In this issue

Editorial....1

Change of telephone number....2

Removal of old Internet domain name....2

METEOROLOGICAL

Changes to the operational forecast system....2

Physics on adjoint models....2

Summary of Technical Memorandum No. 216....6

Summary of Technical Memorandum No. 217....8

EPS verification....9

COMPUTING

MARS....15

High-Availability server....19

36 track cartridges....20

Efficient use of MARS/ecfile....21

GENERAL

Table of representatives....22

ECMWF Calendar 1996....23

ECMWF Publications....23

This Newsletter is edited and produced by User Support and is designed and printed by ECMWF.

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

Cover

Reliability diagram for day 6 probability forecasts of 850 hPa temperature cold anomalies of more than 4 degrees, spring 1995. See article on Ensemble Prediction System Verification (pages 9-15).

Editorial

An active area of research on modelling and assimilation is the inclusion of physical processes in linear forward and adjoint models. Preliminary results of recent work at ECMWF are presented in the first article in this issue.

Two Technical Memoranda have been written on recent developments in the use of satellite data. Memo 216 by Phalippou describes a newly developed scheme which retrieves humidity, surface wind speed and cloud liquid water path from SSM/I radiances through a one-dimensional variational analysis. Comparisons with the equivalent model first guess fields are presented. Memo 217 by Gaffard and Roquet describes the impact of the ERS-1 scatterometer data on the analysis both in the southern hemisphere and northern Pacific and shows some positive impacts in the short-range forecasts. Summaries of these two Technical Memoranda are on pages 6 and 8.

ECMWF has been running an Ensemble Prediction System (EPS) for some years (see ECMWF Newsletter 65, pp 3-15, 1994). Recent EPS verifications are discussed on page 9.

One of the major services offered by ECMWF is an archive of meteorological data, covering observations, operational forecasts, research forecasts, re-analysed data, etc. Because of the growth both in size and the variety of data it handles, the present system has had to be redesigned internally to cope. An overview of the redesigned system is on pages 15 -19.

In a related issue the Centre has recently changed its data storage medium to 36 track cartridges (see page 20), while on page 21 are some hints on how best to use the various archive facilities without adversely impacting the service.

With more and more reliance being placed on the workstations and servers, the Centre has recently purchased a high-availability system to improve their reliability. The article on page 19 provides an overview of the chosen system.

Finally, note that the Centre's telephone number is changing, again! See page 2 for details.

Change of telephone number - advanced warning -

ECMWF's telephone number is changing:
Current number: +44 1734 499000
New number: +44 118 949 9000.
In more detail the changes are
- Reading area code changes from 1734 to 118
- all existing Reading numbers have 9 added at the front, e.g. 499000 becomes 949 9000.

There are similar changes to all direct dial numbers (direct dial to staff extensions), and to the Centre's fax number.

The new number can be used now. The old number will be withdrawn 1 January 1998.

Removal of old Internet domain name

ECMWF's original Internet domain name was "ecmwf.co.uk". About 18 months ago this was replace with "ecmwf.int", but with the original name still available. It has now been decided to terminate the original name as of 1 July 1996, hence please note that as of this date "ecmwf.co.uk" will no longer be valid.

Changes to the operational forecasting system

Recent changes

On 4 March 1996 the generation of initial perturbations for the Ensemble Prediction System (EPS) was changed to be based on singular vectors in the southern hemisphere as well as the northern hemisphere.

On 25 March 1996 the amplitude of the initial perturbations for the EPS was increased by approximately 15%.

On 23 April 1996 a new definition of the sea surface temperature and sea ice was introduced. The input to the scheme is the 1 degree SST analysis from NCEP Washington and the gridded ice fields derived from SSM/I data from NESDIS.

Planned changes

A comprehensive observation screening and quality control will be implemented in the 3D-Var assimilation system.

Bernard Strauss

On the inclusion of physical processes in linear forward and adjoint models. Impact of large-scale condensation on singular vectors.

1. Introduction

Adjoint equations have been applied in meteorology to solve, among others, problems of data assimilation (*Marchuk*, 1974; *Lewis and Derber*, 1985; *Talagrand and Courtier*, 1987), sensitivity analysis (*Cacucci*, 1981), instability analysis (*Lorenz*, 1965; *Farrell*, 1982; *Lacarra and Talagrand*, 1988), and Kalman filtering (*Epstein*, 1969).

In data assimilation, the central part of a variational assimilation system is the minimization of a cost-function measuring the departure between model variables and corresponding information (observations, previous forecast). Given the dimension of the model state ($\sim 10^6$), the minimization requires the adjoint of a linear forward version of the model. In fact, the adjoint model provides the gradient of the cost-function with respect to the initial state, a necessary quantity for the iterative minimization algorithm. So far, the adjoint model developed at ECMWF does not include realistic physical processes, apart from a simple vertical diffusion and surface drag scheme acting as a momentum sink (*Buizza*, 1994). The lack of physical processes is not too detrimental at mid-latitudes, where most of the evolution of the flow can be interpreted in terms of dynamic instabilities, but the use of the current adiabatic adjoint model is certainly more critical in tropical regions. As a first step of system upgrading, the linear forward and adjoint versions of a large-scale condensation scheme have been implemented in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System.

In instability studies, as for data assimilation because of the large system dimension, linear forward and adjoint model versions are needed to compute perturbations with the fastest growth over finite time intervals (*Buizza and Palmer*, 1995). At ECMWF, these type of perturbations, named singular vectors, are used to construct the perturbed initial conditions of the Ensemble Prediction System (*Molteni et al*, 1996). Again, the inclusion of the most relevant physical processes in the linear forward and adjoint model versions is essential to compute meteorologically relevant perturbations. [For example, the inclusion of the simplified vertical diffusion and surface drag scheme avoided the computation of non-meterological structures.]

In this work, first the general problem of the inclusion of physical processes in linear forward and adjoint model versions is discussed, and then preliminary results on the impact of the inclusion of large-scale condensation processes on the singular vectors' characteristics are reported. Specifically, after this Introduction, Section 2

discusses the problem of linearity of physical processes and its implication especially for data assimilation. Section 3 describes the change in amplification rates and structures induced on the singular vectors by the inclusion of a large-scale condensation scheme in the linear forward and adjoint model. Conclusions are drawn in Section 4.

2. Data assimilation and physical processes

Meteorological data assimilation systems attempt to provide the best initial conditions for numerical weather prediction models. This is achieved by taking into account in an optimal way all the available information describing the state of the atmosphere at a given time: observations, model forecast, climatology, balance criteria. Variational assimilation systems, like the 3D-VAR system currently operational at ECMWF (*Anderson et al.* 1996), can easily include observations of different origins (radiosondes, surface data, drifting buoys, satellite data) and related non-linearly to model variables (satellite radiances, scatterometer wind data).

At mid-latitudes, the validity of the geostrophic approximation allows the definition of dynamical constraints to produce a consistent analysis of the mass and wind fields (balance equation, projection of the flow on slow Rossby modes). These constraints are no longer valid in the tropics where the dynamics tends to be in equilibrium with diabatic processes such as radiation and deep convection. The lack of physical constraints on dynamical fields (mostly on the divergent part of the flow) and on thermo-dynamical variables (temperature and moisture) is evidenced in data assimilation systems by the so-called spin-up problem. The hydrological cycle (precipitation, evaporation, clouds) is generally not in equilibrium with the model at initial time. Therefore, during the first hours of the forecast, evaporation and precipitation increase or decrease in order to achieve a quasi-balance.

Future satellite systems like the Tropical Rainfall Measurement Mission (TRMM) will provide high resolution measurements of precipitable water, rainfall rates and cloud cover in sparse data areas like the tropical oceans. But these new observations will only be useful for data assimilation if they can be compared with model counterparts given by physical parametrization schemes (sub-grid scale processes).

Physical processes are characterized by strong non-linearities, as transitions from unstable to stable regimes in the planetary boundary layer, and on/off switches (e.g., deep convection is activated when moisture convergence is positive and a low level vertical instability is diagnosed). This means that the range of validity of the tangent linear approximation can be reduced by including physical parametrizations in linear forward/adjoint models. In data assimilation, for example, this can cause convergence problems of the minimization algorithm (wrong estimate of the gradient of the cost-function, existence of multiple minima).

Very few studies have been undertaken to include physical parametrizations in linear/adjoint models. However, their preliminary results are quite encouraging.

In data assimilation, depending on the model and the physical parametrizations involved, convergence can be obtained either without particular treatment of the thresholds or with some regularization (*Zou et al.*, 1993; *Zupanski and Mesinger,* 1995). Still, the proof that the additional constraints provided by physical processes can actually improve the estimate of the initial state of numerical weather prediction models remains to be given.

In instability studies, *Errico and Ehrendorfer* (1995) have shown that the inclusion of moist processes in the computation of the singular vectors of a limited area model leads to faster growth, and to an increased number of phase-space directions along which perturbations may grow.

At ECMWF, we have devised a strategy to include simplified physical parametrizations in order to represent the most important diabatic sources and sinks that occur in the real atmosphere (large scale condensation, deep convection, radiation, clouds, turbulence, sub-grid scale orographic effects). For example, moist processes (clouds, large scale condensation) are represented using the former operational diagnostic schemes, and the long-wave part of the radiation scheme has been linearized. Preliminary sensitivity studies have shown that the range of validity of the tangent linear approximation can be improved when thresholds present in the parametrizations are smoothed.

Although the upgrading of the linear forward and adjoint model versions has not yet been applied in data assimilation, it is thought that it should improve the current system for the reasons discussed hereafter.

3. Singular vectors and large-scale condensation

In the first part of this section, the methodology applied to compute the singular vectors is described, while some preliminary results on the impact of large-scale condensation processes on singular vectors is described in the second part.

3.1 Methodology

The generation of effective perturbations is one of the major problems in Ensemble Forecasting. They are obtained at ECMWF by perturbing the initial conditions along the most unstable directions of the phase space of the system: the singular vectors.

The solution $\boldsymbol{x}$ of the model equations linearized about a non-linear trajectory:

$$\frac{\partial x}{\partial t} = A_1 x \qquad (1)$$

may be written in the form:

$$x(t) = L(t_0, t)\, x_0 \qquad (2)$$

where $\boldsymbol{L}$ is the linear forward propagator, the time interval t-t_0 is named optimisation time interval, and $\boldsymbol{A_1}$ is an approximation of the tangent linear version of the

ECMWF model ***A***. Specifically, the (linear) forward propagator is the product of operators representing large-scale condensation ($\boldsymbol{L_{LSC}}$), vertical diffusion and surface drag ($\boldsymbol{L_{vd/sd}}$), adiabatic processes ($\boldsymbol{L_{dyn}}$) and non-linear normal mode initialization ($\boldsymbol{L_{NNMI}}$),

$$L \equiv \left(\prod_{t_0}^{t} L_{LSC} L_{vd/sd} L_{dyn}\right) L_{NNMI} \tag{3}$$

Denote by *(..;..)* an inner product based on total energy,

$$\begin{aligned}(x;y) \equiv \frac{1}{2}\int_0^1\int_s (&\nabla\Delta^{-1}\zeta_x \bullet \nabla\Delta^{-1}\zeta_y + \nabla\Delta^{-1}D_x \bullet \nabla\Delta^{-1}D_y \\ &+ \frac{C_p}{T_r} T_x T_y + w_q \frac{{L_c}^2}{C_p T_r} q_x q_y)\, d\Sigma \left(\frac{\partial\rho}{\partial\eta}\right) \partial\eta \\ &+ \frac{1}{2}\int_s R_d T_r P_r \ln(\pi_x)\ln(\pi_y)\,\partial\Sigma\end{aligned} \tag{4}$$

where ζ stands for vorticity, ***D*** for divergence, ***T*** for temperature, ***q*** for specific humidity, π for surface pressure, C_p is the specific heat of dry air at constant pressure, L_c is the latent heat of condensation at 0ºC, R_d is the gas constant for dry air, T_r=300ºK is a reference temperature, and P_r=80000 *Pa* is a reference pressure. Note that the parameter w_q defines the relative weight given to the specific humidity term.

The total energy of a perturbation ***x*** at time *t* can be computed as

$$\|x(t)\|^2 = (PLx_0; PLx_0) = (x_0; L^* P^2 L x_0) \tag{5}$$

where ***L**** is the adjoint of the linear propagator ***L*** with respect to the total energy scalar product *(..;..)*, and ***P*** is the self-adjoint local projection operator (*Buizza and Palmer*, 1995). The application of the local projection operator permits the identification of singular vectors characterized by maximum growth over a selected region such as northern hemisphere (NH) above 30ºN.

