III. WORKSHOP REPORT

1. INTRODUCTION

The Fifth Workshop on Meteorological Operational Systems was held at ECMWF, 13-17 November 1995. The programme and the list of participants are given in the front part of these proceedings.

The objective of the workshop was to review the state of the art of meteorological operational systems and address future trends in the use of medium-range forecast products, data management and meteorological workstations. The workshop was organised under the following main subjects:

1. Use and interpretation of medium-range forecast guidance

   The session addressed the problems and solutions related to the use of numerical guidance in medium-range weather forecasting. With the current developments in numerical modelling, the medium-range forecaster is faced with yet a new challenge: to prepare weather forecasts based on the output from one or several high resolution global models as well as the output from ensemble prediction systems, which can be exploited through appropriate operational procedures to estimate in advance the uncertainty in the numerical forecast guidance.

   Operational centres presented their approaches to medium-range weather forecasting and reported on their experiences with a combined use of output from high resolution deterministic models and from ensemble prediction systems. The concept of probability forecasts and tailored products for certain categories of end users were also addressed and further discussed in a working group.

2. Operational data management systems

   The ever increasing use in meteorological applications of UNIX systems and associated software tools and utilities to provide data storage and access methods means that for users the distinction between databases and archives is becoming unclear. Operational database and archive systems were reviewed showing how use of standard software utilities and languages facilitates the development and portability of such systems. Special attention was given to ease of use and speed of access to data, as well as data transformations performed by retrieval systems on behalf of the users. The current role of commercial databases in meteorological applications was also reviewed.

3. Meteorological UNIX workstation applications

   Current and planned meteorological workstation applications were presented and demonstrated at the exhibition during the workshop.

   De facto standards were used in some areas, e.g. the X-Window System and Motif, whereas elsewhere different solutions were developed. A working group discussed areas where no standardisation had emerged including different user interface paradigms, operational aspects, data access methods, handling of repetitive tasks and 2D/3D presentation techniques.

The reports from the Working Groups are summarised in this section of the proceedings while the papers from the presentations are given in Section IV.

2. REPORT OF THE WORKING GROUP ON USE OF MEDIUM-RANGE FORECAST GUIDANCE

COMBINED USE OF RESULTS FROM DETERMINISTIC MODEL AND EPS PROBABILISTIC FORECASTS

Use of the ensemble forecast

The Working Group discussed a wide range of possible applications of the ensemble forecast, from simply smoothing the ensemble into a single deterministic forecast, to advanced applications with probabilities of single and combined parameters.

Forecast of forecast skill

There is now evidence that small ensemble spread tends to indicate relatively high skill of the control forecast and possibly of the T213. This is supported by various statistics (mainly area average), and also on the synoptic scale by several examples of converging "spaghetti plots".

The positive experience with the ensemble mean was appreciated, but it was noted that the use of the EPS should not be restricted to this.

Forecast of alternative scenarios

In cases with large spread when no particular weather scenario has a strong majority the aim should be to express this forecast uncertainty in terms of alternative scenarios, subjectively, or with probability estimates if possible.

The usefulness of the so-called "spaghetti plots" was discussed. Although it provides a very handy way to display the ensemble, its importance should not be overstretched since the horizontal spread of the isolines depends also on the gradient in the area. This might make important spread in areas of strong gradient hardly visible, whereas in areas with flat gradient and not necessarily interesting areas, there might be large divergences.

In the cases with small ensemble spread and T213 within the ensemble, a traditional deterministic approach is the obvious one, although unpredictable small scale features should still be filtered out. The most difficult cases are those cases when the T213 forecast is outside the ensemble. This is mainly due to

- the differences between the T213 and T63 models, where the latter's difficulties to develop blockings or cut-off lows after D+5 is the main problem. Cases when tropical cyclones moved out in the extra-tropics were also mentioned;

- the limited size of the ensemble, which might cause problems in cases with widespread dynamic activity.

