III. WORKSHOP REPORT

1. Introduction

The Third Workshop on Meteorological Operational Systems was held at ECMWF, 18-22 November 1991. 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 therefore organised under the following main subjects:

1. Use and predictive skill of numerical forecasts in the medium-range

2. Operational data management systems

3. Meteorological workstation systems.

The first session of the workshop was partly a continuation of the ECMWF Workshop on Predictability which was held at the Centre during the previous week. Especially during the related working group session aspects of using the results from the predictability research in an operational forecast environment were addressed. The other two sessions dealt with data management and processing issues. In presentations and related working groups the participants discussed the impact of standards in computing, data communication, visualisation and formats on the development of effective operational systems to generate, process, monitor and visualise data from global models and the global observing system.

This report summarises the discussions and recommendations from the working groups and the final plenary session. The papers submitted for these proceedings are given in Part IV.

2. Report of the Working Group on Operational Guidance from Monte Carlo Research

To the operational forecaster and the end users, ensemble and probability forecasts will provide information in amounts and of types not previously encountered. It will not only give them some guidance about the deterministic forecast limit of each forecast run, but also open up new possibilities by providing alternative developements. New forms of presentation will be required to provide both the forecasters and the end users with access to essential information available within ensembles.

The following list was considered by the working group to suggest possible products for presentation to the users:

2.1 Guidance in terms of prediction of the skill

This applies only to a scenario where the ensemble is produced in addition to a forecast based on a more sophisticated and/or higher resolution forecast model.

a. Estimate of the RMS error and Anomaly Correlation Coefficient of the forecast every 24 hours up to day 10, for Z 500 and MSLP, with indication of error bars

b. fields of variances (related to the spread): geopotential, MSL pressure, temperature / thickness, to provide indication of the regional variations of the skill

2.2 Guidance in terms of synoptic flow pattern, to present a condensed view of the outcome of the ensemble

a. ensemble mean fields (the induced smoothing will provide forecast fields with supposedly unpredictable scales filtered out): geopotential, MSL pressure, temperature, thickness, upper air wind

b. clustering, i.e. grouping together forecasts which are "similar" according to certain criteria (500 hPa flow pattern, MSL pressure, precipitation etc.) To each cluster a probability must be ascribed.

   As stated at the Workshop on predictability, the best way of producing the clusters is still open. Three user requirements are:

   - the determination of the clusters should be based on the area of interest to the user, implying that clusters for Europe taking e.g. Northern and Southern Europe as reference areas will not be the same;

- the time aspect should be taken into account, i.e. forecasts with phase differences should belong to the same cluster;

- a forecast corresponding to an extreme event should stand out as a cluster on its own.

c. projection of the forecasts into some subspace, i.e. to represent the whole ensemble in a simplified 2-dimensional diagram to help the forecaster identify those ensemble members which are forecasting extreme events during different lead times.

2.3 Guidance in terms of weather parameters

a. probability charts for certain weather parameters and certain thresholds (e.g.the probability of precipitation of more than 5 mm / 12h at D+3).

b. frequency distributions of weather parameters: temperature, precipitation, wind, cloud cover, humidity, for individual locations.

The user must have access to the connections between the different weather parameters within the ensemble (the most probable rain amount will not necessarily coincide with the most probable wind speed, etc...) This can be achieved only by an interactive process, with the possibility for the user to put the emphasis on the parameters of his choice.

The general reaction of the group was that all these directions ought to be explored further. However, little practical benefit seemed to be expected from a prediction of the skill as mentioned under paragraph 1., unless it is given with a very high accuracy. Much more emphasis was put on the potential benefit of synoptic and weather parameters guidance.

