In this issue

European Centre for Medium-Range Weather Forecasts

Shinfield Park, Reading, RG2 9AX, UK

Fax: (0118) 986 9450

Telephone: National 0118 949 9000

International . . . . +44 118 949 9000

Internet URL . . . . http://www.ecmwf.int

Cover

IFS / ARPEGE is the topic of two of the articles in this newsletter (see pages 2 and 11)

Erratum

The following acknowledgement was omitted from the front cover caption of the last newsletter (issue 74):

“Source FOTO Doede de Vries”

We apologise for this error.

Editorial

The development of IFS (ECMWF’s Integrated Forecasting System) was begun 10 years ago in collaboration with Météo-France. It is now used for both operational and research purposes in both organisations. A review of the first 10 years of this project is given on page 2.

The extension of IFS to enable it to run on a distributed memory computer architecture (CRAY T3D) has been described previously. That version has been made portable, and today is used as the production version on ECMWF’s 46 processor Fujitsu VPP700 system. Some background of these recent developments can be found on page 11.

Further summaries of two more of ECMWF’s Technical Memoranda can be found on pages 7 and 8.

The development of ECMWF’s Meteorological Archive and Retrieval System (MARS) has just passed a significant milestone, namely its successful implementation on the new data handling system. Currently it is running in parallel with the existing IBM ES9000 version. This version is a new one, its design having benefited considerably from the experience gained with the original system. An outline of the design of this new version is given on page 9.

ECMWF is exploring the use of CD-ROM technology to make its data available to a wider audience. The first such set of CD-ROMs will be available shortly, the details are on page 10.

The operational Data Acquisition and Pre-processing system, having run on DEC VAX equipment under its proprietary VAX VMS operating system for many years, has now been replaced by a UNIX-based version. An overview of this is on page 14.

Changes to the operational forecasting system

Recent changes

- A new version of the optional project global wave model (both global and mediterranean) was implemented on the 13 May 1997. It consisted of a change in the representation of the wave spectrum to solve the excessive shadowing effect behind islands that was experienced before. This allowed the re-introduction of islands that were deliberately omitted in the previous version.
- Several modifications to the 3D-var data assimilation system were implemented on the 14 May 1997 (model cycle 16r2):

1. two modifications of the forecast error penalty function (J_b). One allows the use of a refined balance constraint, with most impact in the tropics (better fit of the analysis to the observations). The result is also slightly positive in terms of forecast scores. The other modification is in the formulation of the representativeness error growth used for the assimilation of asynchronic data. The errors are now supposed to saturate at levels that are fully varying with space, and this was shown to have a positive impact both on the analysis increments and on the data quality control;
2. several modifications in the way the TOVS data are used (quality control, modification of the tropical bias correction in line with the J_b re-formulation);
3. a restriction of the dynamic initialisation to the small scales.

Planned changes

- ♦ Changes in the use of satellite data and a revision in the computation of water vapour saturation pressure in the model analysis should be implemented soon.

François Lalaurette

1987-1997: Ten years of research and operational activities with the Integrated Forecasting System (IFS)

Introduction

In 1987 the development of a new forecasting system started at ECMWF, the Integrated Forecasting System (IFS), although this was not the original name. The year 1987 was also probably a major turning point in the evolution of the ECMWF models for the following reasons:

- ♦ As indicated by the "I" of "IFS" this system integrates many (almost all) of the applications necessary for an operational NWP (Numerical Weather Prediction) Centre like ECMWF, involved in global modelling, into a single Fortran code. The main advantage is to keep full consistency between the computations performed in different parts of an NWP suite: analysis, forecasting model, initialisation, etc. Before 1987, the different NWP systems developed at ECMWF always consisted of at least two separate parts, one for the analysis and one for the model, and indeed this was true in all the major NWP centres in the world. This was the case at ECMWF for the spectral model developed before the IFS, for the grid-point model and for the OI (Optimal Interpolation) analysis code which went into operations in September 1986. The first ECMWF OI (operational from 1979 to 1986) was even further away from this "integration property", as it consisted of several different Fortran codes, not necessarily consistent.
- ♦ 1987 was the beginning of a cooperation between ECMWF and Météo-France on the IFS which still goes on 10 years later. Incidently, both 1987 and 1997 are (rugby) "Grand Slam" years for the French. This type of cooperation is almost unique in NWP; it means that the scientists from two major NWP centres have developed and maintained in common a single major code, keeping consistent both the scientific and technical aspects needed for the research and operational runs made in Toulouse and Reading on two different computers. The HIRLAM system has also been developed in cooperation between several European countries, but (as we will see later), the number of scientific features is significantly smaller in the HIRLAM system than in the IFS, and also the HIRLAM analysis code was originally imported from ECMWF and not developed from scratch. In Météo-France, the IFS is called "ARPEGE" (Action de Recherche Petite Echelle Grande Echelle), thus the system is often referred to as "IFS/ARPEGE" or "ARPEGE/IFS" in the literature.
- ♦ There are many reasons to believe that the life-time of the IFS will be significantly longer than the usual life-time of preceding models or analysis systems.

Scientific motivations for starting the IFS

The main scientific motivation for starting a new system in 1987 was the desire to start working on a 4D variational assimilation (4D-Var) system. Previous work by Lewis and Derber (1985) and Le Dimet and Talagrand (1986) had shown that by using the adjoint of a forecast model, one could solve the assimilation problem through a global minimization taking all the available information (observations and other information) in a way fully consistent with the model equations. The work described in Talagrand and Courtier (1987) and Courtier and Talagrand (1987) confirmed that a 4D-Var algorithm was feasible and efficient in assimilating real observations (radiosonde data in this case). To pursue this research work there was a need to have a global model, with specifications close to the operational ones, which could be run together with its adjoint. A short attempt (a few weeks) was made to try to code the adjoint of the ECMWF model which was operational at this time. It showed rapidly that this could not be done without a major rewriting of the code. The day when this fact was noted and accepted can be considered as the birthday of the IFS concept. It was then decided (initially by Philippe Courtier and Adrian Simmons over coffee in the ECMWF restaurant) to code a new global spectral model, together with the associated tangent linear and adjoint versions, at least for the adiabatic part of the model. The tangent linear model is almost a necessary step to coding the adjoint. In addition, by definition, it forecasts the evolution of a small perturbation of the initial conditions of the model, which is a useful application in itself. Also it was already clear at this time that an adjoint model could have several other interesting scientific applications, in addition to 4D-Var, see Talagrand (1985).

In the 4D-Var concept, the model and its equations are at the core of the algorithm; the assimilation is built "around" the model. 4D-Var then pushes naturally towards technical solutions where the model and analysis parts are fully integrated.

One particular case of 4D-Var is 3D variational (3D-Var) analysis, in which the time interval used to perform the assimilation is reduced to one instant. 3D-Var is then an intermediate and natural step toward the operational implementation of 4D-Var. From the very beginning of the IFS project, 3D-Var was a scientific motivation as important as 4D-Var, because of two obvious advantages which were foreseen:

(a) better use of observations and the possibility to use more observations;

(b) better spacial consistency of the 3D-Var analysis.

Indeed, a 3D-Var analysis is performed by solving a global minimization problem, which allows nonlinear observation operators, provided these operators are differentiable so as to be treated by the adjoint technique: this is the explanation of advantage (a) (the best example for illustrating this aspect is the use of TOVS radiances). Advantage (b) comes from the global minimization which uses all the observations in one go, avoiding the local data selection which is necessary in all operational OI schemes and which introduces noise into the analysis. More details can be found in Pailleux (1993).

When the IFS project started, 3D-Var was also seen as the first main operational goal for the data assimilation side for the following reasons: 3D-Var is not in principle more expensive than an OI analysis, and it is much cheaper than the corresponding 4D-Var, as the minimization does not have to include the forecast model and its adjoint at each iteration of the minimization algorithm. It has also some potential to make better use of the satellite data. Also a change from an OI to a 3D-Var analysis does not change the overall structure of an operational data assimilation suite, which can be kept with four analyses per day at 00, 06, 12 and 18UTC.

Although the main scientific motivation was on the data assimilation side, some new scientific aspects on the model side were also pushing to code a new model. The main aspect was the concept of variable resolution in a global spectral model which was studied by Courtier and Geleyn (1988), following an idea from Schmidt (1977). A variable resolution model is an alternative to limited area models for studying high resolution features in the forecast. But the highest motivation for the variable resolution was in 1987-88 on the Météo-France side, where studies started for designing a new operational system: from this time it was envisaged to replace both the global model of Météo-France "EMERAUDE" and the limited area one "PERIDOT" by a single global stretched model where the highest resolution is put over France: this is almost a natural choice for a model devoted to short-range forecasting; this was the beginning of the French project ARPEGE, and almost naturally ARPEGE and IFS started as one single project... and one single Fortran code.