The singular vectors $\boldsymbol{v_j}$ of the propagator ***PL*** are the perturbations with maximum localized energy growth over the optimisation time interval *t*. They are computed solving the eigenvalue problem

$$L^* P^2 L v_i = \sigma_i^2 v_i \tag{6}$$

Specifically, they are the eigenvectors $\boldsymbol{v_j}$ with maximum eigenvalues of the operator $\boldsymbol{L^*P^2L}$. The square root of an eigenvalue, s_j, is named singular value.

3.2 Impact of large-scale condensation processes on singular vector structures

Singular vectors have been computed for a case study, 5 December 1994, at T42 spectral triangular truncation, with 19 vertical levels, with a 48-hour optimisation time interval, in four configurations, either including or excluding the large-scale condensation operator $\boldsymbol{L_{LSC}}$, and assigning two different weights w_q to the specific humidity term of the total energy norm (Table 1).

We analyze first the impact of the weight parameter w_q on the singular vectors. The comparison of the NOLSC.0 and NOLSC.1 singular vectors shows that, when the large-scale condensation scheme is not used,

Experiment	LSC scheme	w_q	σ_1
NOLSC.0	*off*	0.	18.6
LSC.0	***on***	0.	31.7
NOLSC.1	*off*	1.	19.4
LSC.1	***on***	1.	35.4

Table 1. Experiments list, with the growth rate of the leading singular vector.

the inclusion of the specific humidity term in Eq. (4) has a very small impact on the growth rates (Figure 1). A measure of the impact on the singular vector structures is given by the similarity index of the unstable sub-spaces generated by the first 16 singular vectors, which measures the degree of parallelism between the two sub-spaces (two parallel unstable sub-spaces have similarity index 1.00, while two orthogonal sub-spaces has similarity index 0). Table 2 shows that the similarity index between the NOLSC.0 and NOLSC.1 is 0.91, indicating a high degree of parallelism. Similar conclusions can be drawn when considering experiments LSC.0 and LSC.1, although a slightly stronger impact on the growth rates and a slightly higher parallelism between the unstable sub-spaces can be detected.

We now focus on the experiments with w_q=1, and examine the impact of the large-scale condensation scheme. Figure 1 shows that large-scale condensation processes increase the growth rates, on average, by 50%, with peaks of the order of 80% for the leading singular vectors. Considering the singular vector structures, the scheme induces a shift of the total energy spectra, at optimisation time, towards higher wave numbers (Figure 2). Consequently, the structural change corresponds to a substantial reduction of the unstable sub-spaces' parallelism (Table 2). (For reference, the similarity index between the unstable sub-spaces generated by the first 16 singular vectors computed for consecutive starting dates and 48-hour optimisation time interval, is about 0.30.)

t=0	NOLSC.0	NOLSC.1	LSC.0	LSC.1
NOLSC.0	1.00	0.91	0.77	0.73
NOLSC.1		1.00	0.75	0.75
LSC.0			1.00	0.96
LSC.1				1.00

Table 2. Similarity indices computed among unstable sub-spaces of different experiments.

It is worth mentioning the result of a more complete analysis of the sensitivity of the singular vectors to the weight w_q, in the range $0.01 \le w_q \le 100$.

If $w_q << 1$, at initial time the singular vectors have very large specific humidity component, and if the large-scale condensation scheme is activated, growth factors much higher than LSC.1 can be produced through latent heat release. By contrast, if the large-scale condensation scheme is not active the impact on the singular vectors is almost negligible.

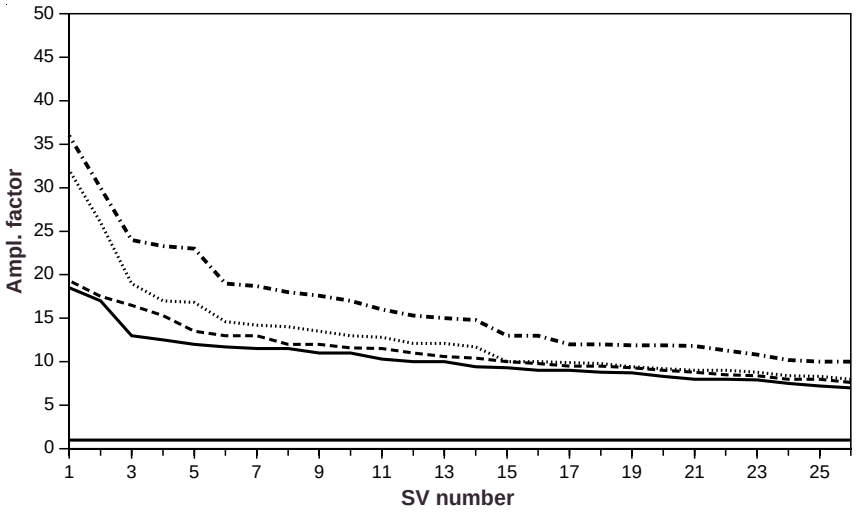

Figure 1. Growth rates of the first 26 singular vectors of experiments NOLSC.0 (solid), NOLSC.1 (dashed), LSC.0 (dotted) and LSC.1 (chain-dashed).

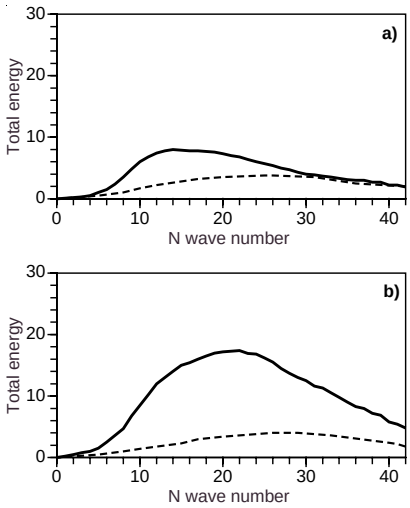

Figure 2. Mean total energy spectra at initial (dashed; increased by a factor of 100) and optimisation (solid) time, averaged among the first 16 (a) NOLSC.1 and (b) LSC.1 singular vectors.

If $w_q >> 1$, instead, at initial time the singular vectors are characterized by very small specific humidity components, and thus the activation of the large-scale condensation scheme has a very small impact. Nevertheless, growth rates much higher than LSC.1 can be produced simply by increasing the singular vectors' specific humidity component. This is achieved by changing the singular vectors' wind field $\boldsymbol{v'}$, so that the singular vectors' specific humidity $\boldsymbol{q'}$ is increased by advection of basic state specific humidity $\boldsymbol{q}$, essentially through a contribution during the linear evolution of the term $\boldsymbol{v' \bullet \nabla q}$.

4. Conclusions

In the first part of this work we have discussed the rationale of including physical processes in linear forward and adjoint model versions. As a first step of the upgrading of the ECMWF Integrated Forecasting System, which is used for data assimilation and instability analysis, a large-scale condensation scheme has been introduced in the forward and adjoint models. Some preliminary results of the impact of the scheme on the singular vectors' structures have been discussed in the second part of this work. They confirm the results of *Errico and Ehrendorfer* (1995), that higher amplification rates can be achieved through latent heat release. Moreover, they indicate that the inclusion of large-scale condensation processes induces a shift of the singular vectors' total energy towards larger wave numbers.

Future work is aimed to a more complete analysis of the impact of the scheme on the singular vectors' structures, especially in regions with high condensation rates. Moreover, non-linear integrations of singular vectors computed with the large-scale condensation scheme activated will be compared with linear evolutions, up to optimisation time, to verify the degree of validity of the linear approximation.

Jean-François Mahfouf and Roberto Buizza

References

Anderson E, Courtier P, Gaffard C, Haseler J, Rabier F, Unden P, and Vasiljevic D, 1996: 3D-Var - the new operational analysis scheme. *ECMWF Newsletter 71,* 2-5.

Buizza, 1994. Sensitivity of optimal unstable structures. *Q. J. R. Meteorol. Soc.*, **120**, 429-451.

Buizza, R, and Palmer, T N, 1995. The singular vector structure of the atmospheric general circulation. *J. Atmos. Sci.*, **52**, 9, 1434-1456.

Cacucci, D G, 1981. Sensitivity theory for non-linear systems. I: nonlinear functional analysis approach. *J. Math. Phys.*, **22**, 2794-2802.

Epstein, E S, 1969. Stochastic dynamic prediction. *Tellus*, **21**, 739-759.

Errico, R M, and Ehrendorfer, M, 1995. Moist singular vectors in a primitive-equation regional model. Preprints of the American Meteorological Society *Tenth Conference on atmospheric and oceanic waves and stability*, 5-9 June 1995, Montana, 272pp.

Farrell, B F, 1982. The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, **39**, 1663-1686.

Lacarra, J F, and Talagrand, O, 1988. Short range evolution of small perturbations in a barotropic model. *Tellus*, **40A**, 81-95.

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

Lorenz, E N, 1965. A study of predictability of a 28-variable atmosphere model. *Tellus*, **17**, 321-333.

Marchuk, G I, 1974 (Russian version 1967). *Numerical methods in weather prediction*. Academic Press, New-York, 277 pp.

Molteni, F, Buizza, R, Palmer, T N, and Petroliagis, T, 1996. The ECMWF Ensemble Prediction System: methodology and validation. *Q. J. R. Meteorol. Soc.*, **122**, 73-119.

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

Zou, X, Navon I M, and Sela J G, 1993: Variational data assimilation with moist threshold processes using the NMC spectral model. *Tellus,* **45A**, 370-387.

Zupanski, D, and Mesinger, F, 1995. Four-dimensional variational assimilation of precipitation data. *Mon. Wea. Rev.*, **123**, 1112-1127.

Summary of ECMWF Technical Memorandum 216

Variational retrieval of humidity profile, wind speed and cloud liquid water path with SSM/I: potential for numerical weather prediction

L Phalippou

The Special Sensor Microwave Imager (SSM/I) is a seven channel radiometer which operates at 19.35, 22.235, 37., 85.5 GHz in vertical and horizontal polarizations except at 22.235 GHz where only the vertical polarization is available (*Hollinger*, 1990). Several atmospheric parameters can be retrieved from SSM/I observations (*Hollinger*, 1991). Among them, the total precipitable water vapour (TPW), the surface marine wind speed, the cloud liquid water path (CLW) and sea-ice are of primary interest for Numerical Weather Prediction (NWP). These SSM/I products have been used for the validation of NWP fields (see for instance *Liu et al.*, 1992; *Phalippou*, 1992 and *Tiedtke*, 1993). Assimilation experiments of SSM/I total precipitable water vapour and wind speed have already been made in Global Circulation Models at the National Meteorological Centre (NMC) (*Wu and Derber*, 1994), at NASA Goddard Space Flight Centre (*Bloom and Atlas*, 1991) and at the European Centre for Medium-range Weather Forecasts (ECMWF) (*Filiberti*, 1993). The results are encouraging, and *Wu and Derber* (1994) and *Derber et al.* (1994) envisage assimilating SSM/I total precipitable water vapour and wind speed into the NMC operational data assimilation system.

After the launch of the first SSM/I on board the F-8 DMSP satellite, SSM/I total precipitable water vapour, wind speed and cloud liquid water path were retrieved separately with regression algorithms (*Hollinger*, 1991). Several problems were found in the products derived from these algorithms, and many 'improved' regression algorithms have flourished afterwards. However, there is one critical point common to any of these regression methods: their performances are limited by a simplified handling of non-linearities and/or by the poor quality of the *a priori* information. This problem is particularly critical when clouds are present in the field of view (FOV) as the retrieval problem becomes highly non-linear and more ill-posed in cloudy conditions. Recently *Wentz* (1992) and *Greenwald et al.* (1993) used *a priori* information coming from climatology and a simplified radiative transfer model for inverting SSM/I radiances. The weakness of their approach is that the errors in the *a priori* information, in the radiances and in the radiative transfer model are not taken into account in the retrieval process.

Variational methods offer an alternative approach for the estimation of the geophysical parameters from SSM/I observations. The geophysical parameters are estimated simultaneously following the theory of optimal estimation, and are therefore the best set of parameters that explain the observed radiances, while being consistent with the available *a priori* information. In addition, an estimation of the errors of the 'retrieved' products can be computed and the quality control is easier. The variational method is used at ECMWF for the operational preprocessing and quality control of data from TIROS Operational Vertical Sounder (TOVS) and is called 1D-Var as the temperature and humidity profile are retrieved along the vertical (*Eyre et al.* 1993).

The ECMWF Technical Memorandum No. 216 (Phalippou, 1996) presented the theory, the implementation and the results of a 1D-Var method for the estimation of the humidity profile, the wind speed and the cloud liquid water path over oceans, using SSM/I observations and ECMWF forecast fields. The method is based on the minimization of an objective cost function and is mathematically similar to the data assimilation problem in NWP. The *a priori* information comes from the ECMWF first guess fields, the forward operator i.e. the radiative transfer model, their associated covariance error matrices and the observation errors. The main purpose of this study was to demonstrate the feasibility of the method with the present knowledge of the radiative transfer modelling and to show its potential for NWP.