Cases are seen of large day-to-day changes in the main direction of the forecast. In such cases, the synoptic investigation by the forecasters should, if possible, include the output of the EPS from one or two days before as well as the current one. The examination of the patterns of the initial perturbations can also be beneficial in these cases.

A higher resolution model like T106, later to be followed by improvements in the perturbation technique, should overcome most of these problems.

Ensembles based on models with different physics, as is being implemented at CMC, and multi-model ensembles were seen as interesting approaches.

The positive experiences with the object oriented clustering developed in Switzerland was appreciated. In general, object-oriented analysis of flow pattern are promising and should be explored further.

Probabilistic forecasting

This was seen as the real aim of the ensemble forecasting technique. To obtain statistically reliable estimates of probabilities requires that the spread of the ensemble, the frequency of extremes etc. be realistic. In addition, the number of elements should be sufficient, and the resolution of the forecast model high enough so as not to introduce systematic errors. Probabilities should be applied to the forecast of synoptic scenarios and of weather elements.

PPM technique or adaptive filtering should be used for the probabilistic forecasts of weather elements.

Appreciation was expressed over the probabilistic charts of weather parameters implemented during last winter.

The possibility of generating probability products by giving different weights to the control forecast and to the ensemble members was discussed. There is no answer to that problem at the moment, but the issue should be kept under review.

Probabilistic forecasting can also be applied at the application level. Ensemble ship routeing could be tested, provided that initial perturbations are introduced in the southern hemisphere.

Validation of EPS, verification of EPS products

The evaluation of an ensemble prediction system should be done on two levels which complement each other:

- the validation of the system itself, i.e. the assessment of the statistical characteristics of the ensemble;

- the verification of the products, which should cover a wide range of aspects: from the verification of the ensemble mean (in a categorical sense), to the verification of predicted probabilities (in a stochastic sense), including synoptic verification of the predicted flow patterns.

Significant progress has been achieved in these two areas, mainly concerning the ensemble spread and the probability products. An initial cluster verification has been implemented, with promising results. The evaluation of the probability products through reliability diagrams is well established. However, much remains to be done:

- the validation methodology needs to be developed further. At the moment, there is no well established way of comparing ensembles from two different systems. This has serious practical implications, in particular for the testing of changes to the system configuration: change to model formulation, to the initial perturbations, increased ensemble size;

- additional verification measures appropriate to the evaluation of the attributes of the EPS products are needed. The relative operating characteristic (ROC) curve is particularly recommended because it allows for the comparative evaluation of deterministic and probabilistic

forecasts. Also, probabilistic forecasts of weather elements (such as likelihood of precipitation amount above certain thresholds) can be created using past track record of the control forecast, by computing the frequency at which observed rain reached the selected threshold, given different forecast limit values in the forecasts. Probabilistic forecasts from the EPS system should be able to surpass the performance of such a simple technique based on the observed conditional frequency values. Other important aspects to be pursued are the verification of the day-to-day consistency of the ensembles and of the clusters and the reliability of the dependency between small spread and high skill.

Another possible approach to estimating the differential benefits of deterministic and probabilistic products from the EPS, while also placing value on these products directly, is through the application of the decision analytic model. Such models are readily available, but it was noted that each application is unique and so a substantial effort would be required to produce data from a series of applications.

Development of workstation applications

The range of derived products from the EPS that can be produced at ECMWF cannot and is not supposed to cover the variety of requirements from member States users. For example, the best thresholds for probability products vary from country to country, and also from time to time. The required clustering area may depend on the weather situation and the current demands. Ideally, each user should have the possibility to perform its own post-processing according to its requirements.

With the development of workstation applications and the increased telecommunication capacities, the Centre plans to develop METVIEW applications for possible release to interested Member States. The plotting of "spaghetti diagrams" has already been tested (any field, any contour level). Other suitable applications are the computation of probability products and the generation of clusters.