In 1987-88, both the French and ECMWF operational models were using Eulerian numerical schemes. However, a common motivation was to work on semi-Lagrangian schemes, also for operations, with the idea that longer time steps could be used with the same horizontal resolution, leading then to a higher resolution model for the same computer time. It was then decided at the design level that the IFS should contain several options, some Eulerian, some semi-Lagrangian.

Technical features of the IFS design

In order to satisfy the above motivations, the following features were included in the IFS (at least in the design) at a very early stage of its development:

- Global spectral primitive-equations model, with a triangular truncation in the horizontal, and a hybrid vertical coordinate;
- Corresponding tangent linear and adjoint models (initially for the adiabatic part);
- Barotropic versions of the model for carrying out some research studies: one is a vorticity equation model, the other one is a shallow-water model; both have their corresponding tangent linear and adjoint versions;
- All the software needed to carry out 4D-Var assimilation experiments; this software includes that necessary to carry out 3D-Var:

(i) A minimization procedure based on the software imported from INRIA (Institut National de la Recherche en Informatique et Automatique);

(ii) The observation operators together with their adjoint and tangent linear versions, plus the code necessary to compute the "Jo cost function" (distance of the model to the observations);

(iii) The code necessary to compute the "Jb cost function" (distance of the model to the background – originally in a very simplified form);

(iv) The code necessary to compute the "Jc cost function" (a constraint for the control of gravity waves);

- Some utility software, such as that necessary for preparing the IFS runs and postprocessing.

All these facilities were (and still are) integrated in a single Fortran code which has been progressively used by more and more people. Some precise norms for developing and documenting the code were followed by the scientists working on it; this was necessary to make it manageable. In most of the applications several options were developed, which increased the scientific possibilities of the IFS, but also its complexity. As an example, the adiabatic part of the IFS can be run in Eulerian or semi-Lagrangian modes, with a reduced grid or not, and with a grid-point at the pole or not. Considering the physical parametrizations, two different packages were developed and can be used in the IFS: the ECMWF package and the French one.

At the beginning, no OI analysis was planned in the IFS, which was mainly a variational tool for the data assimilation scientists. At ECMWF, the need of a new OI code was not felt at this time for a simple reason: the ECMWF OI had been entirely recoded from 1984 to 1986, and the new code had been operational only since 1986. However a need for an OI code started to be felt on the French side in 1988, and in 1989 it was decided to develop an OI analysis code, also integrated to the IFS: this means that it could use the same observations as the variational analysis code, and also the same observation operators. This OI code has been called "CANARI" by the French, and so far has been used only by them.

Other configurations, not planned in the initial design, were added to the IFS later on. Let us quote:

- Software for computing the most unstable modes in the initial state of the model;
- Software for computing the gradient fields with respect to the initial state of any model output function, for example the forecast RMS error over a given area; this is useful for determining the sensitive areas of the model;
- An Extended Kalman Filter (EKF) which has been used for the evolution of forecast error, see Bouttier (1994);
- Comprehensive post-processing software (called "FULL POS") useful for many applications, from the production of model output fields for forecasters to the space interpolations which are necessary when going from one model geometry to another;
- An option to use distributed-memory computer systems.

Although not planned in the initial design either, the IFS/-ARPEGE code has been used from 1990 onward for developing a limited area spectral model. This particular application, called ALADIN, has been carried out by Météo-France, in collaboration with several other European countries. Its scientific aspects will be discussed in the next section.

Evolution of the main IFS/ARPEGE scientific features in 10 years

Forecast model

In the first design, the model was an "In-Core Model": it was called "ICM". A version of this in-core model was already available in 1988, together with its tangent linear and adjoint parts. It was used mainly for basic studies on 4D-Var and on the adjoint technique. Then, under the pressure of computer technologies, an IO scheme was developed and the model was renamed "IFS" at ECMWF (ARPEGE remaining the French name). People working on IFS experiments can still find some historical residuals of the old name "ICM" in the names of some files.

The physical part of the model was developed immediately afterwards, first in France (development of a new package), then at ECMWF (adaptation of the existing physics). At ECMWF, a T213 semi-Lagrangian version of the old code with 31 levels in the vertical was implemented operationally in September 1991: it was only afterwards that efforts were concentrated on the IFS, in order to make a proper operational version which was implemented in March 94 (also at T213/31L). In France the first operational version of ARPEGE was implemented in September 1992, at the same geometry as its predecessor (T79/15L, without stretching), together with the CANARI OI analysis.

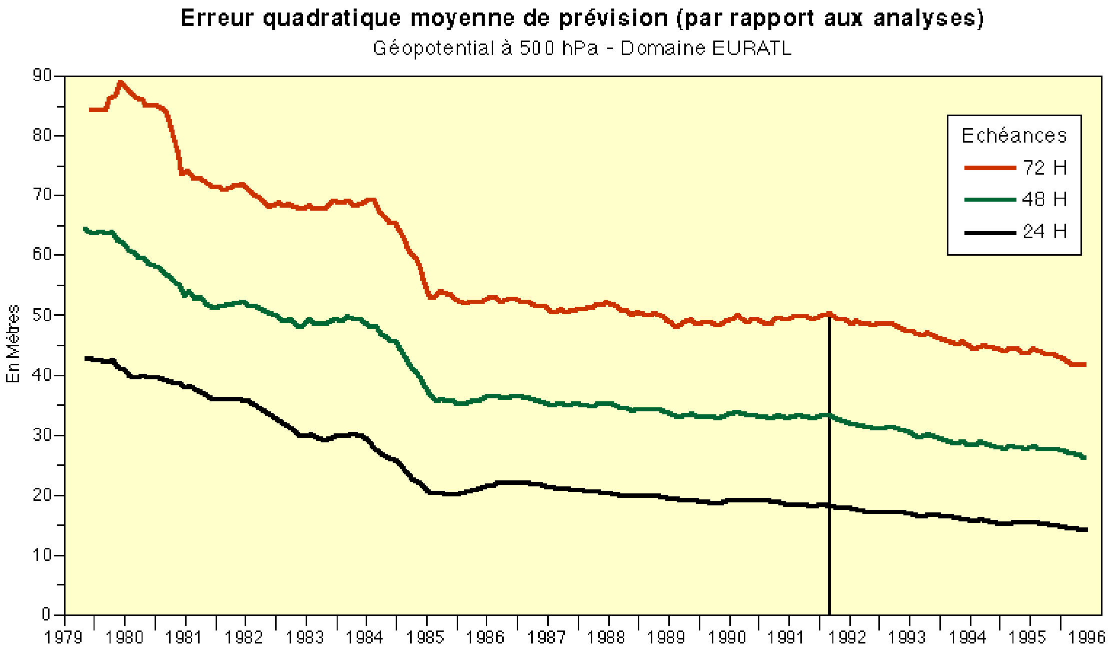
Many other changes occurred after these first operational implementations of IFS/ARPEGE, especially at Météo-France. It took until October 1993 to implement a variable resolution version of ARPEGE with a stretching factor 3.5(T95/21L/C3.5); it used an Eulerian scheme whose resolution was increased to T119/24L two months later. A lot of work had to be done on various aspects of the semi-Lagrangian scheme and on the horizontal diffusion scheme, in order to implement operationally at Météo-France a semi-Lagrangian version of ARPEGE at T149/27L, with a stretching factor 3.5 (October 1995). This is still the French operational version in 1997.

Assimilation

As soon as the first forecast model became available in 1988, together with its adjoint, reduced to its adiabatic part, and in its in-core version, it was used to perform basic 4D-Var experiments, to demonstrate its feasibility and to evaluate its cost. This was done with simulated data (assimilating fields) first, then with real observations when the first observation operators became available in the IFS, see Thépaut and Courtier (1991) and Rabier and Courtier (1992). These results, obtained in close collaboration between ECMWF and Météo-France, happened to be crucial for understanding the power of 4D-Var in some particular situations, and to design an operational scheme, although they were obtained at low resolution (T21 or T42).