The TPW field retrieved with 1D-Var compares well with TPW retrieved from the regression algorithm (Alishouse et al., 1990). However, our results suggest that the Alishouse TPW exhibit local biases. The present method has two advantages for the retrieval of TPW. That approach is fully non-linear and no approximation of the radiative transfer is needed. The use of the humidity profile as control variable, the use of *a priori* information and the control of supersaturation allow us to distribute the humidity increments in an optimal way. The theoretical accuracies of the 1D-Var TPW are estimated to lie between 1 to 2 Kgm^{-2} depending on weather conditions, representing an improvement of a factor 2 to 8 on the ECMWF TPW error.

The effects of cloud and wind speed on SSM/I radiances can be similar in some cases. If separate retrievals for wind speed and cloud liquid water path are attempted with no *a priori* information, there might exist ambiguities in the retrieved products. This effect is believed to be the source of the speckled appearance of the wind field retrieved with the regression algorithm of *Goodberlet et al.* (1992).

In the 1D-Var approach, the background wind speed is known with a good accuracy (2 ms^{-1}). This *a priori* knowledge of the wind speed is used as an 'anchored' point that enables us to resolve the wind speed-cloud ambiguity. As a consequence, the wind speed field retrieved with 1D-Var is very close to the ECMWF background. The theoretical standard deviation of 1D-Var retrieved wind speed varies between 1.6 ms^{-1} and 0.5 ms^{-1} for low to high wind speed. This last value is believed to be underestimated due to the very simple approximation used for the sea foam emissivity. However, we are conscious that it is difficult to demonstrate that we improve the accuracy of the wind speed without comparing the 1D-Var retrieved wind speed with independent wind measurements.

The modelling of the sea surface emissivity/scattering has been identified as one of the most critical areas for the success of a 1D-Var method for a SSM/I-like instrument. The sea surface model described in this paper gives sensible results, as no unreasonable feature appears in the retrieved fields.

The present method is believed to be the first attempt, to the knowledge of the author, to retrieve the cloud liquid water path from SSM/I and collocated NWP cloud fields. The main advantage of using a NWP cloud field is that the vertical location of the cloud, the phase mixing (ice and liquid), the temperature and humidity inside the cloud are *a priori* defined in a consistent manner. However, the retrieved fields are almost insensitive to the value of the background CLW. The theoretical CLW errors have been found to vary between 0.02 Kgm^{-2} to 0.1 Kgm^{-2} depending on atmospheric conditions. The validation of CLW still remains an open question because of the lack of independent accurate ground truth. However, the retrieved humidity profile, wind speed and cloud liquid water path are the most probable solution which explains the observed radiances. This is obviously not checked when regression-type algorithms are used for retrieving each variable separately.

A threshold on the value of the cost function at the end of the minimization has been shown to be a simple and efficient quality control, which allows one to detect cloud ice and precipitation.

Since the publication of this Technical Memorandum, a fast radiative transfer model has been developed and SSM/I data (brightness temperatures) are received from NESDIS on a daily basis. Since 15 March, SSM/I data are processed in real time using the 1D-Var method presented in this Technical Memorandum. The 1D-Var retrieved total precipitable water vapour, wind speed and cloud liquid water path are used for monitoring the quality of analyses/forecasts over ocean.

References

Bloom, S C and R Atlas, 1991: Assimilation of satellite surface wind speed data and its impact on NWP. Proceedings of the 9th AMS Conference on Numerical Weather Prediction. Denver, Colorado, October 1991.

Derber, J C, D F Parrish, W S Wu, Z Pu and S R H Rivzi, 1994: 'Improvements to the operational SSI global analysis system'. Pp 149-150 in Proceedings of the 10th AMS Conference on Numerical Weather Prediction, Portland, Oregon, July 1994.

Eyre, J R, G A Kelly, A P McNally, E Anderson and A Persson, 1993: Assimilation of TOVS radiance information through one-dimensional variational analysis. *Q J R Meteorol Soc*, **119**, 1427-1463.

Filiberti, M A, 1993: Assimilation dans des modèles météorologiques de données en sapeur d'eau integrée mesurées par radiométrie hyperfiquenne spatiale. PhD dissertation, Université Paris 7, UFR de Physique, 304 pp.

Greenwald, T.J., Stephens, G.L., Vonder Haar, T.H. and Jackson, D.L., 1993: A physical retrieval of cloud liquid water over the global oceans using the Special Sensor Microwave/Imager observations. J. Geophys. Res., **98**, D10, 18471-18488.

Goodberlet, M A and C T Swift, 1992: Improved retrievals from the DMSP wind speed algorithm under adverse weather conditions. *IEEE Trans Geosci Remote Sensing*, Vol 30, No 5, 1076-1077.

Hollinger, J, J L Peirce and G A Poe, 1990: SSM/I instrument evaluation. *IEEE Trans Geosci Remote Sensing*. Vol 28, No 5, 781-790.

Hollinger, J, 1991: DMSP special sensor microwave/imager. Calibration/validation. Naval Research Laboratory. Final Report Volume 2. Washington DC.

Liu, W T, W Tang and F J Wentz, 1992: Precipitable water and surface humidity over global oceans from special sensor microwave imager and European Centre for Medium-Range Weather Forecasts. *J Geophys Res*, Vol 97, No C2, 2251-2264.

Phalippou, L, 1992: Comparison between SSM/I and ECMWF total precipitable water. Pp 22-26, Proceedings of the Specialist Meeting on Microwave Radiometry and Remote Sensing Application. E R Westwater Editor. Boulder, Colorado. June 1992.

Phalippou, L, 1995, 1996: Variational retrieval of humidity profile, wind speed and cloud liquid water path with SSM/I: potential for numerical weather prediction. *ECMWF Technical Memorandum No. 216, also in QJR Meteorol. Soc.* Vol. 122, 327-355.

Tiedtke, M, 1993: Representation of clouds in large-scale models. *Mon Wea Rev*, Vol 121, No 11, 3040-3061.

Wentz, F J, 1992: Measurements of oceanic wind vector using satellite microwave radiometers. *IEEE Trans Geosci Remote Sensing*, Vol 30, No 5, 960-972.

Wu, W S and J C Derber, 1994: Inclusion of SSM/I precipitable water observations in the NMC spectral statistical-interpolation analysis system. Proceedings of the 10th AMS Conference on Numerical Weather Prediction, Portland, Oregon, July 1994.

Summary of ECMWF Technical Memorandum 217

Impact of the ERS-1 scatterometer wind data on the ECMWF 3D-Var assimilation system

C Gaffard

The ERS-1 satellite was launched on 17 July 1991 by the European Space Agency (ESA), and is carrying on board a C-band scatterometer. The ERS-1 scatterometer is a radar instrument, measuring the power backscattered towards the satellite by the earth's surface on a 500 km wide swath. It has three antennae pointing in a horizontal plane towards a direction of 45°, 90° and 135° with respect to sub-satellite track. The return signal is sampled every 25 km both along and across-track directions and can, with a good precision, be related to the 10 metre wind vector over sea (*Stoffelen and Anderson*, 1993). The major problem when retrieving wind vectors from ERS-1 scatterometer measurements is the so-called directional ambiguity, which leads most of the time to a first wind solution and a second one 180° apart, for the same backscatter measurement triplet.

ESA has implemented a ground segment, which enables numerical weather prediction centres to get ERS-1 scatterometer data in near real-time with a global coverage. These data are processed by ESA, which provides retrieved 10 metre wind vectors with ambiguity removal, as well as the instrumental parameters, necessary for the wind retrieval. The retrieval and ambiguity removal scheme, used by ESA and called CREO (*Cavanié and Lecomte*, 1987), was implemented at ECMWF, and significantly improved by *Stoffelen and Anderson* (1995), in particular from the ambiguity removal and quality control point of view. This version, which uses the CMOD4 transfer function, has been operational at ECMWF since July 1994 and is called PRESCAT.

An assimilation experiment of ERS-1 scatterometer wind data was performed in ECMWF's 3D-Var system for a 2-week period from 6 to 20 December 1994 (*Gaffard and Roquet,* 1995). In this experiment, the retrieved wind speed given by PRESCAT has been modified according a bias correction obtained using a collocated data set of buoy measurements. This correction removes the former CMOD4 low bias for high wind speeds. Scatterometer wind data which have passed the quality criteria of PRESCAT are assimilated in the 3D-Var system as pairs of ambiguous wind vectors with a 100 km horizontal sampling. The ambiguity removal is done implicitly during the 3D-Var analysis, using all the information from the background and the other observations.

During this assimilation experiment, the implicit ambiguity removal within 3D-Var has proved to work well, giving very similar results to the statistical scheme in PRESCAT. The results of the experiment assimilating scatterometer wind data were compared to the ones of a parallel experiment, without scatterometer wind data. When assimilating scatterometer wind data, the departures between first-guess wind fields and buoy, scatterometer or altimeter winds are decreased. An improvement of the first-guess pressure field is also found. The reduction of the departures from scatterometer and altimeter winds is maximum in the Southern Hemisphere.

The impact of scatterometer wind data on the short and medium range forecasts has been investigated for 1000 hPa geopotential. For short range forecasts, a significant impact is found in the Southern Hemisphere only. Taking the 3D-Var analysis using scatterometer wind data as reference, the scatterometer wind data have a clear positive impact on the 12 hour forecasts in the Southern Hemisphere, a significant part of which is lost in the 48 hour forecasts. For medium range forecasts, and taking the operational analysis as reference, the impact of scatterometer wind data is found slightly positive or neutral whatever the geographical area of the globe, with a maximum positive impact over Northern America. This feature is likely due to the particular meteorological regime during this period, with storms developing in the eastern part of the northern Pacific.

Scatterometer wind data, as other surface observations, are difficult to use in sequential assimilation schemes as optimal interpolation or 3D-Var in which the horizontal and vertical structure functions used for the background errors are predefined and do not depend on the meteorological conditions. To that respect, the 4D-Var assimilation methods are very promising (*Thépaut et al.*, 1993). However, scatterometer wind data can already be beneficial for the analyses and short term forecasts, in particular in the areas where few conventional observations are available. For this reason, the assimilation of scatterometer wind data in the numerical weather prediction systems should also be beneficial for wave and ocean modelling, which are very dependent on the quality of surface wind fields.

References

Cavanié, A and P Lecomte, 1987: "Wind retrieval and dealiasing subroutines", Final report of ESA contract No 6874/87/CP-I, vol. 2, ESA publication.

Courtier, P et al., 1993: "Variational assimilation at ECMWF", ECMWF Research Department *Tech. Mem. 194.*

Long, A, 1985: "Towards a C-band radar sea echo model for the ERS-1 scatterometer", proceedings of the third international Symposium on spectral signatures, pp. 29-34, ESA publication.

Guillaume, A, 1994: "Accuracy of the ERS-1 altimeter-derived Fast Delivery Products for wave data assimilation", ECMWF Research Department *Tech. Mem. 198.*

Hoffman, R, 1993: "A preliminary study of the impact of the C-band scatterometer wind data on global scale numerical weather prediction", *J Geophys Res*, **98** (C6), pp. 10233-10244.

Stoffelen, A, C Gaffard and D Anderson, 1993: "ERS-1 scatterometer data assimilation", proceedings of the second ERS-1 Symposium, Hamburg, Germany, 11-13 October 1993, pp. 191-194, ESA publication.

Stoffelen, A and D Anderson, 1993: "Characterisation of ERS-1 scatterometer measurements and wind retrieval", proceedings of the second ERS-1 Symposium, Hamburg, Germany, 11-13 October 1993, pp. 997-1001, ESA publication.

Stoffelen, A and D Anderson, 1995: "The ECMWF contribution to the characterisation, interpretation, calibration and validation of ERS-1 scatterometer backscatter measurements and winds, and their use in the Numerical Weather Prediction models", ESA contract report.

Thépaut, J-N, R Hoffman and P Courtier, 1993: "Interactions of dynamics and observations in a four-dimensional variational system", *Mon. Wea. Rev.*, **121**, No. 12, pp. 3393-3414.

Verification of the Ensemble Prediction System (EPS)

The use of EPS products as medium-range forecast guidance is becoming more widespread as experience builds up among forecasters and at ECMWF. Systematic evaluation contributes to this process by assessing the quality and usefulness of direct and derived output products. A description of the EPS and of the EPS products can be found in the ECMWF Newsletter 65, 1994, pp 3-15.

In this article various aspects of the EPS verification are presented, mainly based on the ensemble forecasts of temperature at 850 hPa for last winter, December 1995 to February 1996. The spread of the ensemble and the skill of the ensemble members are considered first, then the relationship between ensemble spread and forecast skill is explored. The last section deals with the verification of derived probability products.

Spread of the ensemble

One of the most important requirements to be met by an ensemble forecasting system is that the spread of the ensemble should be sufficient to cover the uncertainties in the forecast, which are due to inaccuracies in the initial conditions but also to model imprecisions.