At this stage, it appeared difficult to progress any further in the discussion without separating two different types of use of Monte-Carlo products:

a. provision of information for the forecaster: all forecasters present were willing to explore the use of guidance as described under 2. and 3. During the experimentation planned for winter 1992 - 93, forecast charts from up to 6 or 7 clusters could be disseminated daily for that purpose (at low resolution).

b. direct provision of products of category 3. to specific end users (after application of proper procedures for bias removal): this will be an extension of what already exists in many Member States

It was stressed that much more experimentation is needed before more specific product definition can be formulated. In this respect, it is recommended that during the winter 92-93 experimentation, some forecasters from Member States come to ECMWF for a period of about 10 days to work on the evaluation; in that way, they would have access to all the output from every ensemble run. This should take place after 4 to 6 weeks of experimentation.

3. Report of the Working Group on Data Bases for Meteorological Applications

3,1 What are the benefits offered by commercial DBMS systems for the management of:

* data?
* metadata?

Some participants regarded the use of commercial data base management systems for the management of meteorological data and/or metadata as inevitable. However, it was pointed out that commercial systems are designed principally to manage information (facts, knowledge, generally non-numeric data), so the meteorological community should not place their expectations too high.

Several sites represented had already made use of commercial systems to some extent; most of the larger sites had plans either to move to the use of such systems over the next few years, or to carry out extensive studies and/or prototyping with a view to implementation in the medium term. Attention was drawn to the fact that the American AWIPS would use a relational data base, and thus would be a driving force for change in that country. In this context, as a step towards modernising NMC's data bases, cammercial data base management software would be studied in a workstation environment.

The view was expressed that, especially for large implementations, the commercial data base management was perhaps more appropriate for metadata. Commercial data bases could also be used as a dictionary to indicate where data are stored; data bases supporting "binary large objects" or "blobs" are perhaps suitable for this, offering also the ability to manage the storage of the data in standard binary forms.

The following had been quoted as advantages by Pauli Rissanen:

* interactive data availability
* user-friendly updating
* what you see is what you've got
* sometimes fast
* simple data manipulation (sum, mean, s.d., etc.)
* easy to form and update tables
* system support

It was also suggested that it is not sufficient simply to buy a commercial system "off the shelf". The requirements must be fully defined, including expected performance, and the correct advice must be sought; this would probably involve the purchase of consultancy.

3.2 What are the disadvantages?

The following had been quoted by Pauli Rissanen:

* "why don't they use SQL?"
* the need for trained people
* difficulties linking FORTRAN
* sometimes slow
* SQL supports simple computations only
* very large queries jam the system

Some of these disadvantages could be rectified; important factors are the correct use of indexes (tables without indexes could be 30 times slower), avoidance of "stupid" requests, causing all table items to be accessed, optimisation of data retrieval, and attention to archiving strategy.

A number of participants raised the issue of data security, asking if potential users of commercial data base software plan to back up data.

3.3 Should SQL be used directly, or should interfaces be provided which handle the generation of appropriate SQL for the user?

The view was expressed that many users would need a more meteorological interface to data than SQL, whereas others would rapidly learn to use SQL directly. Some site have established interfaces to data used by many applications; supporting the same interfaces but adapting the underlying application to a commercial system would be helpful in terms of implementation.

3.4 What are the Users' Requirements?

Where the user's requirements can be met cheaply and effectively with simple software, there may not be a case for using a commercial package. The "cross over" point at which a commercial package becomes sensible must be assessed. What is clear is that as commercial systems improve, and as computer hardware also improves, the point at which commercial software is cost effective is likely to continue to become within range of an increasing number of users.

Representatives of the suppliers of commercial data base software indicated that it was essential that potential users define fully their requirements and expectations.

The Finnish participant had given some figures relating to expectation:

* 40 retrievals per second (each of 6 parameters plus simple calculations)
* 10 additions per second when loading data into the data base
* extraction of all observations for one hour from the SYNOP table within 4 seconds

Various participants pointed out that the needs differ from user to user. In particular, there are very different requirements for a small national climatological data base than those appropriate to a large centre for near real time applications. There are also factors relating to where the data and data base software will reside - a network of workstations, or a large mainframe, for instance.

The opinion was expressed that a relational model does not necessarily fit meteorological requirements - they can often be better expressed in an object oriented manner; again, the type of application has a bearing on this issue.

