 3.8).

(iii) Analysed or forecast "three-dimensional" data converted back to grid-point form on standard grid (probably $1\frac{1}{2} \times 1\frac{1}{2}^{o}$ latitude/longitude). The data now on standard pressure surfaces and a standard horizontal grid arestored in the post-processing data base, available for dissemination to Member States, etc.

On the face of it this may seem a somewhat circular way of proceeding but there are several advantages. Firstly, construction of the standard interface is defined in a very clear and logical way. Complete transparency of the post-prosessing system is achieved - to changes of model resolution, co-ordinate system or even whether a spectral or grid point model is used. Secondly the spectral inversion and subsequent derivation of a new grid point field provides a meteorologically consistent treatment of the fields and truncation of short wave noise. Thirdly, the processed data to be archived is provided directly in a consistent and compressed manner.

Once the post-processing data base has been constructed, in this way, products can be displayed locally (see 3.6), or disseminated to Member States. For this latter step, some fields of data may be converted to a polar steregraphic form, and stored in a sub-data base containing data in this format. From these data bases, the products requested by Member States are extracted and encoded using a version of a GRID code (either the WMO GRAF type of code or a compact, efficient ECMWF GRID code). The products in this encoded form are then disseminated to users over the telecommunications network linking ECMWF and Member States. For a more detailed description of this aspect, see Söderman, page 389 describing the technical facilities for handling the communications between ECMWF and the Member States.

3.6 Graphical aspects and local display of data and results

As already explained, a human monitor will be on duty as the operational cycle runs at ECMWF. In real time, this human monitor will have the responsibility for evaluation of how that particular day's operational cycle is proceeding and of examination of input data and output results - to check the overall standard of the meteorological data processing and to see that meteorological consistency is maintained. More specifically this will involve :

(i) supervision of the coverage and quality of input data;

(ii) study of the analyses produced to see if all salient features have been included; special action if, for example, areas of input data missing;

(iii) study of the forecast results to observe if realism is maintained.

To enable this work to be carried out, several facilities will be included in the general ECMWF forecasting system. For example, there will be programmes for the display and correction of input observational data that is too corrupt for automatic handling. There will be a range of statistical programmes, displaying in tabular form the coverage of data in each WMO block, or analysis box. There will be the possibility of calling for display and examination of observations in which errors have been found (this is particularly important for crucial observations in data sparse areas such as ocean weather ships). Also, as each operational run proceeds, first a range of charts of analysed fields, and then forecast fields will be automatically produced on the ECMWF plotting system (i.e. by taking the appropriate data from the post-processing data base, and operating on it with the ECMWF graphical software package). There will also be the facility to request the display on a graphical terminal, or production on the plotting system of any other field of data that the human monitor might wish to examine.

3.7 Supervision, control and scheduling of the operational suite

The whole operational suite is a complex network of programmes and asynchronous data flows. The system will be controlled by an automatic supervisor and scheduler; this is a complex technical problem, involving detailed knowledge of an interfacing with the computer operating system. Basically the overall forecasting system will be controlled from the CYBER. There will be the possibility of manual intervention in the automatic progress of the suite - for example, to run alternative versions of the analysis with different background fields if much input observational data is missing, to re-run steps of the analysis or forecast, etc.

3.8 Archiving of observational and processed data

In the course of ECMWF's work, several important data streams will be handled or generated. In addition to their operational use, some of these data will be accumulated and stored for a shorter or longer time for later reference. In the first phase of operation, the accumulation of data in, and the structure of the archive will be constrained to some extent by the fact that data will have to be stored on high density magnetic tapes, in the absence of a sophisticated data bank facility (e.g.cartridge or mass storage sub-system).

There are two main streams to archive :

(i) raw observational data - all the raw observational data received in real time from the GTS via the link to the Bracknell RTH (see 3.1) will be archived (i.e. theoretically the data listed for global exchange, and that listed for regional exchange within RA VI). However, the data archive is being designed in a flexible way, so that its scope could be extended and that it would be as complete as possible (e.g. by adding observational data from other sources which may not be available in real time on the GTS). The data will be stored in an efficient compact form, and it is expected to accumulate data in this form at the rate of about 3.6×10^7 bits per day. Such an archive of observational data is the cornerstone of much meteorological investigation and extensive use by ECMWF and users in Member States is anticipated. It is worth noting that the data archive will be carefully quality controlled with appropriate flags set (see 3.2); the archive will also contain the flags set by the analysis scheme. Evidently the value of having a uniform checked set of observational data available to users in Member States could be considerable.