A necessary condition to be fulfilled is that the observed values should fall into the range of predicted values. More generally, are the distributions of the ensemble members compatible with the distribution of the observations? This can be investigated by counting the number of occurrences of the observed values in the intervals defined by the single ensemble members, at each grid point for the ensemble forecast to be verified. Given that on the long run the ensemble members should all behave with the same characteristics as the real atmosphere, the frequency of occurrence of all intervals should be the same, including the two extreme intervals outside the entire range of ensemble values (method suggested by O. Talagrand, personal communication). Figure 1

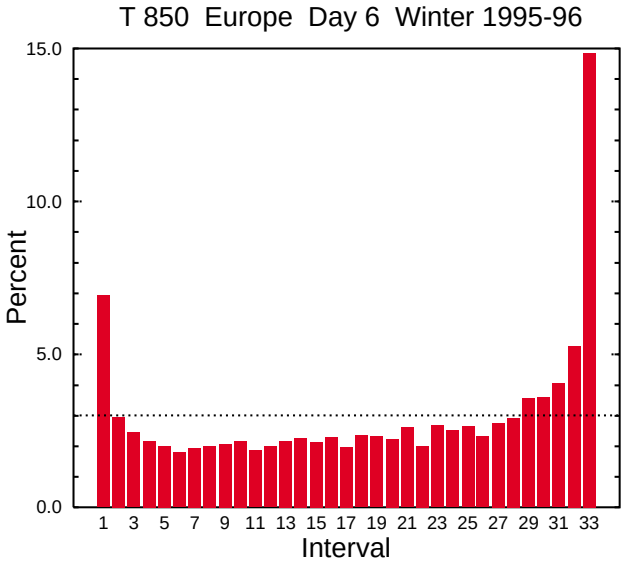

Figure 1. Talagrand diagram for temperature at 850 hPa, ensemble forecasts at day 6, winter 1995-96, over Europe

shows such an evaluation in graphical form, for the day 6 forecast of temperature at 850 hPa over Europe. The distribution shows an excess in the extreme classes, indicating that the ensemble spread is too small to cover all uncertainties.

With an ensemble of 32 members, each class should occur in 100/(32+1) = 3% of the cases (dotted line in figure 1). This applies in particular to the two extreme classes, which means that, even with a "perfect" 32 member ensemble, the analysis will not be captured by the ensemble at about 6% of the grid points. Figure 2 shows the performance of the current ensemble over Europe in this respect. In the range of day 6 to 7, the percentage is about 15 to 20%. This result is complemented by subjective evaluations indicating that the actual evolution of the atmosphere is too often not encompassed by the ensemble.

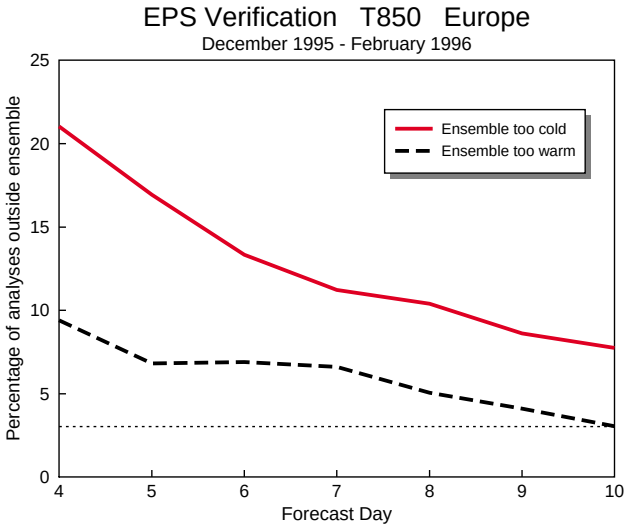

Figure 2. Percentage of grid points where the verifying analysis lies outside the ensemble range

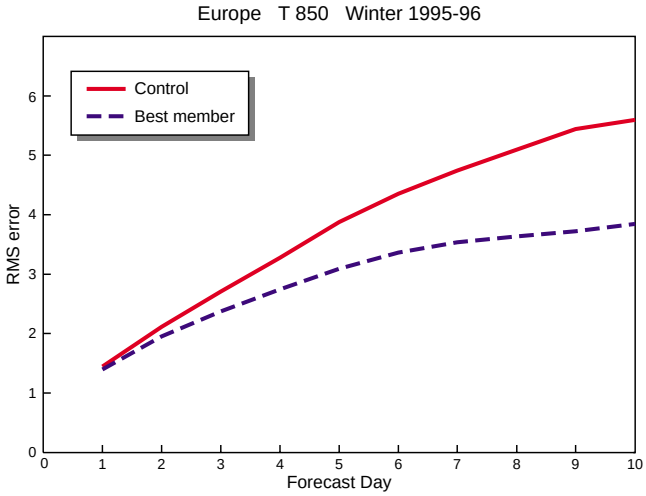

Figure 3. RMS error of the control forecast and of the best ensemble member over Europe

The increases of ensemble size and of model resolution which are planned for the EPS development should improve this aspect of the performance.

It should be noted that this condition on the spread of the ensemble is a necessary condition if the ensemble is to be skilful, but it is not sufficient. An ensemble composed of fields taken at random among analyses for the same month over the past 10 years would provide a correct spread, but yield no skill above climatology.

Skill of the ensemble members

When looking at the synoptic guidance from the EPS, one tends to expect that at least some ensemble members should be significantly more skilful than the control. However, although this expectation looks reasonable, a closer analysis shows that it is not properly formulated, because it does not specify the size of the area on which the ensemble members are supposed to be skilful.

It is not reasonable to expect that some ensemble members should always be correct over the whole globe, the northern hemisphere or even Europe. For example, in a situation with a ridge building up over northern Europe and a cut-off low moving over the east Atlantic, there is no reason to expect that some ensemble members should accurately forecast both the ridge and the cut-off low. Some members should correctly depict the ridge, and some members the cut-off low, but they should not necessarily be the same members. For a synoptic use of the EPS output by forecasters, a proper general requirement would be that ensemble members should be skilful on areas of a size corresponding to the synoptic patterns under consideration.

The importance of the size of the verification area can be shown by comparing the skill of the control forecast with the skill of the best ensemble member over various areas. This is done first for the whole of Europe, figure 3. The RMS errors of the best ensemble member and of the control forecast computed over a 5x5 grid are plotted from day 1 to day 10. At day 1, the best ensemble member was only marginally better than the control, which should be expected with the present system, as the control is based on the best available analysis of the real atmosphere and therefore should yield the best possible forecast in the short (linear) range. In the medium range however, the differences grow rapidly. At day 6 the best ensemble member had typically the same level of skill as a control forecast at day 4.2, and the best day 10 ensemble member was on average equivalent to a day 5 control forecast.

When looking at smaller areas, the difference between the control forecast and the best member increases. Figure 4 displays for the day 6 forecast over the same period the ratio between the best skill and the control skill for areas ranging from the whole of Europe to the grid point closest to Reading. This ratio decreases with the size of the area, from 80% to less than 20% at one grid point. A plot of the verification from day 1 to day 10 for an area of 4x4 grid points, or 15x15 (figure 5) shows that the RMS score of the best ensemble member at day 6 was similar to the score of a day 3 control.

It should be noted that the best member of any ensemble, even a trivial one, will display a reduction of RMS error when the size of the area decreases. The implication is important for both the verification and the use of the EPS: the ensemble should not be regarded as a collection of single global forecasts which should be as accurate as possible. It should rather be treated as providing statistical (probabilistic) information on individual patterns or grid points.

Relationship between ensemble spread and forecast skill

An essential requirement for the practical usefulness of the ensemble is that there should be a correlation between the spread of the ensemble and the deterministic skill of the individual forecasts. When the spread of the ensemble is small, i.e. when most of the ensemble members produce similar forecasts, these forecasts

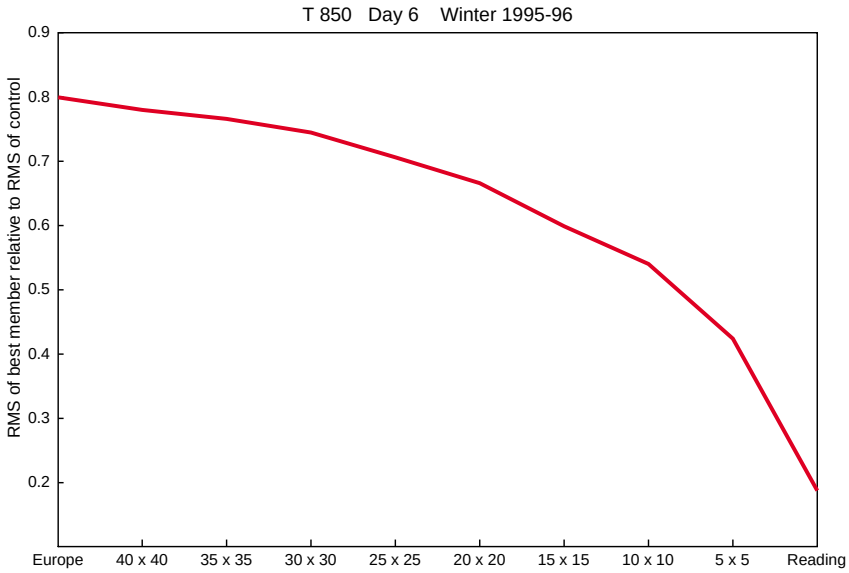

Figure 4. Ratio of the RMS error of the best ensemble member to the RMS error of the control at day 6 for areas ranging from Europe to the grid square nearest to Reading

should be more skilful than average. This is actually required for any application of the ensemble, including the generation of probabilities.

The relationship between spread and skill is best explored at the grid point level. Figure 6 shows a result for all grid points over Europe last winter. It is based on the comparison between the absolute error of the control forecast and the spread of the ensemble defined as the difference between the largest and the smallest values predicted by the ensemble members at the grid point. For each forecast range a curve is plotted consisting of three points. The values along the x-axis are the lowest decile, median and highest decile of the distribution of spread values at all grid points, the values along the y-axis being the median of the distribution of control absolute errors for that spread category. Looking for example at the situations with large spread at day 6, the figure shows that in 10% of all cases the spread was about 17K or more, and that the median of the control forecast error for these 10% was about 4K. Looking at small spread at the other end of the distribution, 10% of the cases had a spread of 4.5K or less, with a median of control error of 1.8K.

A measure of the effectiveness of the relationship between spread and skill is the reduction of error achieved by the cases with small spread compared to the whole sample (figure 7). For the cases with small spread, the skill at day 6 is the same as the skill of an average forecast at day 3.6. Perhaps more important in practice, a day 10 forecast with small spread can be expected to score like an average day 5 forecast.

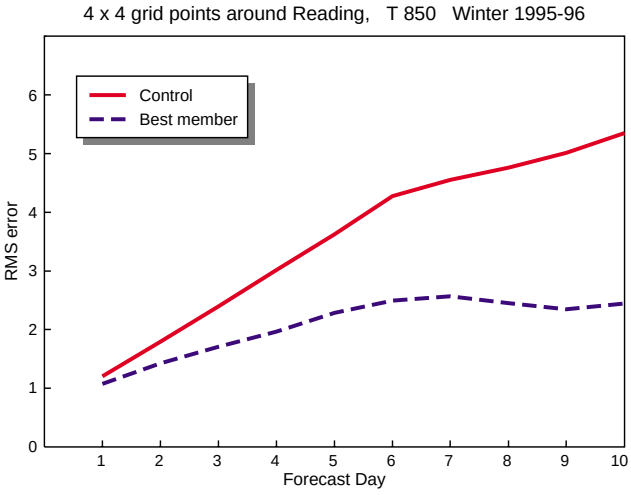

Figure 5. Same as figure 3 for a square of 4x4 grid points around Reading

Derived probability products

One important aspect of the EPS is the forecast of probabilities of meteorological events, e.g. the probability of precipitation to occur in a defined period. Skilful medium-range prediction of probabilities of weather events certainly would be of great value to end users if used adequately in decision making processes. A range of methods to verify probabilistic forecasts exists, and descriptions of techniques and applications are given, e.g., in *Stanski et al.* (1989).