Some participants advocated the requirement that data base queries be independent of the data base used, pointing out that this was of importance if users were to be able to develop applications which would not be affected adversely by changes to the data base.

3.5 What are the costs involved in terms of:

* purchase
* maintenance
* support resources

Participants from NOARL and FNOC indicated that 10 to 14 man-years had been invested to date since 1987 at the NOARL level, possibly twice this at the FNOC level.

Approximate figures quoted for one actual implementation were of the order:

* Licences (40 VAX users + 10 - 15 micros)	£93,000
* Training (30 persons)	£45,000
* Consultancy	£66,000
* manpower	5 man years

A participant from a commercial company suggested costs are likely to be of the order of £5,000 to £10,000 for development licences for Sun/Sparc 1 or 2 workstations, and up to £250,000 for a large supercomputer; run-time licences could be expected to be about half these costs, with additionally a percentage for maintenance, and a further percentage for support.

Manpower support could not be expected to be reduced as a result of installing a commercial system; the gains are more likely to be experienced with respect to the possibilities for enhancing applications and service.

3.6 How would the use of a commercial DBMS system interface to an archive system?

3.7 Are atomic data units for on-line access co-incident with those appropriate to an archive?

Questions 6 and 7 were taken together. Two different types of archive were discussed - the migratory type of archive, where appropriate files migrate out, and have to be recalled by the systems software (or other means) when required, and the extended support offered by some commercial data base systems such that they are able to manage the archive and recall of selected tables.

It was pointed out that optical disks with up to 100 gigabytes of storage are now available, but that accessed files must first be staged to magnetic storage. It was also suggested that most UNIX based archive systems can not address data below the level of a file.

Participants noted that, whatever the means of archiving and managing the data, if they were BUFR and/or GRIB they could be accessed directly by non-data base applications if required.

The question of very large amounts of "directory" type information was raised. It would seem that, unless there are efficient means for the archival and recall of directory type information, the file sizes used for archiving purposes should be at least an order of magnitude greater than those often used on-line.

In the final plenary session one participant suggested that some attention be given at the next workshop to the format of data for archives.

Concluding Remarks:

It is believed that this discussion would have taken a rather different course had it taken place 10, 5, or even 2 years ago. There have, in recent years, been continued advances in the performance of computer hardware and software, the general migration towards UNIX style operating systems, and the effectiveness of the commercial data base products available. The consensus of the working group considered the choice of whether, and when to move to a commercial data base management system must depend on a full analysis of the precise needs of each individual centre. Most if not all participants considered it essential to monitor the progress of commercial systems, and to gain experience where feasible. It was further recommended that centres moving towards the implementation of data base management systems should, where possible, move gradually, building on initial experience with small systems before incurring the considerable expense of full scale implementation on large systems.

4. Report of the Working Group on UNIX for Meteorological applications

4.1 What considerations are important when porting Meteorological Applications to a UNIX environment?

Under this topic, the group discussed some of the topics touched on in the earlier presentations. It was agreed that UNIX offers considerable advantages in terms of flexibility and efficiency, but that attention is necessary to adopting, and even enforcing conventions within an operational environment. The suggestion was made that naming conventions might be controlled through common access to a dictionary of names.

Otto Pesonen also elaborated on the guidance material adopted by ECMWF for the migration of the EMOS system; this material will be reproduced as an annex to the corresponding paper in the Workshop Proceedings.

4.2 What are the respective merits of FORTRAN, C, C++, etc.?

C was seen as a more natural language for the development of non-numeric intensive applications, especially those requiring the manipulation of characters. The principle problems envisaged are the lack of C knowledge amongst the older analysts, the potentially cryptic nature of some C programs, the linkage with FORTRAN, and the enforcement of standards, good programming practice, strategy for portability, etc.

It was believed that there will be a rôle for FORTRAN for the foreseeable future, especially for numerically intensive applications such as forecast models.