(ii) Processed data - the analyses carried out by ECMWF at the main synoptic hours will be stored in the archive as well as a substantial subset of the forecasts produced. Fields of data will be held on standard pressure levels in spectral format (resolution T80 for analysed data, T40 for forecast data); spectral format is used because it provides a highly compact way of archiving data of this sort. However, surface fields, which cannot be suitably represented in this form will be held in grid point form as produced by the analysed or forecast model. Regarding the archive of analysed data, it is noted that ECMWF's operational analysis will be carried out much later after observational time than is customary in national meteorological centres, and therefore may include data arriving later on the GTS. Further they will be global in scope and also, as described in 3.3, use statistical interpolation and assimilation techniques which are especially effective in taking account of asynoptic data or data with different error structures (such as satellite observations). Thus it is anticipated the set of ECMWF's analyses will form a uniquely valuable set for reference by Member States and research activity within ECMWF itself. The spectral format could be of direct use for several applications (e.g. picking out synoptic situations characterised by particular values of the spectral coefficients). However, the data can also be recovered from the archive in grid point format if this is required. The archive of forecast results will be of considerable value to ECMWF itself in long term assessment,evaluation and improvement of the forecasts, as well as to Member States in making use of the forecasts. For example, long series of forecast results are needed to isolate and study typical failure situations, systematic errors, and in assessment of statistical methods of forecasting. Member States are likely to use local statistical methods to exploit the Centre's forecasts, and long series of forecast results for that local area are needed to establish the appropriate model output statistics. The forecast data will be able to be recovered in either spectral or grid point format according to user requirements.