The two main properties of probabilistic forecasts are reliability and sharpness. Reliability indicates the correspondence between forecast probability and the observed frequency of occurrence of an event. It is best depicted in graphical form, in a so called reliability diagram, as shown in figure 8. The reliability curve is constructed by splitting the range of forecast probabilities into intervals (or a set of discrete values as usually in the case of subjective probability forecasts), counting the observed occurrences of the event in all forecast probability classes and plotting the relative frequency of occurrence in every class against the interval centre. For perfect reliability the points in this curve lie on the diagonal. Points below the diagonal indicate that probabilities were over-forecast, points above the diagonal mean under-forecasting. The example in figure 8 is the reliability curve for the day 6 forecast of the event 850 hPa temperature anomaly less than -4 degrees, verified against

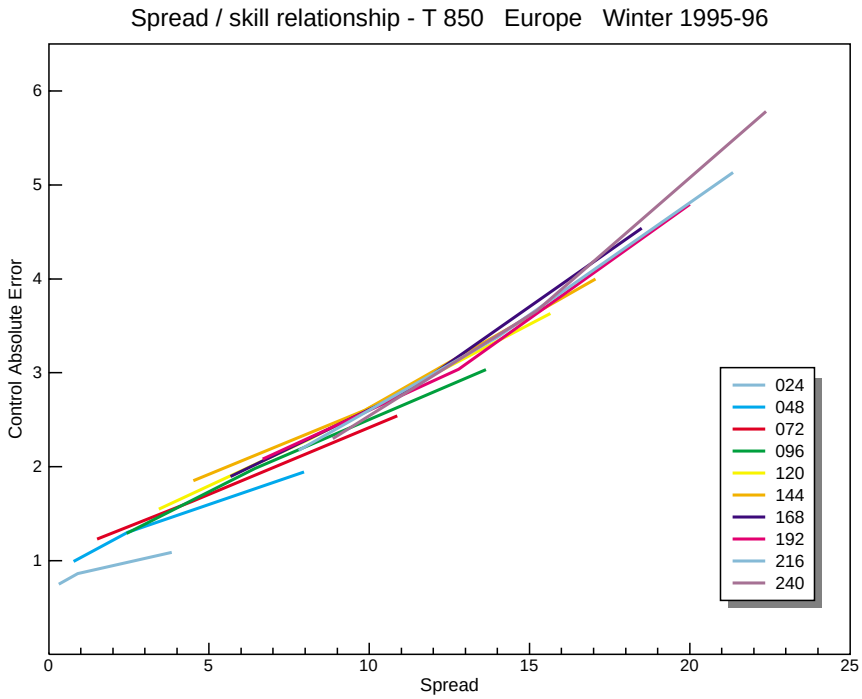

Figure 6. Relationship between ensemble spread and control forecast absolute error (refer to text for explanations)

analyses over Europe, for spring 1995. It shows quite good reliability, however, high probabilities were predicted slightly too frequently, whereas low probabilities were under-forecast. In other words, the EPS was slightly too confident in predicting this event.

The histogram plotted next to the reliability diagram in figure 8 shows the relative distribution of forecasts in probability intervals, which reflects the sharpness of the forecast. A probabilistic forecast is sharp if it often predicts probabilities close to 0% or 100%. A reliable probabilistic system which has minimum sharpness will always predict a probability equal to the sample climate frequency of occurrence of the event (which is a non trivial forecast).

A widely used measure of accuracy of probabilistic forecasts, which takes account of both reliability and sharpness, is the (half) Brier score, BS (Brier, 1950). BS is the mean square error of the probabilistic forecast, whereby the observations have either the value 1.0 for occurrence of the event or 0.0 for non-occurrence. BS does not take into account how close the observed values are to the threshold of the defined event. The score is 0 for a perfect and 1 for the worst possible set of forecasts. Note that these extreme values can only be obtained by a categorical forecast, i e using exclusively probabilities 0 and 1. Since BS is strongly dependent on the sample climatology of the event, comparison between scores from different samples is not very meaningful.

To answer the question of skilfulness of a probabilistic forecast a score obtained by a reference forecast is needed for comparison. A possible reference is a forecast which uses only observed long term climatological frequency as forecast probability. A skill score can be computed which expresses the relative improvement of the forecast against the reference score. Such a skill score is 1 for a perfect forecast, 0 for a probabilistic forecast which is no more accurate than a trivial forecast using long term climatology, and negative for even worse forecasts. In the reliability diagram in figure 8 Brier score and skill score are indicated in the title. BS = 0.123 compares to a climate score BSlcl = 0.184, yielding a skill score of BSS = 0.325, i.e. a 33% improvement in accuracy against a climatology forecast.

In addition, isolines of Brier score are plotted in the diagram to help understand the relation between the position of points in the reliability curve and the contribution of the sub-samples to the overall score. The minima (good BS) are in the bottom left and top right corners, and relatively low values stretch along the diagonal. This illustrates that high reliability contributes towards good overall scores. However, since the overall score is the sum of sub-sample scores weighted by their relative frequency (depicted in the adjacent histogram), also some sharpness, i.e. relative concentration of forecasts at high and low probabilities, is needed to yield a good score. The degree to which sharpness can be realistically achieved depends on the climatological likelihood of the event. Forecasts for very rare events will naturally concentrate strongly in low probability classes, and generally it is quite hard to achieve high skill, as defined here, for such events.

In figure 9, the evolution of reliability with forecast range, for 850 hPa temperature cold anomalies of more than four degrees, verified against analyses, for spring 1995, is depicted. While reliability remains quite good over the entire forecast range, sharpness (see histograms) declines due to increasing ensemble spread. As discussed above, this leads to a reduction in skill score, despite almost identical reliability curves. The 10-day scores show that there is still significant advantage of the EPS over a climate estimate of probabilities. The reliability curves for day 5 warm anomaly forecasts

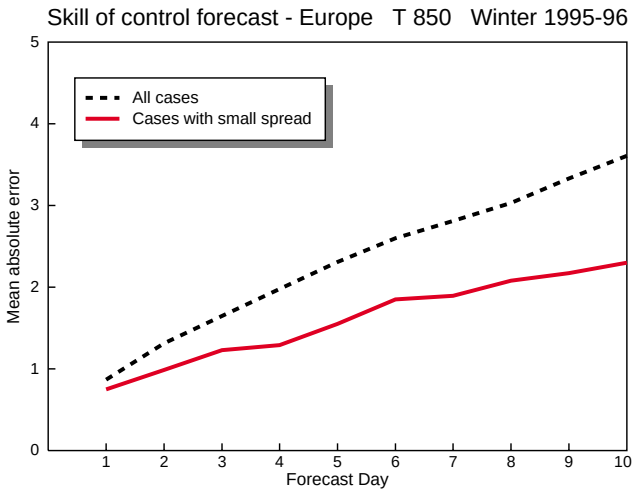

Figure 7. Overall mean absolute error of control forecast (dashed line), and mean absolute error of control forecast over the 10% grid points with the smallest spread

for different seasons in figure 10 show quite good results for all seasons, but especially the winter curve reveals significant under-forecasting of low probabilities. Since most forecasts in this sample lie in this range this can be interpreted as being - at least partly - due to a cold model bias. This is also confirmed by over-forecasting of cold anomalies over the entire range of probabilities above 0.1 in winter 1994/95 (not shown), without "compensating" under- forecasting of very low probabilities.

Reliability curves for precipitation accumulated from t+120 to t+144 above 1 mm, for winter 94/95 and summer 95, are displayed in figure 11. Both curves indicate over-forecasting of high probabilities, especially in summer. Verification here is against SYNOP observations from about 200 stations all over Europe.

Summary

Current evaluation of the Ensemble Prediction System shows that the ensemble forecasts provide non trivial information in the medium range, out to and including day 10. In particular the connection between small ensemble spread and relatively high skill of the control forecast is well established, which has positive implications for the quality of the probability products and for the synoptic use of the ensemble. However, the spread of the ensemble is not satisfactory, as there are too many occurrences where the ensemble does not include the actual evolution of the atmospheric flow.

Results on probability products for weather events show a significant degree of skill, particularly for the probability of temperature anomaly. Reliability varies for different parameters and thresholds, and has also seasonal dependency. A common feature in practically all reliability statistics is over confidence, i.e over-forecasting of high probabilities and under-forecasting of low probabilities. This is partly explained by systematic model errors and by the lack of ensemble spread, which should be addressed by the planned increase in model resolution and in ensemble size.

References

Brier G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.* **78**, 1-3.

Stanski H. R., L. J. Wilson, W. R. Burrows, 1989: Survey of common verification methods in meteorology. Second edition, published as *WMO WWW Rep. No 8.*

Bernard Strauss, Andreas Lanzinger

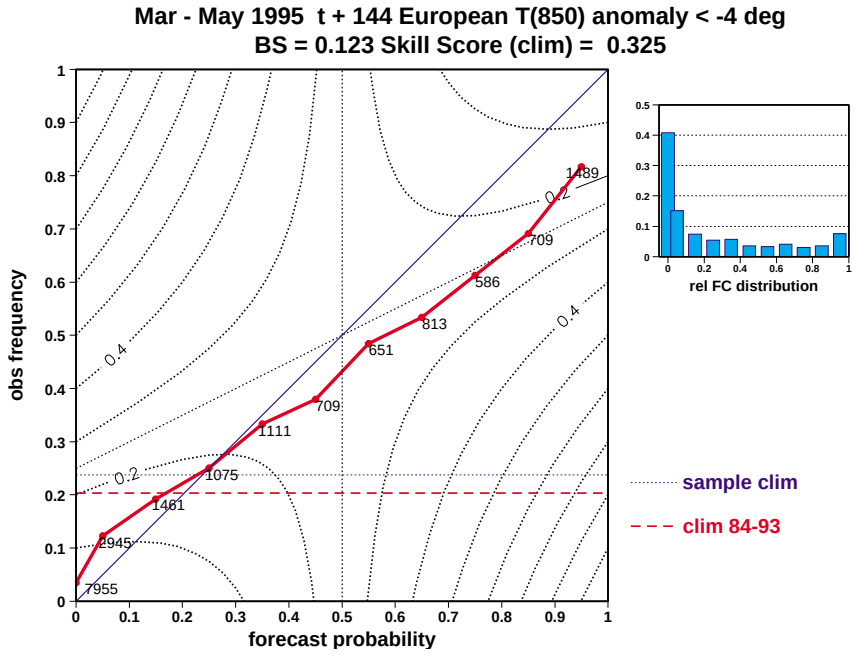

Figure 8. Reliability diagram for day 6 probability forecasts of 850 hPa temperature cold anomalies of more than 4 degrees, spring 1995. Verification against analysis over the European area. Numbers next to reliability points indicate the absolute number of cases (forecasts) in the probability interval. Horizontal lines denote the levels of sample (dotted) and long term (dashed) climatologies. The histogram shows the relative distribution of forecasts in probability intervals (sharpness).

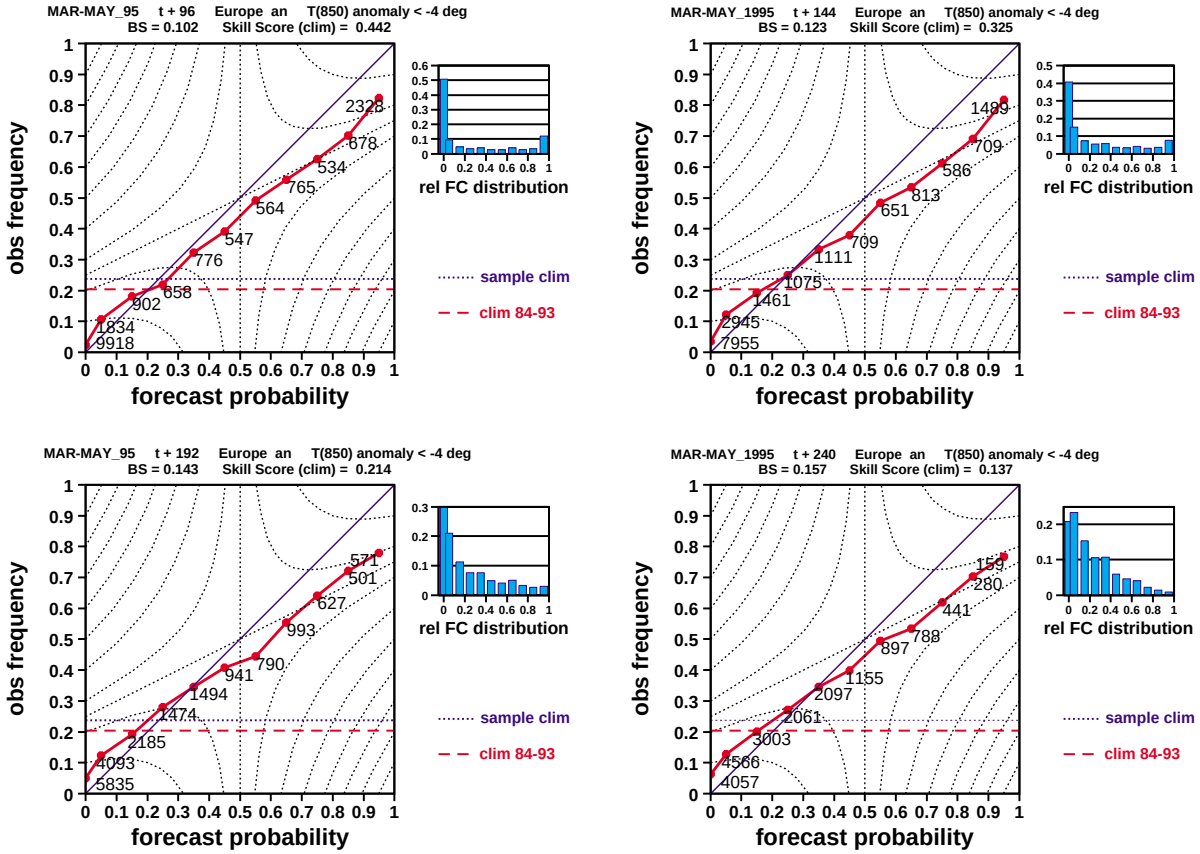

Figure 9. Reliability diagrams for 850 hPa temperature cold anomalies of more than 4 degrees, spring 1995, at day 4 (top left), day 6 (top right), day 8 (bottom left) and day 10 (bottom right)