While, strictly speaking, there may as yet be no international standard for C++, a de facto standard could be seen to exist in that the most widely used compiler emanates from the vendor which has played the leading rôle in the development of UNIX. C++ was seen as a language in its own right; it is useful for object oriented applications, and especially so for image manipulation. Since it is virtually a superset of C, mixed applications containing C and C++ can be compiled together, and can be used for specific functions within the framework of controlling master tasks written in C. It was noted that SMALLTALK and C++ don't talk to each other.

One participant asked what packages are there for enforcing standards? Responding participants quoted as examples Toolpack from NAG, Smart Systems, and Sabre C.

4.3 Is there a standard and portable UNIX based system for programme maintenance?

No, although various potential maintenance systems are planned to be investigated by various participants.

5. Report of the Working Group on Meteorological Workstations

5.1 THE METEOROLOGICAL WORKSTATION

The most important potential benefit of the MWS is the potential for integrating various data sources. The forecaster has access to and sees a large amount of data, which is growing. The MWS should provide a way for the forecaster to assimilate large amounts of data very quickly. The forecaster not only wants to see the raw data but has also a need to manipulate data and perform quality control. When performance is crucial, it is seen as an advantage to precompute weather maps, in a "display-ready" format.

Thus the MWS can centralise the forecaster's tasks. It should also be noted that paper plots will not go away in the foreseeable future. Another key point is that some forecasters may initially be sceptical of the benefits the MWS, but experience shows that once they get used to it, they do not want to work without it.

The working group also realizes that there is a significant difference between operational and research requirements for an MWS.

5.2 GENERAL OBJECTIVES AND REQUIREMENTS

Caution was urged in not being too ambitious in MWS implementation. Hardware and software develop very fast, meaning potentially rapid obsolescence. Also, long implementation times are unrealistic in most weather centres due to resource limitations. Rather a step-by-step approach is recommended, with useful introductions at each stage, and a flexibility to add functionality over a period of time.

The most important performance requirement, is that the MWS must respond **within the timeframe of the decision at hand.** This may mean 1-3 seconds for short term forecasts to less stringent requirements for longer term forecasts and research applications. However, it should also be noted that high performance may carry a cost penalty, such as data duplication. Forecasters will prefer a system with a consistent and predictable response. One of the ways to solve this is to dedicate a processor to each forecaster.

Although overall reliability is important, the highest cost-benefit is achieved by maximising the reliability of **critical capabilities,** which means unavoidable duplication. For such capabilities a 99.99% uptime is realistic, but not for the whole system.

Although interactivity is seen as the basic function of an MWS, batch is still a very important requirement in operational weather centres.

Portability and standards are inextricably linked. Although several written standards exist, many vendors do not always follow them. Therefore, it seems most beneficial to use **accepted standards,** i.e. the lowest common denominator of the computing platforms that are used or under consideration. This is further compounded by a very fast moving market. However, standards may often impose a performance penalty. These two considerations must be balanced.

5.3 PRODUCT PRESENTATION

Stationary 3D is not seen as a useful tool. To be meaningful 3D pictures must be animated (4D). However, today this tool is seen as useful only for teaching purposes to illustrate meterological concepts, such as vorticity and divergence. It is also agreed that it might be useful soon to researchers. However, it may be some time before such tools are mature enough for operational forecasters. To advance these ideas more packages must be made available for experimentation.

5.4 USER INTERFACE

It must be realised that operational forecasters are often not familiar with computers, as opposed to meteorological researchers. Thus the operational forecaster is facing many obstacles in the path to accepting a graphical user interface, such as using a mouse, perceiving multiple windows etc.

The user interface must facilitate the user's tasks. Specifically, it must provide the following:

- as many defaults as possible;
- minimise the number of layers of menus (1-2 maximum);
- inform user that the system is working, if response time is long;
- give status information at the highest level of the user interface;
- non-overlapping static windows are preferred to avoid screen cluttering.

The "desktop metaphor" is a useful comparison, i.e. the cluttered desk vs the organised desk.

Ergonomic studies may also provide insight into the productivity of users. This involves user interface considerations such as the choice and number of colours used, style guide etc.