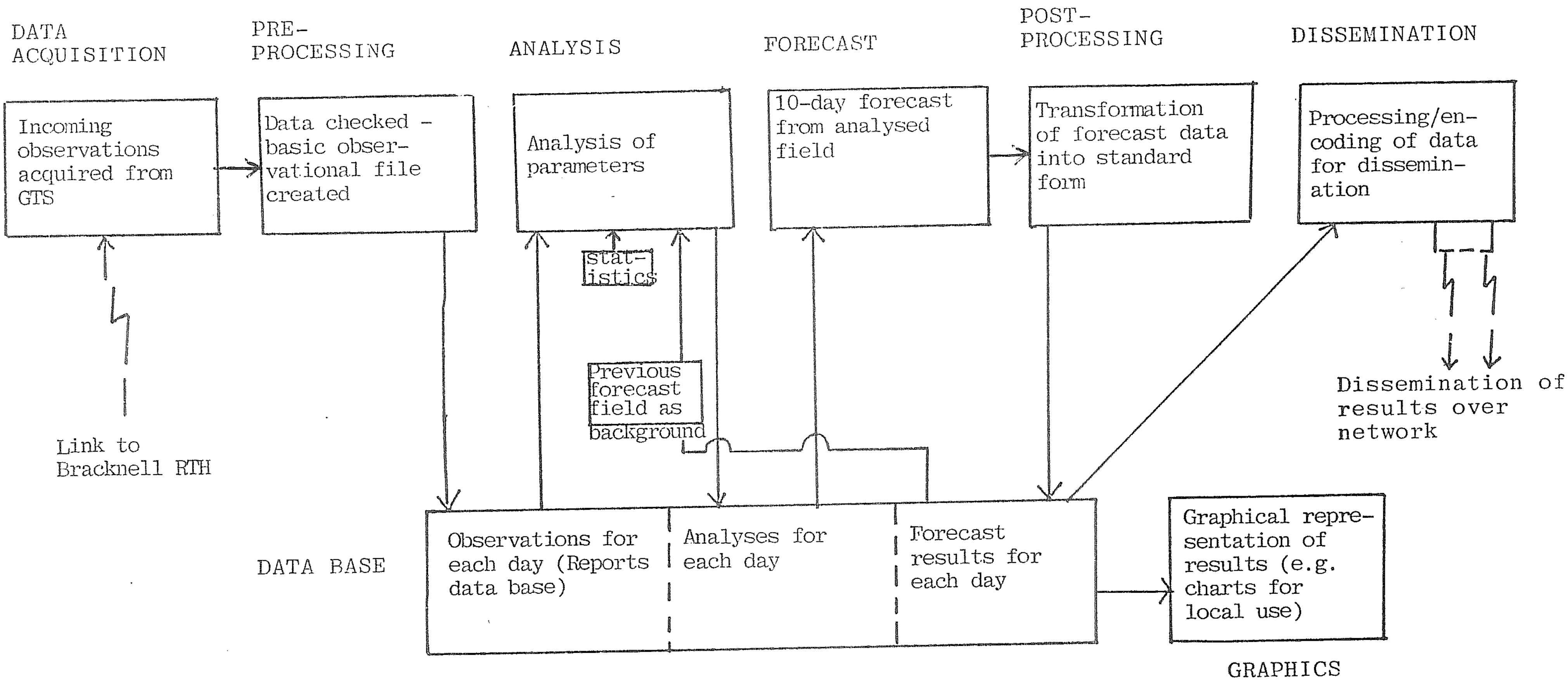

Fig. 1 Diagrammatic representation of operational system

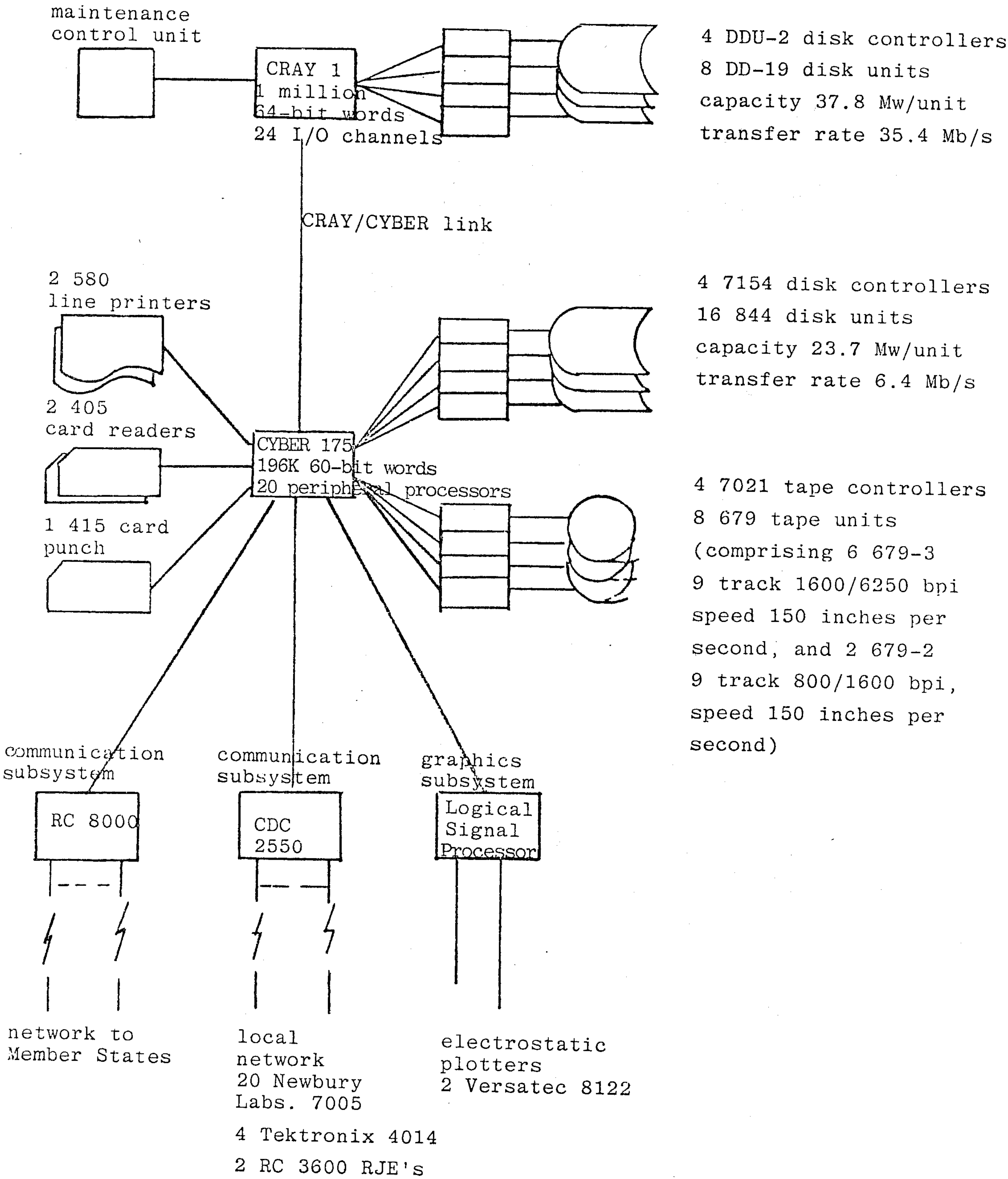

Fig. 2 Configuration diagram of ECMWF Computer System

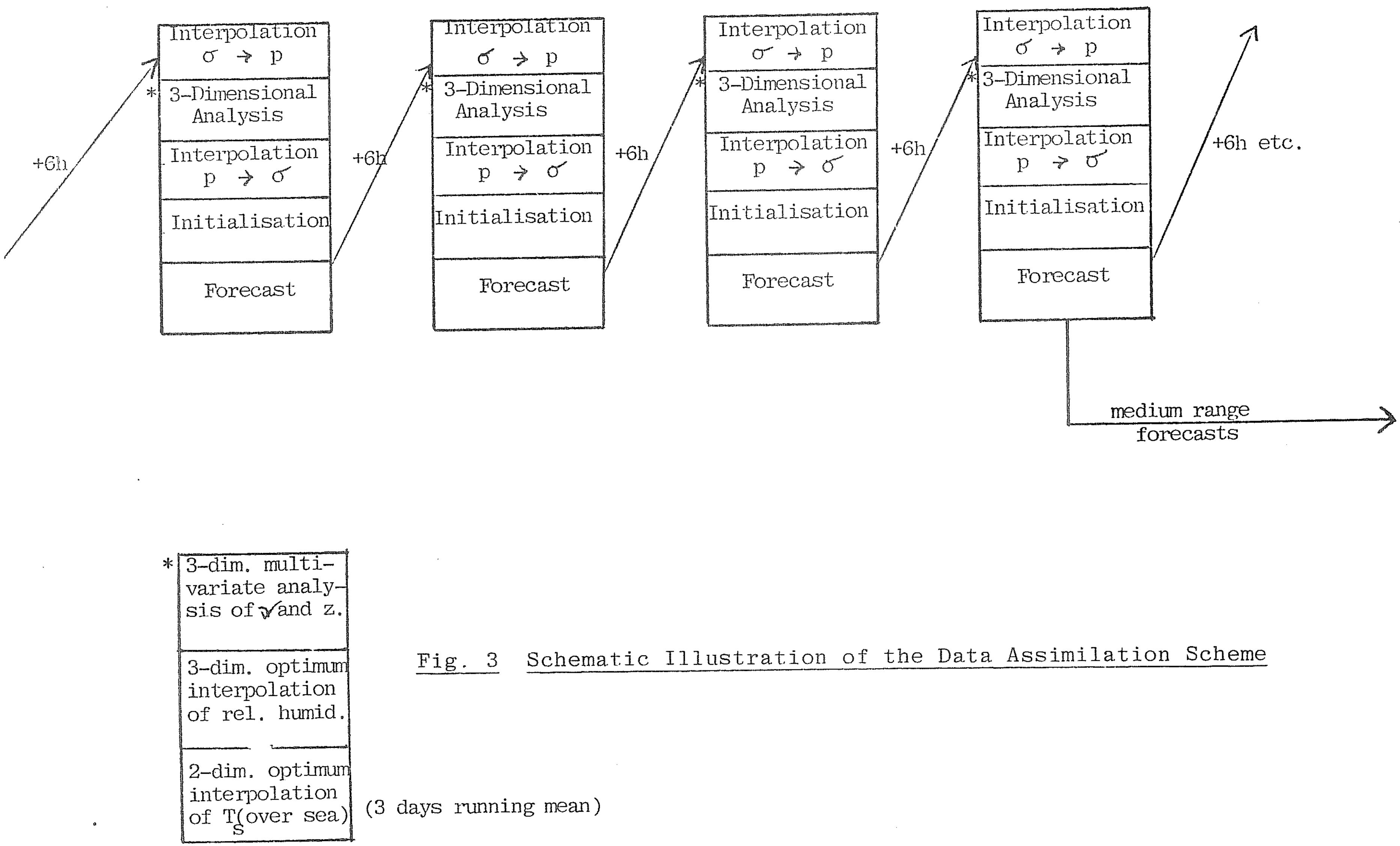

Fig. 3 Schematic Illustration of the Data Assimilation Scheme

References:

Bengtsson, L.	1976	Initial data and some practical aspects of weather forecasting. NCAR Summer Colloquium 1976, Weather Forecasting and Weather Forecasts. Models, Systems and Users, pp 354-419
Bengtsson, L.	1978	ECMWF Global Forecasting System. ECMWF, 1978.
Burridge,D.,and Haseler, J.	1977	A model for medium-range weather forecasting - adiabatic formulation. ECMWF Technical Report No. 4.
ECMWF Seminar	1977	The Parameterization of Physical Processes in the Free Atmosphere. ECMWF Contributions.
Larsen, G.,Rutherford,I., and Lorenc, A.	1977	The ECMWF Analysis and Data Assimilation Scheme - analysis of mass and wind fields. ECMWF Technical Report No. 6.
Rutherford, I.	1976	An operational three-dimensional multi-variate statistical objective analysis scheme. Proceedings JOC Study Conference on 4-dimensional numerical experimentation. Report No. 4, pp. 98-121.
Temperton, C	1977	Normal modes of a barotropic version of the ECMWF grid point model. ECMWF Research Department Internal Report No. 12.