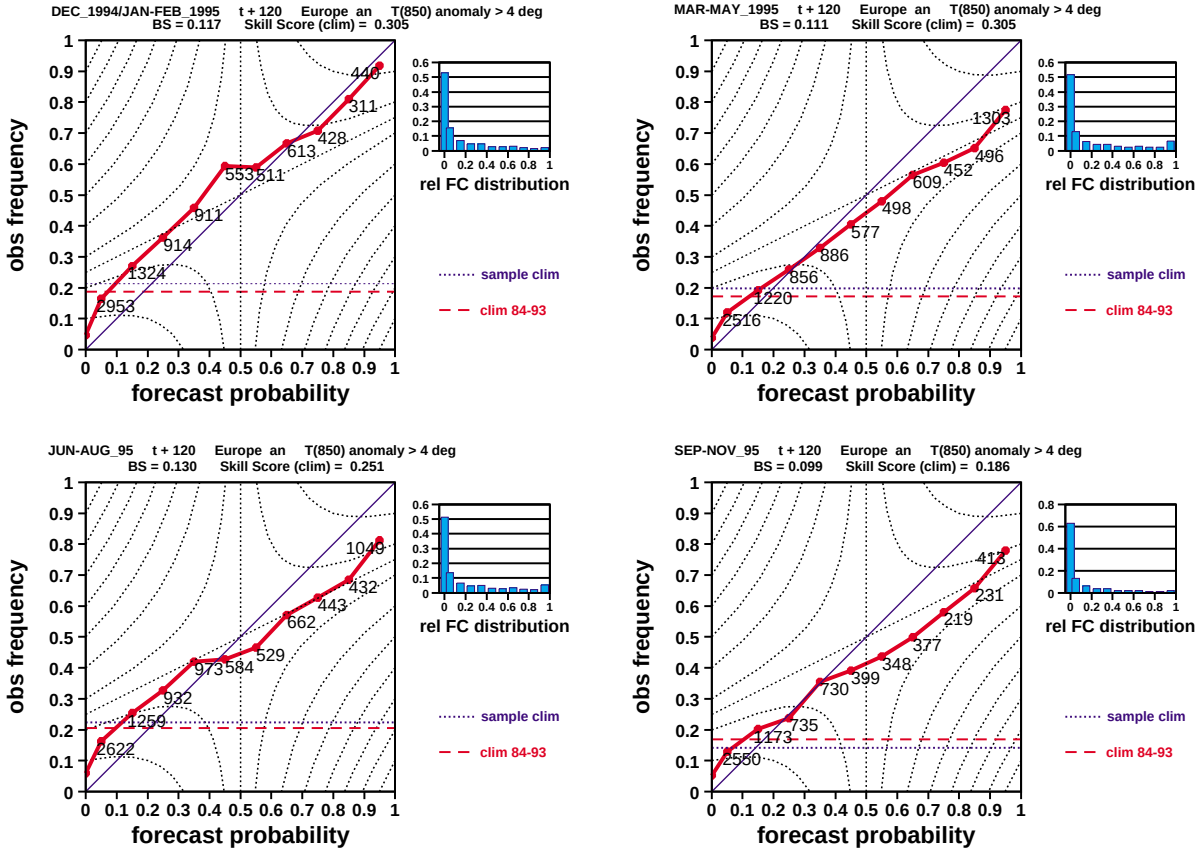

Figure 10. Reliability diagrams for 850 hPa temperature warm anomalies of more than 4 degrees for t+120 hours over Europe. Top left: winter 1994-95; top right: spring 1995; bottom left: summer 1995; bottom right: autumn 1995.

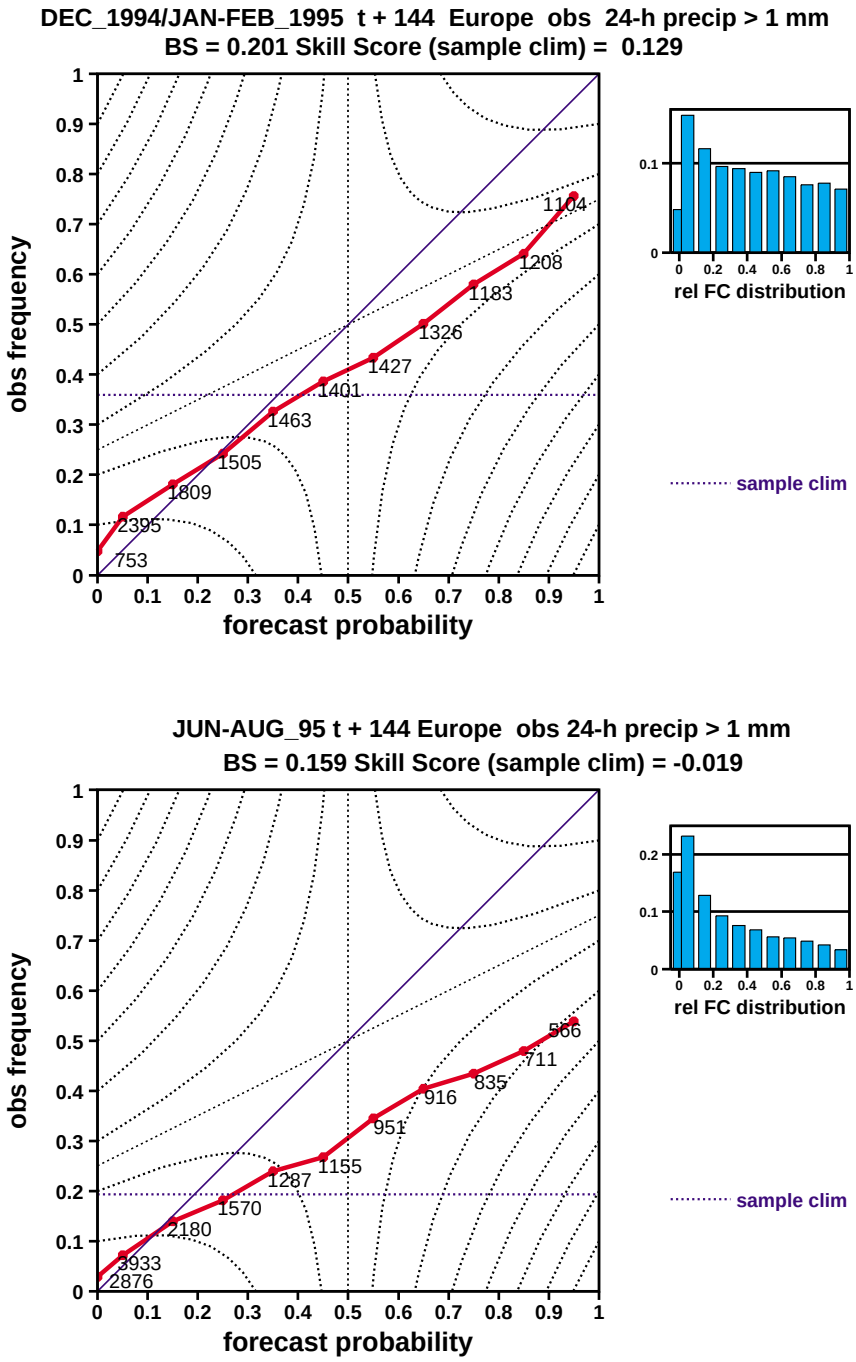

Figure 11. Reliability diagrams for precipitation accumulated between t+120 and t+144 hours greater than 1 mm. Verification against observations from about 200 SYNOP stations in Europe. Top: winter 1994-95; bottom: summer 1995.

Data Handling via MARS at ECMWF

Introduction

MARS (Meteorological Archive and Retrieval System) handles the Centre's data archive. The current MARS archive has been growing both in size and variety. It is now reaching its limits and needs to be redesigned in order to cope with the increase in generated data foreseen after the arrival of a new Fujitsu system and the never-ending requirement to archive new kinds of data types.

Growth

In 1985, the operational archive was storing 70 megabytes per day. By 1995, the total archive had reached 6 terabytes, with a daily growth rate of 8 gigabytes of operational data plus 9 gigabytes of research data.

The number of items archived per day is now very large: 120 000 operational fields, 140 000 reports and 200 000 research fields.

The meteorological data is not yet all under the control of MARS. New data types are first checked by the Research Department. Then, when a new data type is brought into use in the operational suites, it is added to the archive. For this reason, the archive has grown in variety and complexity. (See Figure 1 which shows when each new data type was added to the archive).

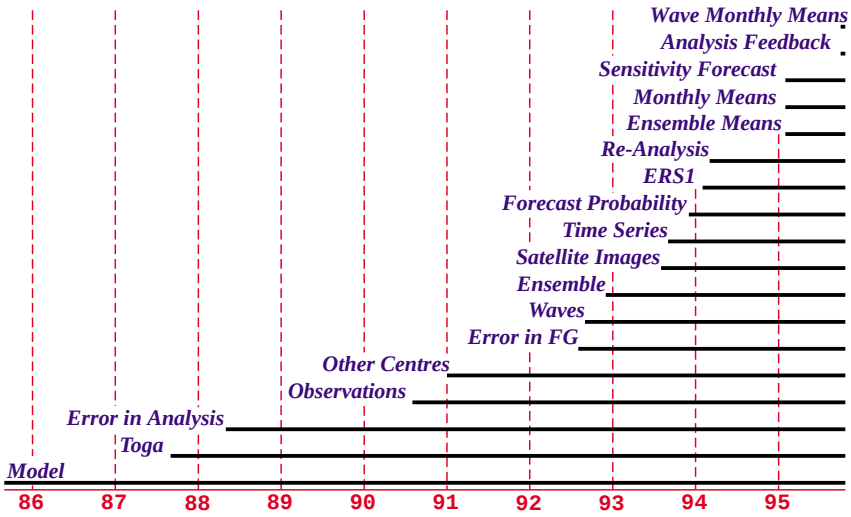

Figure 1. Introduction of new data types in MARS.

In the near future, data from 3D variational analyses, seasonal forecasting, ocean models and ensemble forecasts from other centres will also be archived. Further in the future, experimental analysis feedback and private datasets will have to be handled by MARS.

In order to accommodate these future requirements, the internal architecture of the software has been changed, while its external user interface has remained untouched.

A bit of history

The Cyber

MARS is based on a file store from Los Alamos named Common File System (CFS) which runs on an IBM MVS mainframe. The first client was a CDC Cyber. At that time, the client-server concept did not exist. Clients were called worker machines. The protocol was file based and networking was done with RHF.

There was no direct communication between MARS and CFS. Requests would be passed in files to the MARS server which would translate them into a list of commands in the CFS native language mass. This list would be passed back to the worker machine which would then issue the commands and retrieve the data. (See Figure 2).

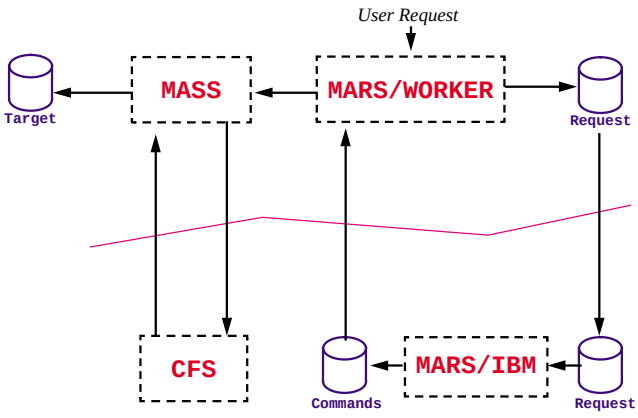

Figure 2. MARS Architecture - Cyber

The Cray

Later, the MARS client was ported to a Cray running COS. The MARS client could access an on-line field database (FDB) that contained model output. Communication with the IBM was done via SUPERLINK, and ecfile was used as a file transport mechanism. An interface between MARS and CFS was written. The MARS server would receive a request and translate it into a list of CFS files, offsets and lengths with the help of a basic database using a VSAM file. This information would be passed to CFS which would extract only the relevant fields to pass back to MARS. These results would be saved again in CFS to be fetched by the client, as shown in Figure 3.