Training

With a new system like the EPS and the need of developing new methodology for the use of the products, extensive training and information on the system is essential. The effort that ECMWF has put into that aspect, through expert meetings, training courses and liaison with Member States, was appreciated and should be continued.

Data access

Users made it clear that they are not interested in whether data bases are accessed through SQL, embedded SQL or calls to C or Fortran language interfaces. What they are interested in is meteorological data and they want to address it in meteorological terms. This requires additional software layers to be provided between the database and the end user. The lowest level should be SQL to maintain the portability of the software and stay independent of any particular database vendor. Performance can be improved by using language interfaces but this normally means using vendor specific features. A majority view expressed was that, if performance was an issue, the solution should be found by acquiring faster hardware and maintaining the SQL level interface.

The WMO binary codes, BURF and GRIB, used for the representation of observations and field data were criticised as being unsuitable formats for many data processing/display applications and were really suitable only for transmission and storage purposes. The BUFR format especially created problems. Different user requirements mean that data stored in BUFR format needs to be retrieved in different formats eg in BUFR format as stored, as integers or as floating point numbers. BUFR is not a working format and individual sites create and use their own working formats.

The question was raised as to whether it might be useful to try and coordinate on standard formats but it was generally felt that there was too great a diversity of end user requirements to make this feasible. Utilities should be provided to access the data bases and other utilities should provide facilities for users to operate on the retrieved data. Specific user utilities should be provided for specific user needs. One system returns BUFR data to users as a structure, but this limits access to C and Fortran 90 programs. It would be nice to have a general query language for meteorological data but it was pointed out that the maintenance of such a language would be extremely difficult due to new data types and new requirements from users.

Research data bases and operational data bases can have very different retrieval patterns and requirements. It is sometimes necessary to have the same data in two different data bases, with different database models, to provide raw retrieval speed for production purposes and detailed relational searches for research users.

In the case of distributed data bases, or data bases accessed through networks (either LAN or WAN) the system needs to have redundancy built in to avoid having a single point of failure. They should be as reliable as when data is accessed and used on the same machine. Client/server mechanisms are very user friendly as users do not have to know where the data is or where it is retrieved from, even if a backup system has been used. WMO's ideas on distributed data bases were discussed and there was complete agreement that any work in this area should be actively encouraged to enable data to be acquired on an ad hoc basis.

Off-line data was also discussed. The point was made that users did not care where the data was stored, but just wanted to access it. However, a strong requirement was expressed for facilities for users to find out if data was archived or on line and to bring archived data on line in advance of the time it was needed for use. It was felt that general users did not really need to have complete control over where their data was archived but that operational data did need archiving in a more controlled fashion. Finally, it was said that we could look forward to new systems which would provide terabyte/petabyte desktop storage, together with the necessary database software and tools to manage the data.

Data manipulation on retrieval

It was pointed out that manipulation of retrieved data should be done in several layers, having the database access at the lowest level and the end-user's software at the highest one. Users with different requirements should provide intermediate layers in order to obtain their data in the preferred format.

It was noted that different units/tables are used for meteorological data in GRIB format. This gives more difficulty in data manipulation for users when the units depend on the originating centres' models. This should be transparent to users. However, it was pointed out that only SI units were legal in the GRIB code and that this was more a GRIB code question than a data manipulation one. The meeting also discussed other shortcomings of the GRIB code and lent its support to the current WMO initiatives to provide a new edition of GRIB which would address many of the present and foreseen requirements. It was underlined that one of the most urgent requirements was for an international GRIB standard for the representation of control and perturbed forecasts from ensemble prediction systems.

When a particular data manipulation is done frequently, a new data type should be created and the stored back into the database to make it available for other users. This obviously applies to routine operational forecast products, but should also be considered for other common requests eg conversion of spherical harmonic format to grid point. However, a balance between storage space and CPU usage must be achieved. Some mechanisms to save CPU time in manipulation are either performing the manipulation at a pre-processing stage, knowing in advance the most common format/units in which users want the data, or applying some cache facility which stores the most recently manipulated data.