It is desirable that the user be able to tailor the user interface to his particular work and particular situations that he is facing. This way each user can optimise the use of an MWS individually and be able to customise the user interface for his particular tasks.

Whether or not to use the local language in each country as the preferred medium of communication seems to be culturally dependent. Three choices were brought up:

(i) all communication in English;

(ii) use a dual language system, where English is one language and the local language is the other;

(iii) use only the local language since concise meaning in short words are more easily communicated this way.

5.5 METEOROLOGICAL DATA

The data access problem seems to be one of the most crucial problems to solve in order to make the MWS successful. For users, the potentially greatest benefit of an MWS is its ability to assimilate data.

It is realised that with today's workstation technology, time-critical applications require that graphics data are pre-computed. However, the trend in the evolution of computer technology is such that, within 5 years we might expect a low-cost desktop workstation with sufficient power to handle even critical applications without data pre-computation. It was noted, though, that the user demand for increasingly higher model resolution may be matched by the increase in workstation CPU power. It is concluded that throughput is the critical factor in every application.

To handle data more efficiently, a reference was made to the EGOWS meeting in Norway. There it was suggested that data header information could be put in a database management system, referencing a data file located elsewhere.

It is generally agreed that GRIB is accepted for representing fields and satellite images, and BUFR for observations. However, some questions were raised regarding the efficiency of the proposed extension to GRIB for satellite images. It should be noted, though, that the purpose of the standard formats is to make data portable, rather than an efficient input format for computer processing. Also, since the satellite image data are already represent in an unpacked, binary format the performance penalty is negligible.

No one seems to employ a standard format for radar data, although WMO recommends BUFR. Presently, every site defines its own. However, the EEC sponsored program, COST73, is attempting to define a standard format for radar images, based on BUFR. Radar archival is only being done in a few places in Europe, although it was agreed that there is a general need for this.

Future data representation requirements include possible 3D field representations in GRIB format.

5.6 HARDWARE REQUIREMENTS

The UK project manager for the Horace project quoted their requirements for a high-performance MWS. These requirements seem to reflect the requirements for other meteorological centres. The requirements are as follows:

- more than 50 MIPS CPU power;
- 60-70 MB memory;
- 3 GB local disk (1-3 GB is the range of requirement for all centres);
- 21" monitor, 1280x1024/256 colours
- diskless X terminals;
- FDDI and Ethernet networks;
- 600K 3D vectors per second;
- Motif.

It is generally agreed that some distribution of processing is always necessary. Also, the trend is for rapidly increasing memory requirements due to more complex vendor and user software.

5.7 SOFTWARE REQUIREMENTS

The following components are required:

(i) Operating System.

Unix is by far the preferred solution, although it is realized that there are important differences between different vendor's Unix implementations.

(ii) Network.

Some experiences with NFS/RPC have given interesting results in terms of portability and of processing on CPU with different data formats. However, it is regrettable that OSF has released a competing, but incompatible, network product also called RPC.

(iii) Graphical User Interface.

The conclusion is that Motif is the most powerful and has the most features but that Xview is the easiest to program. It is still not clear which of the two will be the preferred product.

(iv) Graphical User Interface Tools.

Experience has shown that 40-60% of X based application code is in the user interface code. Therefore, the emerging GUI builders, such as Guide (Xview) and XDesigner (Motif), are very important programmer productivity tools.

(v) Graphics Libraries.

Many existing European meteorological applications rely on GKS as the preferred graphics standard. Standard GKS and X are incompatible since both want to be in control of the drawing surface. Since X controls all windows system-wide, it seems reasonable to require that GKS packages rely on X for window management. With respect to graphics output, it is agreed that the GKS standard can be matched with X. With respect to graphics input both X and GKS offer powerful, but incompatible features, and developers may consider either in their implementation approach. It was noted that PHIGS and GKS has the same input model.

(vi) Object-Oriented Programming (OOP) Tools.

It was agreed that OOP Tools should be evaluated for implementation of MWS projects. These tools are now mature enough for realistic usage, but no firm standards have yet emerged. C++ seems to be the OOP language closest to standardisation.