Post-processing facilities were added: sub-area extraction, interpolation and derived fields.

UNIX

With the coming of UNIX, new tools were available: application level networking with TCP/IP, better control of operating system resources through the C language, and text parsing with lex and yacc.

So a new MARS client was developed. The previous system could access two sources of data: the FDB and CFS. The idea was to implement an architecture that could access any number of data sources or "databases".

The "database" concept

Data access should be done in an object oriented fashion, dealing with different data formats in a transparent way (data hiding) and allowing the same manipulations to be carried out on data with different contents (polymorphism).

A database in the MARS context is a software package that supports four calls - open, read, write and close - and understands a MARS request (Figure 4). A client program scans all available databases and calls them one at a time. A MARS request is passed in an open call., and subsequent reads return one field at a time. Writes would be called for archiving data.

Using this technique, any new data source can be added to the MARS system by simply implementing one of these database packages.

A database can be installed spread across a network with the help of a "network glue" software that uses TCP/IP to send function parameters across the network and receive return values. Because the network calls are hidden in the "glue", neither the database nor its client has to be changed to use this feature. None of the components involved has to be aware of the presence of the network. (See Figure 5).

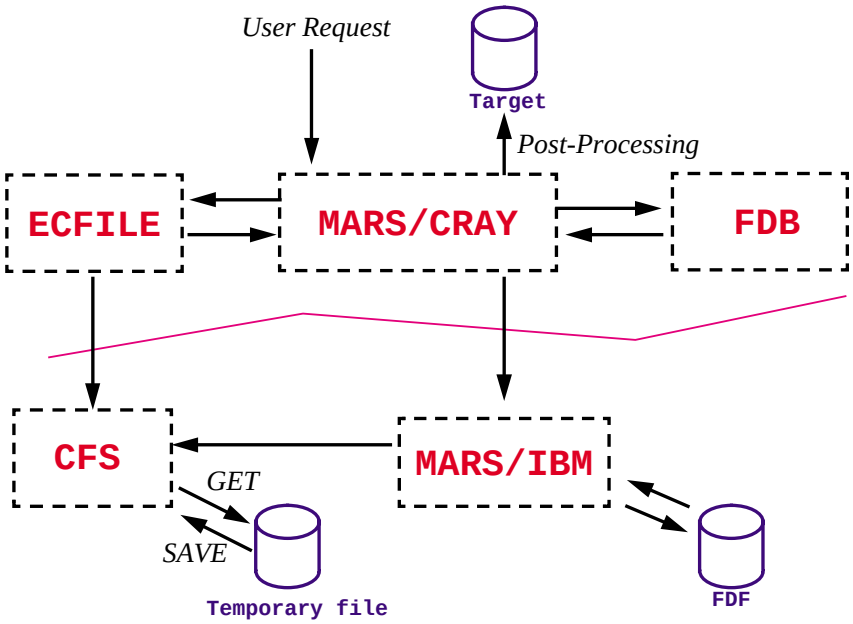

Figure 3. MARS Architecture - Cray

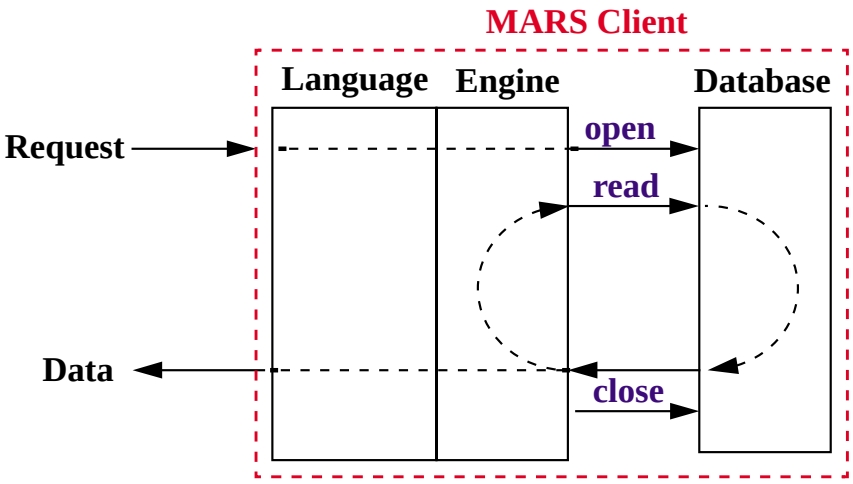

Figure 4. The "Database" Interface.

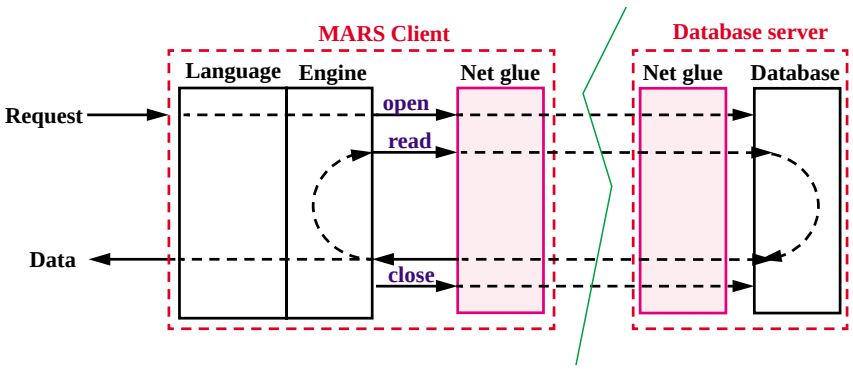

Figure 5. Networked databases.

Caching

Several of these database packages have already been implemented: one to access the IBM archive, one to access the Cray FDB, and one using the commercial SQL database "Empress". The latter is used to implement the MARS caching system (MCS) for the fast access to archived data needed for interactive work using Metview. Data retrieved from one database can be written to another. In this way, a field that is being accessed often is moved "closer" to the user. (See Figure 6).

To show the effect of caching in improving retrieval times, the following request was run twice with an interval of one minute between requests. The first run went to the IBM archive and took some time:

```
MARS - INFO   - Request 1 is :
RETRIEVE,
              TYPE     = FC,
              LEVTYPE  = SFC,
              PARAM    = T,
              DATE     = 940601/TO/940605,
              TIME     = 1200,
              STEP     = 12/TO/240/BY/12,
              TARGET   = "foobar",
              GRID     = 1.5/1.5
MARS - INFO   - Calling IBM
MARS - INFO   - Got 5.55 Mbytes from IBM
MARS - INFO   - 100 fields retrieved from 'IBM'
MARS - INFO   - Request time: 44 min 21 sec
```

The second run found the data still in the MCS cache and was much faster:

```
MARS - INFO   - Request 1 is :
RETRIEVE,
              TYPE     = FC,
              LEVTYPE  = SFC,
              PARAM    = T,
              DATE     = 940601/TO/940605,
              TIME     = 1200,
              STEP     = 12/TO/240/BY/12,
              TARGET   = "foobar",
              GRID     = 1.5/1.5
MARS - INFO   - 100 fields retrieved from 'MCS
Research 2'
MARS - INFO   - Request time:  wall: 10 sec
MARS - INFO   - No errors reported
```

Note the difference in request times between the two runs.

Member States access to MARS

The same scheme will be used for MARS access by Member States. A member state user will run a client MARS on his or her machine which will possibly access some local database, then a site-wide MCS, and then an ECMWF MARS proxy (implemented as a database as far as the member state MARS client is concerned). The MCS database would cache any field retrieved from ECMWF, to avoid unnecessary repeat

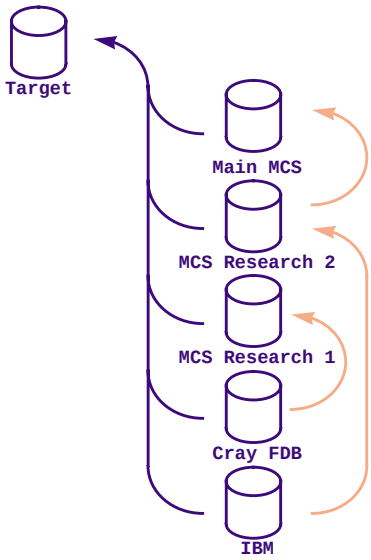

Figure 6. Data Access and Caching.

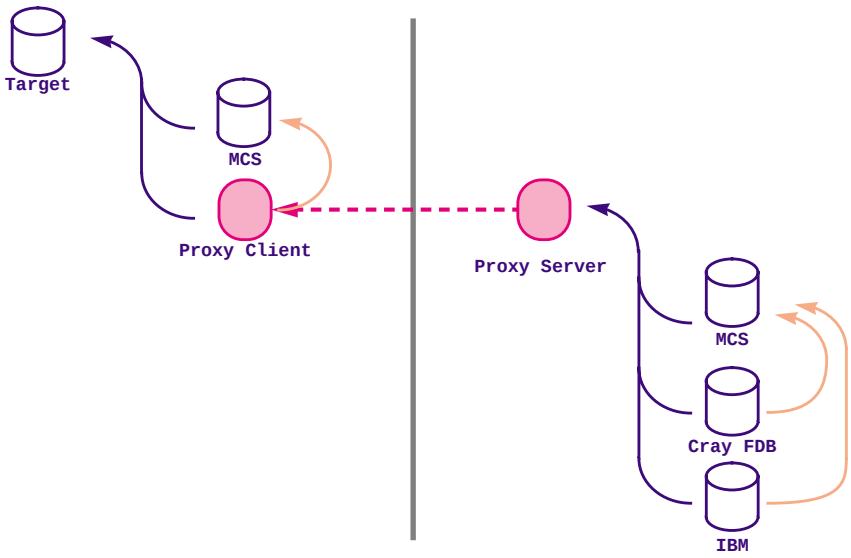

Figure 7. Member States Access to MARS

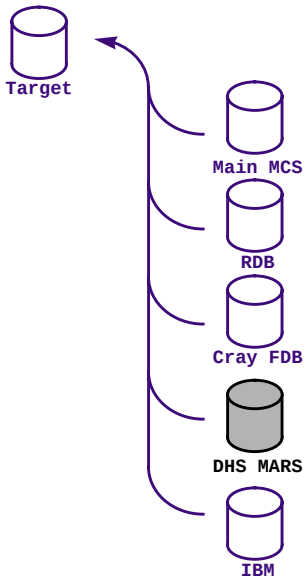

Figure 8. Accessing Both Archives

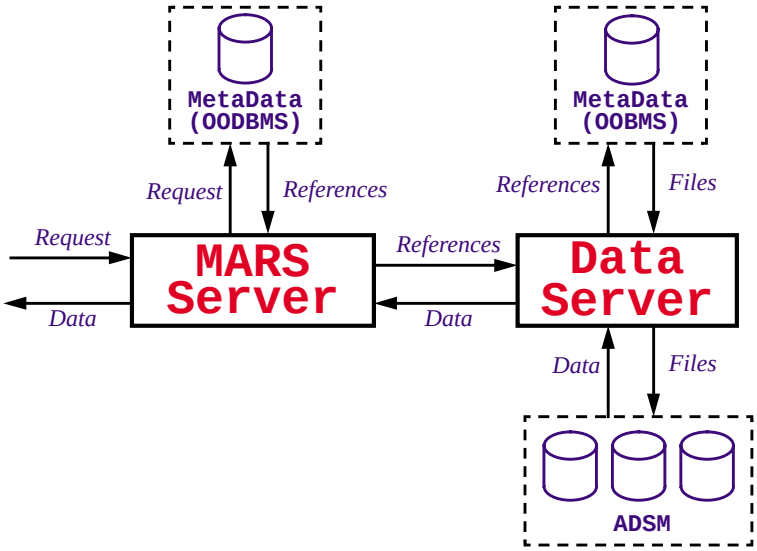

Figure 9. MARS Architecture - DHS

network traffic should a number of users from the same site request the same field. (See Figure 7).

This access scheme is already under test and will soon be made available to Member States.

The DHS project

The present design of MARS has reached its limits. The maintenance of the MARS server on MVS is becoming increasingly difficult. The Cray will be replaced by a more powerful Fujitsu that will produce greater volumes of data at a faster rate. New data types will be added. So MARS needs a new design using newer technologies.

Hardware

The MVS mainframe will be replaced by three RS/6000 R30s from IBM, each with four PowerPC 604 CPUs. Each machine will have 1 gigabyte of memory, 1 megabyte of cache per processor and 3 gigabytes of disk, and will share 240 gigabytes of disk with the others.

A tape robot with 8 drives will be able to mount two hundred 10-gigabyte tapes per hour.

Software

CFS will be replaced by the ADSM data management system and an IBM file store. The MARS server will be written in C++, and its metadata describing the archive contents will be stored in a objected oriented database (OODBMS).