In parallel to this 4D-Var research, the developments went on (from 89 onward) to get all the ingredients necessary for a quasi-operational version of 3D-Var. These developments were almost entirely made at ECMWF, as Météo-France was working mainly on the development of the CANARI analysis (using the same observation format and the same observation operators inside the same IFS code). Interesting results on use of TOVS in 3D-Var were obtained in 1991/92 (see Andersson et al (1993)), and on the use of scatterometer data (see Gaffard et al (1997)). However, in order to reach the operational stage, two basic problems had to be solved, which were not foreseen when the first IFS design was made:

- The need to go for an "incremental" technique, both for 3D and 4D-Var; for 3D-Var the incremental technique implies minimization in a space which is different from the forecast model space and which has a much smaller dimension;
- The necessity to have a proper mass/wind balance included in the Jb computation, which is equivalent to having proper multivariate structure functions for the variational analysis; a lot of work had to be done



12 month running mean of the RMS error (in metres) of the French operational model since 1979, for the 500hPa geopotential height at ranges 24h (bottom), 48h (middle) and 72h (top). The verification area covers Europe and North Atlantic.The vertical bar corresponds to the date when the first operational ARPEGE model is taken into the statistics. Before this date the operational system EMERAUDE was frozen for 3 to 4 years in Meteo-France. After this date, several developments on ARPEGE went into operations. (Figure produced by the French control team)

at ECMWF from 1992 to 1994, which ended up with "non-separable" structure functions in the variational assimilation, another obvious advantage of the variational formulation... not foreseen at all in 1987! See Courtier et al (1997) , Rabier et al (1997) or Andersson et al (1997) for more details.

Eventually, the first operational version of 3D-Var was implemented at ECMWF in January 96. In France 3D-Var was implemented operationally on 5 May 1997. In order to reach the operational stage for 4D-Var, it will require probably about two years of work after the operational implementation of 3D-Var. An operational 4D-Var could then be reached by the end of 97 or the beginning of 98 for ECMWF, and in 1999 for Météo-France.

Other applications of the adjoint technique

3D- and 4D-Var were major initial goals of the IFS in 1987; it took a long time to reach the first operational implementation of a variational analysis. However, when performing the basic research experiments on the assimilation, it appeared that the adjoint technique could be used for some applications in routine mode. The best example is the "sensitivity studies": every day the areas which are sensitive for the forecast error over one given area, or for a particular parameter of the forecast, can be computed through one integration of the IFS model and its adjoint (a simpler version compared to the operational one). This type of computation has been performed every day on the operational suite, at ECMWF since 1994, and at Météo-France since the end of 1995.

The limited area model ALADIN

Although the IFS was designed as a global forecasting system, some work started in 1990 at Météo-France, aiming at producing a limited area version of the model. Indeed a huge part of the global code could be used (general organisation routines, a big part of the dynamics, physical routines). This limited area model, called ALADIN, was developed successfully in cooperation between Météo-France and several other weather services, most of them being from Central Europe. This ALADIN project was triggered by several motivations:

- The desire of several weather services to have a limited area model for operational forecasting, and to be involved in its development;
- The understanding that the global stretched strategy had its own limits, and that a limited area model was also needed to go to very high horizontal resolutions.

The ALADIN cooperation was indeed very successful in spite of many technical difficulties, with a first version of the ALADIN model run every day from May 1994 in Toulouse, and several versions run in different countries in 1997. This is probably the best example of a project which is a direct outcome of the IFS, and very successful in spite of its absence in the 1987 design.

Digital Filter Initialization

Developed and first used for the ALADIN model, the Digital Filter Initialization technique was then applied to the global system IFS/ARPEGE. It went into operation at Météo-France in July 1996, replacing the Normal Model Initialization technique.

Main operational milestones of the IFS/ARPEGE project

September 1992	*ARPEGE model operational at Météo-France (T79/15L) without stretching and in Eulerian mode*
December 1992	*Computation of the unstable modes for the Ensemble Prediction System at ECMWF*
October 1993	*ARPEGE in stretched mode at Météo-France (T95/21L) with a stretching factor equal to 3.5, and in Eulerian mode*
February 1994	*Computation of the sensitive areas at ECMWF*
March 1994	*IFS operational at ECMWF in replacement of the old code (T213/31L) - semi-Lagrangian version*
May 1994	*First routine run of ALADIN model at Météo-France*
October 1995	*ARPEGE in stretched mode and semi-Lagrangian mode operational at Météo-France (T149/27L)*
December 1995	*Computation of the sensitive areas at Météo-France*
January 1996	*3D-Var operational at ECMWF*
July 1996	*Digital Filter Initialization operational in the French ARPEGE model*
September 1996	*Distributed-memory version of the IFS operational on the Fujitsu VPP700 at ECMWF*
May 1997	*3D-Var operational at Météo-France*

Conclusion

The IFS is probably the biggest set of forecasting models and other numerical tools which has been integrated into a single consistent Fortran code. It has been developed since 1987 in cooperation between scientists from ECMWF and Météo-France, and everything suggests in 1997 that this collaboration should go on for at least several more years.

During these 10 years, apart from the scientific problems, the main difficulty has been to keep all the applications working and consistent in the single code. Scientists working on the IFS must have found the code heavy, difficult to understand and to handle, and difficult to maintain too. However, the code has already been through several computer changes, both in Reading and in Toulouse. Many important operational applications are now working, although it took a long time to reach the operational stage for some of them: 8 years at Météo-France from the beginning of the project to get the operational version of the model with all the desired features in (including the variable resolution and the semi-Lagrangian scheme); almost 9 years at ECMWF to get 3D-Var working operationally.

The figure shows the evolution since 1979 of the RMS error of the French operational model, at ranges 24h, 48h and 72h. These RMS errors are annual running means in order to filter out the seasonal variations, and are for the 500 hPa geopotential height over an area covering Europe and a large part of the Atlantic. The vertical bar corresponds to September 92 when the first version of ARPEGE/IFS replaced its predecessor EMERAUDE (because of the running mean, the vertical bar is on March 92, the first month affected by the change). One can see a continuous improvement of the French operational forecast from this date when new ARPEGE developments were progressively introduced. From 1988 to 1992, during the pure development phase of the project, the operations were frozen and no significant progress was achieved.

If some important applications were a bit slow to reach the operational stage, some other applications reached this stage although they were not envisaged at the beginning (sensitivity studies for example). Also, the IFS has been, even at an early stage, an excellent research tool, especially in the area of 4D assimilation, predictability, and applications of the adjoint techniques. Indeed, the integration of so many research aspects together with operational applications into a single system, is an original and strong aspect of the IFS.

References

Andersson, E., J. Pailleux, J.N. Thépaut, J.R. Eyre, A.P. McNally, G.A. Kelly and P. Courtier, 1993: Use of cloud cleared radiances in three/four dimensional variational assimilation. *Quart. J. R. Met. Soc.*, **120**, 627-653.

Andersson, E., J. Haseler, P. Unden, P. Courtier, G.A. Kelly, D. Vasiljevic, Č. Branković, C. Cardinali, C. Gaffart, A. Hollingsworth, C. Jakob, P. Janssen, E. Klinker, A. Lanziger, M. Miller, F. Rabier, A. Simmons, B. Strauss, J.N. Thépaut and P. Viterbo, 1997: The ECMWF implementation of three dimensional variational assimilation (3D-VAR). Part III: Experimental results. Submitted to *Quart. J. R. Met. Soc.*

Bouttier, F., 1994: A dynamical estimation of forecast error covariances in an assimilation system. *Mon. Wea. Rev.*, **122**, 2376-2390.

Courtier, P. and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part II: Numerical results. *Quart. J. R. Met. Soc.*, **113**, 1329-1347.

Courtier, P. and J.F. Geleyn, 1988: Aglobal numerical weather prediction model with variable resolution: Application to the shallow-water equations. *Quart. J. R. Met. Soc.*, **114**, 1321-1346.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier and M. Fisher, 1997: The ECMWF implementation of three dimensional variational assimilation (3D-VAR). Part I: Formulation. Submitted to *Quart. J. R. Met. Soc*.

Gaffart, C., H. Roquet, B. Hansen, E. Andersson and D. Anderson, 1997: Impact of ERS-1 scatterometer wind data on the ECMWF 3D-VAR assimilation system. To appear in *Quart. J. R. Met. Soc*.

Le Dimet, F.X. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, **38A**, 97-110.

Lewis, J.M. and J.C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A**, 309-322.

Pailleux, J., 1993: Assimilation of satellite data using variational techniques. Proceedings of the ECMWF seminar on "Developments in the use of satellite data in operational numerical weather prediction: 1989-1993", Reading, September 1993.