The general feeling was that data manipulation is the users' responsibility, although some software tools could be provided.

Commercial data bases

It appears that among the sites represented at the meeting that the most popular commercial data bases in use are Empress and Oracle, with one site having made a port of software from one system to the other. Other commercial data bases mentioned were Sybase, Ingres and Informix.

The different ways that the Empress and Oracle data bases interacted with operating systems and used files were discussed, as were database administration requirements. Each of these had vendor specific features for enhancing performance, but as stated above, most users stressed that portability of software should not be compromised for the sake of performance.

The Neons software package is proving popular and there are versions running with both Oracle and Empress, both of whom appear to be preparing to add object-oriented features to their systems.

At several sites objects and C++ are in use, interfaced with relational data bases. Object-oriented data bases were also discussed. These data bases are very suitable for storing observations, except in the case of climatological data, where relational techniques have an advantage. Object-oriented technology has, as yet, no standards. The Object Management Group (OMG) is working to define standards and all data base vendors have promised to adhere to these standards when they are approved.

4. REPORT OF THE WORKING GROUP ON METEOROLOGICAL WORKSTATIONS

User interface paradigms

Quality and consistency are the primary aims for a good user interface. There should only be one user interface. This should be relatively easy to use so that beginners and occasional users can use it without difficulty. However, there should be macros available to enable the more advanced and/or ambitious users achieve their goals.

Button clicking and pushing should be kept to a minimum. One approach which may help this is to have direct manipulation (i.e. placing the cursor/mouse in the data) rather than indirect.

The need for uniformity and consistency was expressed. Users should find basically what they expect and should not get surprises.

It should be possible to run frequent or everyday tasks without interaction.

Development tools

People should spend time to find the right tools for their application and ensure that the tools are used in the right way.

A number of development tools were mentioned, including:

- Programming languages: C, C++;
- Presentation "look and feel": Motif;
- CASE tools: PURIFY, CVD;
- Development environments: CENTRELINE, SOFTBENCH;
- Configuration and control: CLEARCASE, CONTINUOUS;
- User interface design: X-Designer, UIMX;
- Interface configuration languages: TCL/TK;
- Commercial data bases.

The growth of the World Wide Web (WWW) as a means of exchanging information is acknowledged as extraordinary. The widespread availability of the HTML language and browsers such as Mosaic and Netscape points to the strong possibility of WWW being used as an interactive application itself and as a support for the dissemination of meteorologically related graphics, data and documents.

Product generation

Meteorologists are going through a period of cultural change leading to a new generation of workstations and new tools which might take some time to adjust to. More interactivity on workstations and an increase in screen size should lead to a reduction in paper output.

Product generation is seen as very important in the workstation area for the future. It can be done automatically or with forecaster intervention.

The workstation can also be used to generate automatic text messages which can be used as weather forecasts on radio or in the press. In some cases, it was found that there was no need for forecaster intervention except when there were problems with data, hardware etc. Others felt that the models were

not yet sophisticated enough to produce consistently high quality text messages.

2D/3D presentation techniques

Reduced costs and improvements in techniques make 3D within reach for many establishments. Research people are becoming more interested in using 3D, operational people less so. One big difference recently is the ability to go from 3D to 2D quickly and vice versa. While volume rendering techniques (i.e. transparent fogs) are interesting, they are qualitative and less used by meteorologists than quantitative techniques such as moveable 2D slices.

Aviation users are excited about 3D but meteorologists in general are still entrenched in 2D and need to invest time in 3D to enable them to learn from it.

There is still a need for exploration to exploit 3D, e.g. in the Ensemble Prediction Scheme (EPS).

Video tape presentation of 3D is useful particularly if accompanied by a commentary and is seen as both a research and a presentation tool.

3D usually requires fairly large data sets and this could represent a problem especially if data had to be transferred from a Centre to regional offices. Compression, truncation and indexing should help to reduce the transmission times. Also, the reduction in the cost of lines could offset the increase in data.