Rabier, F. and P. Courtier, 1992: Four-dimensional assimilation in the presence of baroclinic instability. *Quart. J. R. Met. Soc.*, **118**, 649-672.

Rabier, F., A.P. McNally, E. Andersson, P. Courtier, P. Unden, J.R. Eyre, A. Hollingsworth and F. Bouttier, 1997: The ECMWF implementation of three dimensional variational assimilation (3D-VAR). Part II: Structure functions. Submitted to *Quart. J. R. Met. Soc*.

Schmidt, F. 1977: Variable fine mesh in spectral global model. *Beitr. Phyz. Atmosph.*, **50**, 211-217.

Talagrand, O. and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. *Quart. J.R. Met. Soc.*, **113**, 1311-1328.

Talagrand, O., 1985: The adjoint model technique and meteorological applications. Proceedings of the ECMWF workshop on "High resolution analysis", Reading, June 1985.

Thépaut, J.N. and P. Courtier, 1991: Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model. *Quart. J. R. Met. Soc.*, **117**, 1225-1254.

Jean Pailleux (Météo-France)

Summary of Technical Memorandum 222

Sensitivity of ECMWF model wintertime climate to model formulation

(This summary is based on a revised version of TM222 published in Climate Dynamics 13, 75-101, 1997)

Čedo Branković and Franco Molteni

Four different versions (cycles) of the ECMWF NWP model have been integrated at T63 resolution on the seasonal time scale in order to study the influence of various model formulations on model climate. The model versions, denoted as cycles 36, 46, 48 and 12r1, differ primarily in the representation of physical processes, but in cycle 12r1 a higher vertical resolution was also used for technical reasons (however, without changing the model vertical domain). The successive cycles were introduced in ECMWF operations over a period of about four years, from 1990 to 1994.

For five northern winter seasons (DJF 1986/87 to 1990/91) three 120-day experiments, initiated from initial data one day apart, were run for each model cycle. The four model climatologies were derived as averages of the 15 integrations in each cycle. In the first three cycles the observed sea surface temperature (SST) was updated every five days throughout the experiment integration, but it was updated daily in cycle 12r1.

With the introduction of cycle 48, an apparent improvement in the northern hemisphere wintertime climate of the ECMWF model was observed. A strong systematic bias in the earlier cycles in the northern hemisphere, associated with an overestimation of the zonal flow in the eastern Pacific and a weak diffluence over western North America, was reduced and replaced by more realistic ridging over the north Pacific/Alaskan region. This error was further reduced in cycle 12r1. The improvement is associated with a reduction in the temperature error over the northeastern Pacific. A reduction of errors in zonally averaged zonal wind and eddy kinetic energy is also evident.

While cycle 46 exhibits reduced model errors in the southern hemisphere summer, its contribution to the northern hemisphere winter mean flow is not evident. It was argued that the impact of the revised cloud shortwave optical properties introduced with cycle 46 is strongly susceptible to seasonal forcing. However, other model changes introduced between cycle 36 and cycle 46 may also have contributed to the reduction of the southern hemisphere errors.

The frequency of blocking with cycle 12r1 is much more realistic over the north Pacific region when compared with the other three cycles, though it was slightly improved already in cycle 48. It was hypothesized that this improved representation of the northern Pacific block is primarily due to a more efficient diabatic response in the later model cycles to the warm SSTs over the western equatorial Pacific. Another possible candidate that may influence the representation of the Pacific block, the anomalous transient forcing that comes from high frequency baroclinic waves associated with the north Pacific storm track, seems to be less important. Namely, the study reveals that the north Pacific high-frequency variability in cycle 12r1 was somewhat weakened in comparison with other model cycles.

Over the Atlantic/European region, a slight deterioration of blocking frequency in cycle 12r1 is associated with the strengthening of the Atlantic jet in comparison with the other cycles.

Not all parameters exhibited an improvement in the climate drift in successive cycles. For example, in the lower troposphere a negative temperature error has slightly increased over the oceans from cycle 46 to cycle 48. This is associated with an increase in the lower tropospheric humidity in cycle 48 relative to cycle 46. There was a small increase in the height of the planetary boundary layer (PBL) in cycle 48 due to the introduction of entrainment at the top of the PBL. The effect of the entrainment in shallow convection contributes to the moistening around 850 mb level.

Arelative warming over the tropical continents, on the other hand, implies further drying and larger errors in cycles 48 and 12r1. Astrong polar stratospheric cooling in both hemispheres is still seen in cycle 12r1, however, this has been reduced in the southern (summer) hemisphere. [It is worth pointing out here that stratospheric cooling has been further greatly reduced with the introduction of the semi-Lagrangian dynamics in ECMWF model.] A warming of the tropical lower stratosphere, which became apparent with the introduction of cycle 46, is still very much evident in cycle 12r1. This error is not unique to seasonal integrations - it has also been seen in the ECMWF operational medium-range forecasts.

The Hadley circulation, which was somewhat weakened from cycle 36 to cycle 46, has become more intense again with the later cycles. The lower branch of the large-scale tropical ascent is now overestimated when compared with the operational analysis (which, however, is likely to be deficient in this respect). This is associated with an intensification in convective rainfall within the tropical convergence zones in cycles 48 and 12r1, particularly over the Pacific, and to a more active shallow convection in these cycles. Another prominent feature of the later cycles is the concentration of tropical precipitation into relatively narrow bands that represent the main tropical convergence zones. Lack of observations to provide climatological data makes the verification of the southern hemisphere model precipitation uncertain.

In terms of representing the northern hemisphere circulation, model cycles 36 and 46 are comparable and cycles 48 and 12r1 can be regarded as improvements over the former two cycles. Overall, the most recent ECMWF model versions, cycle 48 and especially cycle 12r1, have generally a better wintertime climatology than earlier versions. On seasonal time scales, some of the model systematic errors have been steadily reduced and model's ability to reproduce interannual variations has been improved. We have demonstrated that the development of more sophisticated physical parametrizations led to a substantial improvements in the ECMWF model climate.

Summary of Technical Memorandum 224

Impact of prognostic cloud scheme and subgrid scale orography on the simulation of the Asian summer monsoon

Someshwar Das, Martin Miller and Pedro Viterbo

The Asian summer monsoon is a well known feature of the general circulation of the atmosphere. Despite its regular seasonal cycle, medium range forecasts of its different characteristics such as the onset phenomenon, northward progress of the rainfall bands, genesis and movement of the monsoon depressions/ cyclonic storms have been far from reality. The prediction of weather elements averaged over a longer duration of time, such as monthly or seasonal features, may look all right, but the real ability of a model could be justified only when it is correctly able to predict extreme weather events on the medium-range time-scale.

A new version of the ECMWF model, CY13R4, was implemented on 4 April, 1995[1]. This version includes:

(i) a new prognostic cloud scheme

(ii) smoothed mean orography with a new parametrization of the subgrid scale orography and

(iii) changes in the numerics involving semi-Lagrangian scheme and some other minor changes.

When tested independently, these changes have respectively shown to provide better cloud cover and cloud water contents (Tiedtke, 1993), a better estimate of the. mountain drag, flow dynamics and low level wake observed in mesoscale analysis (Lott and Miller, 1997) and controlled aliasing arising from quadratic terms over poles in the reduced gaussian grid (Courtier and Naughton, 1994). The Memorandum assesses the performance of the new version of the model, in particular with regard to the Asian summer monsoon.

The performance of this version of the model was studied against the earlier version CY12R1 for a set of 10-day forecasts with initial conditions of 1-20 June of Monsoon-1994. Comparison of the first two changes above (it is assumed that the change in the numerics is of no relevance for this study) against the old version of the model indicates better forecasts of cloudiness and 2 metres temperature, correcting the negative cloud bias and warm surface bias of the old model. With regard to rainfall, while the old version underestimated the total precipitation over the western Ghat mountains, the new version overestimates the rainfall there. The current

[1] See ECMWF Newsletter 70, pp 2-8

model produces a larger positive bias and higher RMSE of rainfall compared to the old version. This is in contrast to the behaviour over Europe where the rainfall forecasts are clearly better with the new model. The track of cyclonic storms/monsoon depressions was slightly better in the old version. The new model produces better scores of geopotential and wind over the Indian region as compared to the previous version.

Study of the performance during Monsoon-1995 shows a reasonably good simulation of the Somali Jet and cross-equatorial flow during the pre-onset phase. However, the model produced excess upper level easterlies, weak low level westerlies and a slightly stronger Hadley cell compared to the verifying analysis. The onset of the monsoon was predicted about two days in advance as compared to the observed date. However, the model produced relatively weak low level jet and monsoon trough during the active phase. In general, the bias and RMSE of rainfall increased from pre-onset to the onset phase. The model continued to predict better rainfall over Europe and overestimated rainfall over the western Ghat as in the previous year 1994. Quantitatively, the best rainfall forecasts were produced for low rainfall amounts. It underestimated clouds over northern India during the active phase. Early morning cold bias over the northern India was about -9 to -10°C during the pre-onset phase but it improved substantially with the progress of the monsoon when there were more clouds over the region.

Sensitivity experiments were carried out to investigate the problem of underestimation of widespread convective rainfall over the Indian region. Experiments with the subgrid scale orography were performed to see whether the new model produced excess rainshadow effect on the lee side of the mountain. Results indicated that the operational version produces a better simulation of the zones of convergence by allowing the flow to go around the flanks of the mountain and producing less blocking. However, the problem of missing rainfall in the forecasts continued. Analysis of results indicated a drying of the lower troposphere in the forecasts. Therefore, two more experiments were carried out by flooding and drying the soil moisture values. Although the problem of rainfall underestimation remained unsolved, the results provided several insights into the physical mechanisms involved. It is argued and anticipated that further investigation of the closure schemes of the convective parametrizations and cloud formulations may lead to the solution of the problem.

References

Courtier, P, and M Naughton, 1994: A pole problem in the reduced Gaussian grid. *Q J R Meteorol Soc*, **120**, 1389-1407.

Lott, F, and M Miller, 1997: Anew sub-grid scale orographic drag parametrization: its formulation and testing. *Q J R Meteorol Soc*, **123**, 101-127.

Tiedtke, M, 1993: Representation of clouds in large-scale models. *Mon Wea Rev*, **121**, 3040-3061.

Implementing MARS using object-oriented technology

This article is based on a paper given at the 77th American Meteorological Society meeting in Long Beach, California, USA, February 1997.

Introduction

Since 1985, the ECMWF Meteorological Archival and Retrieval System (MARS) has grown both in size and diversity. At its start in 1985, the operational archive was growing at 70 Mbytes/day. In 1996, the growth rate was 125 Gbytes/day of operational data and 7 Gbytes/day of research data, the total archive was 22 Tbytes. The number of individual items archived is also very large: 120,000 operational fields/day, 140,000 reports/day and 200,000 research fields/day.

The present MARS system has now reached its limits and needs to be redesigned to cope with the increase in generated data foreseen with the arrival of a new supercomputer and the never-ending requirements to archive new kinds of data types.

The issues

In designing the new system, using the expertise gained from the present system, we have assessed the main issues to be:

The number of addressable items

For most people, the size of a database is measured in the number of Gbytes it holds. But the real problem is the number of addressable items. In the case of MARS, a user must be able to retrieve a single field amongst thousands of millions. No commercial database is able to address this issue.

The number of physical files used to hold the data

Similarly, the number of files that a commercial archive system handles is in the order of a few million. We already have 7.3 million files in our present system.

Support for various data types, past and future

At present, MARS supports three data types: fields, observations and images. But ECMWF's Research Department regularly comes with new requirements and new data types to be supported by the system. It is not possible to foresee these new data types at the time of the analysis of the system. Thus, we need to keep an open design, so future data types can be added during the lifetime of the system.

Using OBJECT-oriented technology

The success of the future MARS lies in how it will represent its metadata, i.e. the "data representing the data". MARS must cope with data types not yet invented, as well as a variety of very different objects such as fields, observations, images, and cross-sections. Using object-oriented technology will help:

- with *encapsulation* MARS will be able to handle new data types, as each data type being "self aware" will know how to process itself;
- with *inheritance* it will easily support new attributes and help code re-use (e.g. an ensemble forecast is a forecast plus an ensemble number);
- with *polymorphism*, the MARS engine will be simple as it will have a single view of all the data types. As with *encapsulation*, the addition of a new data type should not imply any changes in the existing code. *Polyphormism* is the perfect concept to solve the problem MARS has when trying to cater for a new data type that did not exist at the time of the design.

Other concepts of object technology will be used, such as reference counting and attribute sharing. The former simplifies the handling of large quantities of objects, providing automatic garbage collection for those which are not useful anymore. The latter should reduce dramatically the size of the metadata databases by sharing information that otherwise would be replicated (many of our forecasts have identical attributes such as their number of steps or levels).

The new system

The current MARS service runs on an MVS-based IBM ES9000 system. This is being replaced by three RS6000 R30 systems, with four PowerPC 605 CPUs each, again from IBM. Each machine will have 1 Gbyte of memory, 1 Mbyte of cache per processor, 3 Gbytes of disk and will share 240 Gbytes of disk with the other machines.

Two robots with 16 drives a piece will be able to mount two hundred 10 Gbyte tapes per hour each.

The current hierarchical storage manager used is the Common File System (CFS) from Los Alamos. CFS will be replaced by ADSM, an IBM file store. The MARS server will be written in C++, and its metadata will be stored in an objected oriented database (ObjectStore).

Design

To address the problem of size and number of files, it has been decided to split the system in two major components:

- A *Data Server* will sit on top of ADSM and handle files. Each file to be archived is presented to this server together with its "layout" (its physical organisation, e.g. the size and offset of each field within the file). The server saves the data into ADSM and uses an ObjectStore database, known as the "location database" to store a "data reference", an object containing the data file location together with its layout. Layouts may actually be shared as explained above.
- A *MARS Server* will have the semantic understanding of the user's request. With the help of an ObjectStore database called the "semantic database" it will know about the content of the archive in meteorological terms. For each forecast archived it remembers its attributes (date, experiment identifier, etc.), its "shape" (number of time steps, number of levels, meteorological parameters), and the "data reference" it has been archived into. As for layouts, shapes may also be shared amongst forecasts.

By clearly isolating the role of the two servers, the Data Server will be able to organise the files in the most efficient way according to the underlying software and hardware. On the other hand, MARS will have the meteorological understanding of the data and being relieved of the data organisation aspects, will be able to provide a better service than the present system.

Conclusion

The present design has been successful for more than 10 years and is nowadays one of the largest, most complete and easily accessible of meteorological archives. We are now facing the challenge of redesigning MARS for the next decade. New technologies will be used to enhance the present system into more open and flexible software that will cope with the endless growth of data, in quantity as well as in variety.

References

J.T. Hennessy, 1986: MARS - The ECMWF Archive and Retrieval System, Second International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology, *AMS*

ECMWF Newsletter No. 72 (1996), pp. 15-19: Data Handling via MARS at ECMWF

Baudouin Raoult

The ECMWF Seasonal Simulation CD-ROMs

CD-ROMs containing selected data fields from the ECMWF seasonal ensemble simulation project are under development and will be ready for distribution soon.

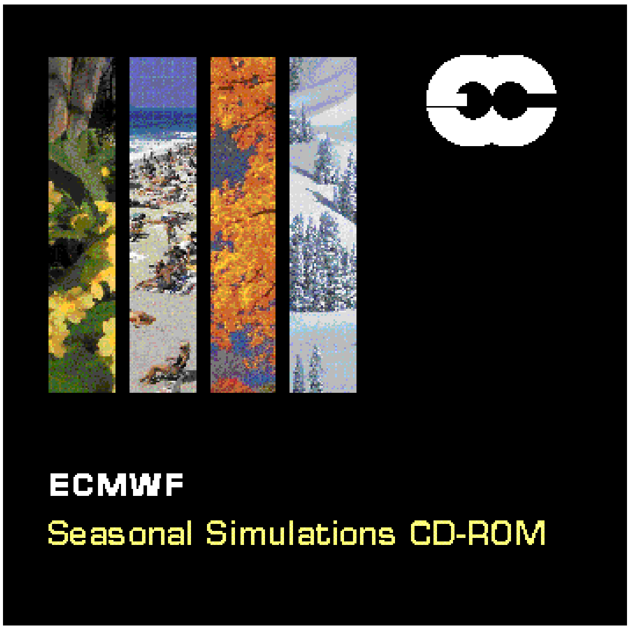
Each ensemble is 120 days long and consists of nine members, started from consecutive 12Z analyses. The integrations are being run over the ECMWF re-analysis period (1979-1993) using re-analysis as initial conditions and observed SSTs taken from the re-analysis project.

In addition to their use as initial data, re-analysis is used to verify the ensemble integrations. The set of CDs being developed contains data from ECMWF ensemble integrations. It provides cost efficient data access to exploit the potential of seasonal ensemble simulations by the wider community of researchers in predictability, particularly those from developing countries.

Results from the ECMWF ensembles simulations for winter, spring, summer and autumn are included in the four CDs. All CDs contain 25 ten-day mean fields for each ensemble member, started from consecutive 12Z analyses immediately preceding the season of interest. That is 1 December for the winter season, 1 March for spring, 1 June for summer and 1 September for autumn. Ensemble member 10 contains ten-day means over 24h forecasts based on re-analysis to attain consistent fields well comparable to the ensemble forecast data.

The following ten-day mean fields (twelve for each 120 day season for each of the 15 years) are included:

- geopotential on levels 700, 500 and 200 hPa (Z), 12 bit
- temperature on 850, 500 and 200 hPa (t), 11 bit
- u- and v-wind components on 850 and 200 hPa (u, v), 11 bit
- vertical velocity on 500 hPa (w), 12 bit
- soil moisture content in the first layer (7 cm) of the soil model used (swl1), 11 bit
- snow depth (sd), 11 bit
- total precipitation (large scale plus convective) (tp), 11 bit
- surface sensible heat flux (sshf), 12 bit
- surface latent heat flux (slhf), 12 bit
- mean sea level pressure (msl), 11 bit
- total cloud cover (tcc), 7 bit
- surface solar radiation (ssr), 12 bit
- surface thermal radiation (str), 12 bit
- top thermal radiation (ttr), 12 bit
- surface wind stress components (ewss, nsss), 11 bit



- minimum temperature at 2m (mn2t), 11 bit
- maximum temperature at 2m (mx2t), 11 bit
- surface temperature (stl1, Member 10 (re-analysis) only), 11 bit
- land-sea mask (lsm, under "season" (winter, spring, summer or autumn, 1 file only) 4 bit.

All of the fields on these CD-ROMs have been converted to a uniform 2.5 degree resolution latitude-longitude grid.

Furthermore, the graphics display systems and GRIB decoding tools are provided on the CD-ROMs.

Bernd Dieter Becker

IFS on the Fujitsu VPP700

Introduction

Previous Newsletter articles have described programming techniques for distributed memory architectures (Programming the Cray T3D - Newsletter 68, page 29), and the nature of distributed memory computer systems at ECMWF(Newsletter 73, page 26). This article outlines the recent technical developments of IFS, with an indication of its performance when in production on the VPP700.

Migration

Migration of the IFS model from the CRAY shared memory platforms required some source code development in order to:

(a) remove CRAY features from the Fortran language

(b) implement parallelism for a distributed memory environment.

The first was initially achieved by removing Cray specific Fortran features (such as POINTERS) to create a portable code which became the starting point for the parallel development. The second was achieved by the use of message passing techniques. Although this approach was rather manual and required a non-trivial effort from experienced programmers, the reward was a portable code with potentially excellent parallel performance. It has resulted in a version of IFS which can be used as a portable benchmark, and which indeed has run on a variety of competing platforms.

Subsequently we realised that it was impractical to maintain two versions of IFS, the portable version and the (then) Cray shared memory production version, on which substantial meteorological development continued. Hence, a further effort was undertaken to merge this (Fortran 77 based) portable version into the (Cray based) production version using Fortran 90 and MPI message passing in order to retain portability. The result is now the only version of IFS in existence. However, it still retains the shared memory Macrotasking options which are required for production use at Météo-France. (The code development project is represented schematically in Fig. 1.)

The parallel implementation

The key to creating an efficient message passing version lies in identifying the inherent parallelism. The first problem is that calculations in the different phases of the

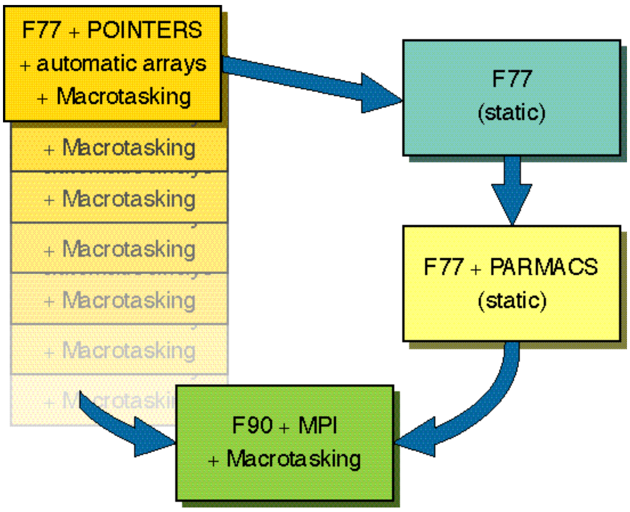


Fig. 1: *Schematic view of the code development project*

IFS spectral model involve complex data dependencies in different directions. For example, during the Fourier transforms, the zonal wave numbers, m, are dependent on the grid-point latitude data. However, in the Legendre transforms they are dependent on the longitudes, and in grid space the physical computations involve vertical data coupling only. In addition, the semi-Lagrangian advection scheme requires access to data from nearby grid columns in order to compute trajectories, thus introducing a degree of local communication.

Atransposition strategy is therefore used to handle the problem of accessing the data. With this approach, the relevant data is redistributed to the processes at various stages of the algorithm in advance of their use, so that the arithmetic computations between any two consecutive transpositions can be performed without further inter-process communication. Such an approach is feasible because data dependencies exist only within one coordinate direction for each algorithmic component. An overwhelming practical advantage of this technique is that the message-passing logic is localized into a few routines. These are executed prior to the appropriate algorithmic stage, so that the computational routines (which constitute the vast bulk of the model code) need have no knowledge of this redistribution of data.

Next, in order to achieve good load balancing, the distribution of work has to be done carefully. In grid-space this is complicated by the use of a 'reduced-grid' in which each line of latitude contains a differing number of points. A careful work distribution strategy allows the total array of grid-columns (138346) to be evenly distributed among any number of available processors such that they all receive an equal amount of work. For example, for a partition of 100 PEs, 54 will receive 1383 data points, and 46 will receive 1384 data points.

Such load balancing is 'static'and is independent of the particular weather situation. However, a dynamic imbalance can also exist due to variations in the amount of computation on particular grid columns arising from the meteorological conditions. For example, a grid-column near to the equator may have to deal with substantial amounts of convective activity, while a polar column may have none. This dynamic imbalance is much more difficult to handle since, in principle, it requires a redistribution of grid-points between processors at each model time step. Currently, no attempt is made to address this inefficiency which is of the order of a few percent.

The reduced grid also presents difficulties in achieving good vector performance. Since the predominant vector dimension is east-west, the vector length varies from 640 for equatorial rows to only 16 at the poles. To overcome inefficiencies due to these short vectors, a buffer is provided into which is packed a varying number of latitudinal rows. The resultant 'super-row' is then presented to the computational routines for processing. The buffer size is defined at run time, thus allowing the user to optimise vector length at the expense of memory space. The success of this technique may be judged by the fact that the IFS model achieves a total sustained floating point computational rate of about 770 Mflops on each VPP700 processor (35% of peak).

Distribution of the cp cost between the most important computational components is indicated in Fig. 2. The Legendre transform (lt) remains at a reasonably low fraction (7%) although as it scales cubically with model resolution, its share will certainly increase in future higher resolution models.

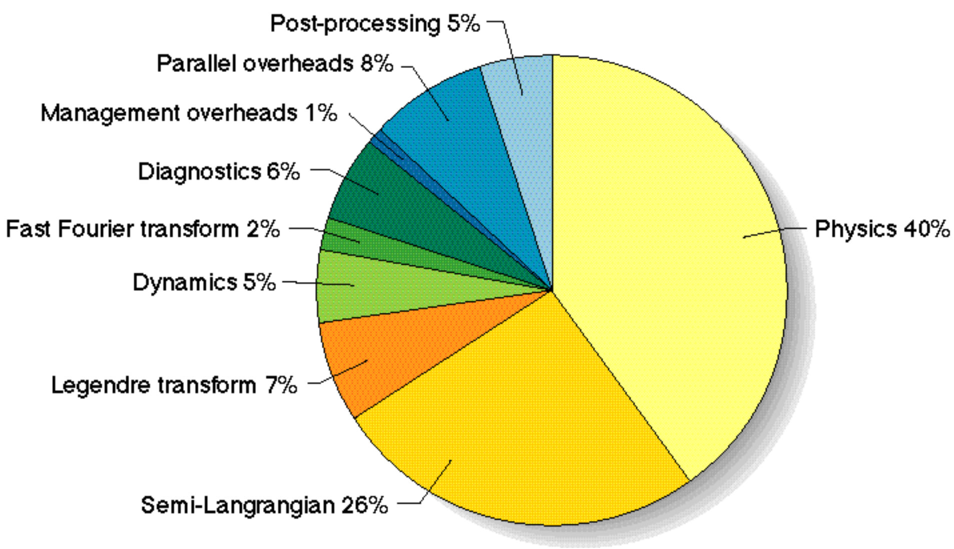


Fig. 2: *Distribution of the cp cost between computational components.*

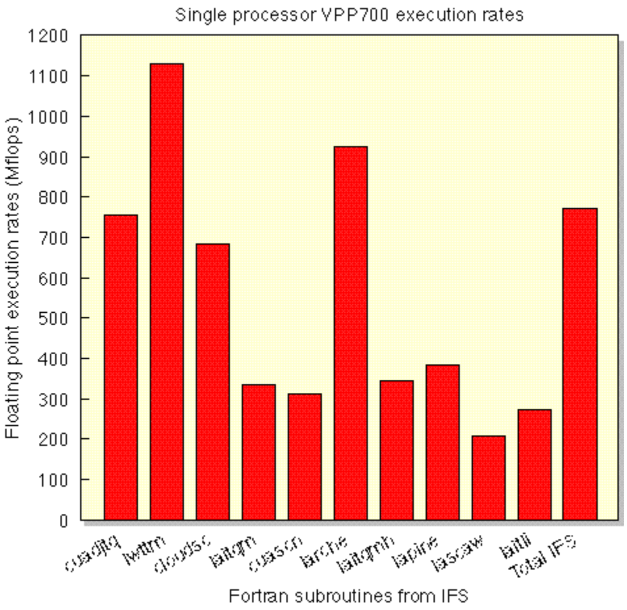


***Fig. 3:** Floating-point execution rates for the 10 most expensive IFS routines*

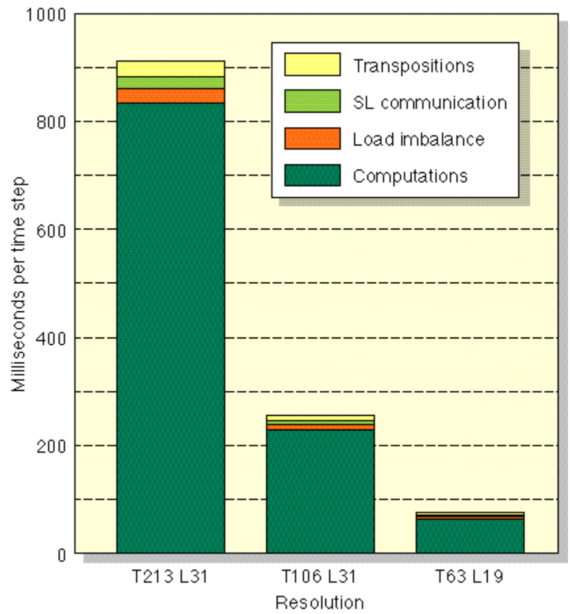


***Fig. 4:** Message passing and load imbalance inefficiencies when run on 46 processors of the Fujitsu VPP700.*

Vector performance

From the vectorising point of view, conversion from Cray to Fujitsu required only a small effort. The Cray code contained many uses of 'CDIR$ IVDEP', the well-known directive to persuade the compiler to 'ignore vector dependencies'. The equivalent directive on the VPP is '!OCL NOVREC' so that migration consisted solely of translating from one form to the other.

With regard to the performance of the vectorised code, Fig. 3 shows floating-point execution rates for the 10 most expensive IFS routines. The most interesting aspect is the wide range of performances. The poorer rates are due either to conditional code, or to indirect addressing probably also incorporating bank conflicts. There is some hope that improvements can be made to some of these figures, but the overall (single processor) rate is reasonable.

Parallel performance

Fig. 4 quantifies message passing and load imbalance inefficiencies when the model runs on 46 processors of the Fujitsu VPP700. It can easily be seen that the inefficiencies are small compared to the computational time.

To investigate scaling, the model has been run using various numbers of PEs of the VPP700, (see Fig. 5). The parallel speedup is 15.4 with 16 processors and 35.5 with 40 processors. These are all based on total elapsed time measurements, and indicate that the message passing overhead is not excessive.

Fig. 6 shows a message trace facility. Using this tool it is possible to obtain detailed information about message passing traffic. Most of this activity within the model proceeds at close to the hardware speed (> 500 Mbytes /sec). This is achieved by ensuring that the numbers of short messages are minimised, i.e. fewer long messages are best for good performance.

In summary, the production forecast (T213L31 10 days) today utilises 18 processors. This is an operational choice made so as to minimise the overall run time of the entire time critical suite of jobs. Typical performance statistics show:

CPU = 94.51%
Vector = 74.62%
Maximum memory used = 776 Mbytes

with a run time of about one hour.

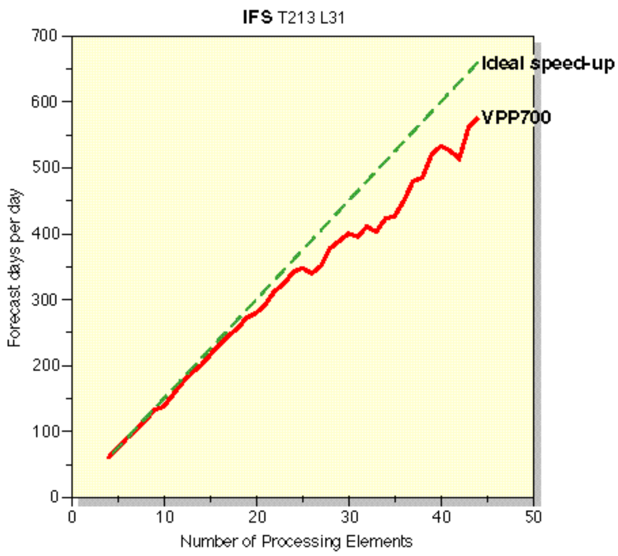


***Fig. 5:** Execution rates as a function of the number of VPP700 processors utilised.*

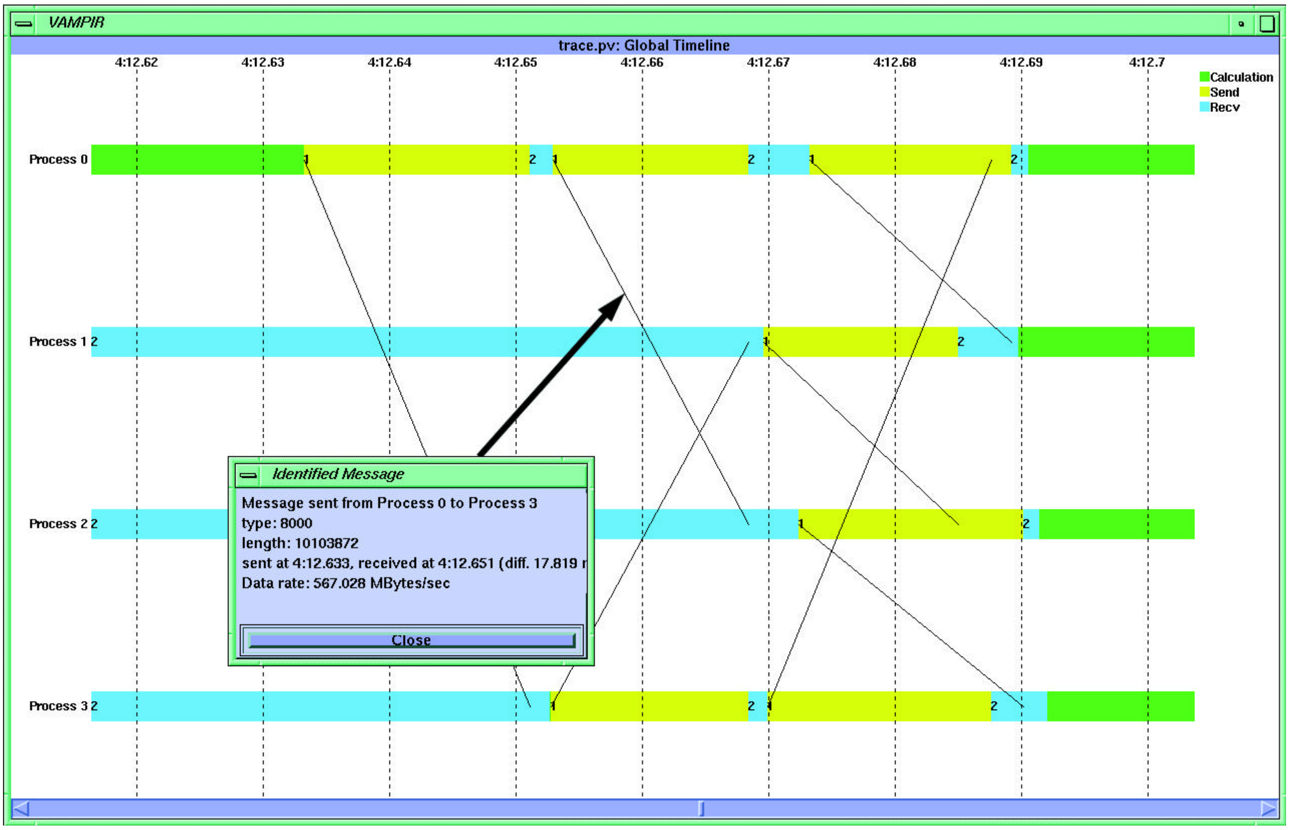


Fig. 6: *Message tracing facility showing detailed information about message passing traffic.*

Source code developments in progress today should substantially reduce the memory requirements. New versions of the compiler are expected to, for example, reduce the scalar time which includes dynamic memory allocation costs.

David Dent

The Data Acquisition and Pre-processing on the High Availability HP System

A final farewell was given to the old VAX VMS based Data Acquisition (DA) and Pre-processing (PP) system on 11 February 1997. It had served us very well for many years, but times are changing leading to the need for a new system. Hence a new UNIX based one has been brought into service.

The transition into operational use went very smoothly. So here we are now with a UNIX version. Let us see what is really new.

Meteorological changes

The basic idea has remained unchanged. The system is still built upon two databases, the Message Database (MDB) and the Reports Database (RDB), but the main characteristic of the new system is the use of commercial relational database software from Empress instead of our original homemade software. That is not the only change. Data Acquisition is a completely new piece of software. It is very simple and robust, and uses Empress to store meteorological bulletins into the database. On the other hand, the Pre-processing software remains mostly as it was on the VAX. The main difference is in its interface to the MDB and RDB. Many new decoding streams have been added to the system, such as SSMI mapped data, SSMI brightness temperatures, RTOVS etc.

There are some changes in the quality control, for example to the position tracking. Also incoming observations already in BUFR format are now subject to quality control. Thus we can carry out quality control for ACARS and 120km TOVS coming in the BUFR format, or we can redo any previous quality controls on the old data in the BUFR format. Most other segments of quality control remain unchanged.

The new DA is handling a lot more data than the old DA. ECMWF is getting about 300 Mb per day by ftp, which is about ten times the volume of conventional GTS data. At the end of the year it is expected to handle about 1.3 Gb per day.

To have more control over the data handling, some new tools have been developed. All the tools are X windows applications. It is now possible to monitor in a graphical way the amount of data received both by reception time and data type. This is used to detect late data. The interactive data coverage tool allows global data coverage of any observation subtype with observations of bad quality marked. The same tool is also very useful in giving information about the number and distribution of observations being extracted from the RDB at different cut-off times.

An integrated tool to access both databases and show the content of a particular bulletin and its corresponding BUFR message has been developed. It makes problem solving much more efficient and easier to do. During decoding when bulletins are found to be in error they are written to error files. The X based interactive correction tool allows easy access to these "bad data" files, and via an editor to correct and reprocess them again.

The RDB data extraction routines for analysis and archiving have remained essentially the same, except for their database access interfaces.

Hardware and software

The new data acquisition and pre-processing software runs on the HPhigh availability cluster. The cluster consists of three Series 800 servers, running the Hewlett Packard version of UNIX called HP-UX, version 10.01. These machines are connected to the local area network through FDDI interfaces on each host, and they also have a private Ethernet network to exchange "heartbeat" and status information between themselves.

Currently there are two "highly-available" packages defined: the Data Acquisition plus Pre-processing suite and the central NFS service. Most of the time, DA and PP run on one machine, the NFS service on another, and the third machine is the backup host in case of problems with either of the other two.

Making the Data Acquisition and Pre-processing software highly available relies on several different components.

The first is the HP software called ServiceGuard, which runs on all three machines. This software monitors the health of the cluster and will restart packages on the backup machine if needed, e.g. if the machine running the pre-processing software crashes, then the backup machine will take over.

Next, the file space needed to store the programs and data for the DA and PP suite is kept on a filesystem on a shared RAID disk device. The RAID disk unit provides protection against a single disk failure by using extra parity information and a hot spare disk, which is automatically activated if necessary. The RAID unit is connected to all of the hosts in the cluster and can thus be accessed by whichever host is running the pre-processing software.

Finally, the suite of scripts that define DA and PP contains the logic necessary to be executed when the package is started on any host. This suite is controlled by ECMWF's supervisor packages (SMS) from one of the operators' workstations. When the package is started it notifies SMS, which will restart any of the jobs as necessary.

Summary

In general, the new Data Acquisition and Pre-processing system is robust and flexible software, suitable now for further development and relatively easy extension.

Milan Dragosavac, Stuart Mitchell

Table of TAC Representatives, Member State Computing Representatives and Meteorological Contact Points

Member State	TAC Representative	Comp. Representative	Met.Contact Point
Belgium	Dr. W. Struylaert	Mrs.L.Frappez	Dr. J.Nemeghaire
Denmark	Dr. P. Aakjær	Mr. N.Olsen	Mr. G.R.Larsen
Germany	Dr. B.Barg	Dr. B.Barg	Dr. D. Meyer
Spain	Mr. T. Garcia-Meras	Mr. E.Monreal Franco	Mr. F. Jiminez
France	Mr. D. André	Mrs.M.Pithon	Mr. D. André
Greece	Mrs.M.Refene Dr. G.Sakellarides	Dr. G.Sakellarides	Mrs.M.Refene Dr. N.Prezerako
Ireland	Mr. J.Logue	Mr. L.Campbell	Mr. M.R. Walsh
Italy	Dr. S.Pasquini	Dr. G. Tarantino	Dr. G.Maresca
Yugoslavia*			
Netherlands	Mr. S.Kruizinga	Dr. J.W. de Vries	Mr. G.Haytink
Norway	Mr. K.Björheim	Mrs.R.Rudsar	Mr. P. Evensen
Austria	Dr. G.Wihl	Dr. G.Wihl	Dr. H.Gmoser
Portugal	Mrs.I.Barros Ferreira	Mr. C.M. Fernandes	Mrs.I.Barros Ferreira
Switzerland	Mr. H.Müller	Mrs. C. Raeber	Mr. P. Roth
Finland	Mrs.K.Soini	Mr. T. Hopeakoski	Mr. P. Nurmi
Sweden	Mr. L.Moen	Mr. S.Orrhagen	Mr. O. Åkesson
Turkey	Mr. M.Selim Yasa	Mr. M.Selim Yasa	Mr. M.Selim Yasa
United Kingdom	Dr. A.Dickinson	Dr. A.Dickinson	Dr. M.J.P. Cullen

* At its 37th Session (December 1992) the Council decided that the telecommunications link between ECMWF and Belgrade would be terminated with immediate effect, and that henceforth no ECMWF documents would be sent to the Federal Republic of Yugoslavia (Serbia and Montenegro).

ECMWF Calendar 1997

2 - 3 Jul	Council	*46th*
25 Aug	ECMWF Holiday	
8 - 12 Sep	Seminar: Atmosphere-surface interaction	
29 Sep - 1 Oct	Scientific Advisory Committee	*26th*
6 - 7 Oct	Computing Representatives' meeting	
8 - 10 Oct	Technical Advisory Committee	*25th*
13 Oct	Policy Advisory Committee	*8th*
15 - 16 Oct	Finance Committee	*58th*
20 - 22 Oct	Workshop: Predictability	
10 - 12 Nov	Workshop: Orography	
17 - 21 Nov	Workshop on Meteorological Operational Systems	*6th*
2 - 3 Dec	Council	*47th*
24 - 26 Dec	ECMWF Holiday	

ECMWF Publications

Technical Memoranda

No. 233 Uppala, S.: Performance of TOVS data in the ECMWF reanalysis 1979-1993, April 1997

No.235 Dent, D., G-R Hoffmann, P.A.E.M. Janssen and A.J. Simmons: Use of the Fujitsu VPP700 for weather forecasting at ECMWF, May 1997

ECMWF Forecast and Verification charts to 31 March 1997