To extend access to both new and old systems, a new “database” will simply be added to the list of existing databases, as shown in Figure 8.

Figure 9 shows the first analysis of the future system. A Data Server will sit on top of ADSM and handle files. Each file archived will be assigned a unique identifier

called a reference. The MARS Server will translate user requests into references and retrieve its data from the Data Server. Because of the clear separation of roles for the two servers, the Data Server will be free to organise files in the most efficient way according to its underlying software and hardware, while MARS, retaining 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.

Metadata

The success of the future MARS lies in how effective it is in constructing its metadata, i.e. the "data representing the data". The main problem in designing metadata is that it is impossible to foresee the ideas coming from research. MARS must cope with data types which have not yet been "invented", as well as a variety of very different "objects" such as fields, observations, images or cross-sections.

Using object-oriented technology will help. With encapsulation, MARS will be able to handle new data types, each data type being "self aware" of its contents and the methods to handle them. Inheritance will make it easier to support new attributes and make it possible to reuse code for different types (e.g. an ensemble forecast is a forecast plus an ensemble number). Using polymorphism, the MARS engine will be kept simple as it will have a single view of all the data types.

Conclusion

The present design has been successful for more than 10 years, and nowadays supports one of the largest and most complete meteorological archives. We now face the challenge of redesigning MARS for the next decade. New technologies will be used to enhance the present system, to keep it more open and flexible to cope with the growth in quantity and variety of the data.

Baudouin Raoult

High-Availability Server

Why high availability ?

Over the past few years the amount of distributed computing has increased atECMWF due to the use of UNIX workstations connected over a local area network to a number of servers and other machines.

As these workstations have a limited disk capacity it is necessary to install large software packages on servers which are then used by the workstations as clients. This places a dependancy between the client and server such that the client always requires the server to function.

If a client workstation fails, only one user is affected, however, if a server fails, a much larger number of users can be affected. Therefore, the server is a potential single point of failure that can affect many people.

High-Availability (HA) is the generic name given to UNIX systems that try to overcome this "single point of failure" limitation and thus improve the overall reliability of the services offered.

What is High-Availability ?

HA is a combination of hardware and software that protects against single faults or error conditions which would normally interrupt a service. This could be failures such as:

- operating system failures (e.g. the machine crashes or hangs)
- network connection failures
- disk drive failures
- disk driver interface (SCSI bus) failures

What HA system do we use ?

As a result of a tender process in 1995 an HA cluster from Hewlett Packard was selected and subsequently installed in December 1995. The cluster consists of three "K" series servers, each with 1 CPU and 512 MBytes of memory. Each server has an FDDI interface as its primary connection to the internal network.

The main disk storage on the cluster is provided by 3 RAID arrays. The term RAID is an acronym for "Redundant Array of Independent Disks" and was originally conceived as a method of providing a large amount of disk space using smaller, cheaper drives working together as "virtual disks".

Depending on the way that these disks are organised within the array it is also possible to make these "virtual disks" more reliable by using some of the capacity for parity information. The different configurations are referred to as RAID levels.

How is HA achieved ?

Hewlett Packard have developed a software product called ServiceGuard, which makes it possible to define a cluster of machines which provide a number of services.

Each service is defined as a network address, filesystems and software (e.g. daemons).

When a server is run as part of a cluster then ServiceGuard controls whether that machine acts as a node, and if so whether it provides any of the cluster services. If there is a subsequent failure then ServiceGuard determines which machine will take over the services from the failed node.

In addition, there are a number of system components which are duplicated/configured to protect against a single point of hardware failure, such as:

- the root (boot) disk is mirrored onto another disk drive;
- there are two private Ethernets connected to all three machines for exchanging "heart-beat" messages;
- each FDDI network interface is dual-attached (using two cables);

- the external RAID disks are connected through two SCSI buses to each machine;
- the RAID units are configured with RAID level 5.

RAID level 5 provides the most reliable storage because it uses some of the disk space for parity information which can be used to recreate data if there is a disk fault. In level 5 this parity information is distributed evenly over the disks in order to improve performance. In addition, there is a dedicated "hot spare" disk in each RAID unit, and the recovery and protection features of the RAID units are monitored and controlled autonomously by the RAID units themselves. The host sees a "virtual disk" that is always available and is error-free.

What happens if one of the machines crash ?

If one of the servers crashes then the following sequence of events would happen:

1) the surviving hosts realise that one of the nodes has crashed after approximately 10 seconds as the "heartbeat" messages have stopped;
2) the surviving hosts reconfigure themselves as a two node cluster;
3) the ServiceGuard software determines if any service that should be provided is not running and could be provided by any of the survivors;
4) if one of the survivors can provide a backup to the machine that crashed then it runs the service start-up script, which:
   a) puts on-line any disk volumes required;
   b) checks and mounts any filesystems required;
   c) starts any daemons/processes required;
   d) modifies the network interface to add any IP addresses required.

All this means that the user should see a downtime of less than five minutes, while the service is transferred from one node to another, rather than having to possibly wait hours while a server is repaired, or replaced manually.

What services do we protect with HA ?

All workstation users at ECMWF have their home directories served via NFS from the HA cluster.

The pre-processing part of the operational suite will shortly be run on the cluster.

Other services such as the central E-mail server, and network registration services will also be moved on to the HA cluster in the near future.

Summary

The purpose of the High Availability cluster is to improve the overall reliability of the services offered to workstation users. It does this through a combination of hardware redundancy and software duplication, the latter being controlled by Hewlett Packard's ServiceGuard software. Since the equipment has been in operational use since 1 April 1995 it has provided excellent service.

Stuart Mitchell

Migration to 36 track cartridges completed

In March 1994 migration of data stored on 3480 cartridges began, with the aim of copying all the Centre's data from 18 track to 36 track media. This would double the capacity of the media, thus effectively extending the storage capability of the main tape store.

In addition to doubling the capacity of each cartridge a variable compression factor was introduced, typically gaining 20-30% capacity per cartridge. Further storage capability was also gained by replacing 540' media with 650', and more recently with 1100' media.

Thus media storage capability has changed from 210 Mbytes per cartridge to ~800 Mb - reaching 1 Gbyte in extreme cases - over the period.

Peter Gray's article on 'StorageTek Upgrade to 36 Track' (ECMWF Newsletter No. 67) stated that 66,000 cartridges would be copied by late October 1994. The initial batch of 54,000 cartridges were migrated successfully over a six month period on an individual volume by volume basis, the remaining 12,000 containing Mars research experiment data to be copied once Research Department had purged redundant data.

When the final migration suite started in November 1994, it quickly became apparent that contention between this migration and production Research experiments was causing data to be interleaved. It was estimated that the migration programme would thus take three to four years to complete, totally impractical and unacceptable, especially as the process would need to be repeated once migration to the new data handling system began. Therefore, software development started in March 1995 to allow data to be migrated file by file rather than by volume.

By that time, 15,000 cartridges with 1.8 million files had to be copied. This involved running four migrate streams at any one time, using up to 8 silo drives. The average number of cartridge mounts rose from -3000 per day to in excess of 5000 per day.

A final migration started on 12 June 1995, with the last cartridge being copied on 11 January 1996, although the bulk was completed by mid-December 1995.

Since the exercise began in 1994 the number of active cartridges stored in our library has increased from 66,000 to approximately 73,000. Without the storage enhancements, plus the migration exercise, the number of cartridges would be well in excess of 250,000. The floor area reasonably holds around 100,000 cartridges!

Within the next year or so, copying data to the new data handling system's 10 Gbyte cartridges will begin.

Meantime, migration continues to recover "holes" caused by deleted data This, hopefully, will extend the storage capability of the main tape store just long enough to see full production on the new data handling system.

Graeme Walker

Efficient use of MARS/ecfile

Introduction

Ecfile is the interface on various user systems (Cray, Fujitsu, workstations) that allows you to store/retrieve your own files on the Data Handling system. MARS is another interface which allows you to retrieve operational or experimental observations/analysis/forecast data that ECMWF has accumulated over the years. CFS is the system running on the Data Handling system which actually stores/retrieves the data for ecfile and MARS. CFS also controls use of the physical storage media (disk, cartridge tapes). Thus CFS is critically important.

Recently, we have seen a number of occasions when the efficiency of CFS has been adversely affected by MARS users. Similar situations can also be caused by ecfile users.

In a sample case, a MARS user was asking for 10 year's of forecast data within a small sub-area. This involved retrieving a large amount of data and hence used a lot of system resources to satisfy. Because many of the tapes were not in the tape silos, it resulted in a large number of requests for tapes to be fetched and mounted. This delayed other users access to data and led to more idle time on the Compute server.

This article attempts to explain what happens in practice and how you can help the system and yourself to get the best service.

Description of current hardware

CFS manages essentially two main types of storage - disk and cartridge tape. The disk-based storage is a very small proportion of the available space (about 0.5% of the total) and, in relative terms, is growing smaller as time goes on. Hence this disk space has to be efficiently used if file transfers are to be as fast as possible. This is because the disk space is a "cache" store and if it fills up too quickly the speed of access to files suffers. Then everyone is affected by poor performance.

For the tape based storage we have four StorageTek robot silos capable of holding a total of 22,000 tapes. Together they hold a maximum of 17 Terabytes which, at the time of writing, is just under half the total data managed by CFS (40 Terabytes). This means not all of the data can be held in the silos simultaneously.

Access to tapes is serial. Most of the time data is accessed by the robot fetching the relevant tape to a free read station. It can take up to two minutes to access the data on a tape assuming there is a free read device in that silo. If not, the tape has to be transported from silo to silo via the so-called pass-through port to find a free station, making the delay even longer. If the tape is outside the silo, it will have to be manually fetched and entered into a silo. This can take many minutes.

Similar considerations apply for writing. CFS will add data to any mounted tape where there is sufficient free space. However, if there is insufficient space, then CFS has to mount another tape.

Each silo has a door for entering/removing tapes with space in the door for some 30 tapes. Thus, inactive tapes can be taken out and stored off-line, and requested or new tapes (for writing to) can be entered. When tapes are entered it implies an equal number of tapes must be ejected because the silos are fully populated to avoid waste.

How data is stored

When a file is initially stored in CFS it is put on to disk, unless it is very large or the user explicitly requests (via the use parameter) that the file goes to tape. Eventually the CFS system will migrate the file off disk and write it to the next available tape. This occurs when there is a space demand on the disk cache, or the usage made of the file declines. Being a demand-based system this is rather random. For example, files written consecutively within a job are unlikely to be written consecutively onto tape(s) even if you request the file to be written to tape immediately. This is because another user's file could be in the "queue" to be written next. This leads to a fragmented system behind the scenes.

To avoid the worst excesses of fragmentation, a so-called Migrate job works in the background copying files from one tape to another. For example a delete operation is not done directly on a tape file. Instead, the directory entry for that file is logically deleted. The migrate job on its next run then makes a new tape with that file missing from the new tape. If this was not done, the number of tapes would increase even more, resulting in more tapes being outside the silos. However, currently no attempt is normally made to copy all of a user's files to consecutive positions on one or more tapes.

MARS has a special concept known as families to ensure that operational and standard research experimental data is written to set of tapes in a sequential and organised way. A copy is also made to disk to allow efficient access to recent data, that copy just being deleted later. In this way the number of tapes is minimised and later access is more efficient. However, all the data for one day's forecasts still occupies a number of tapes, there being separate tapes for analysis, model level, pressure level and surface data.

User requests for data

A MARS job that is composed of requests covering a long time span actually requires a large number of tapes to be mounted and many files to be transferred. Hence it will take a very long time to complete.

Be aware that the MARS archive was designed as a daily (case-based) archive and, as yet, does not have a complete time-series archive for forecast data. However, the Re-Analysis Project will provide one for analyses. Therefore, if you request months or even years of forecast data in one job you could easily create a job that would run for DAYS because of the volume of

data to be retrieved. For this reason it is important to consider, when you submit jobs containing multiple MARS requests, how many (MARS) "files" you are requesting. There is a sizeable overhead for each file not in the disk cache area. It is even worse if each file is outside the tape silos.

The main point is that unless we know in advance of large data requests, the tapes will be requested one by one and will have to be mounted one at a time as they are requested. This is very inefficient. Hence this is where you, the user, can help us to provide a better service to you and everyone else.

We will be particularly grateful if when you request, say, more than a year's worth of data you let us know in advance what specifically you are requesting, in MARS terms. Then we can mount the tapes in advance and avoid unnecessary work for those who have to load the tapes into the silos. It also avoids potential conflicts within the silos.

One other way users can help is to avoid re-requesting the same set of tapes more than once. This is not always obvious but as a general rule avoid running a job to get back all of one parameter over a period and then running another job over the same period to get additional parameters using data of the same type. Here, model level, pressure level and surface data are all regarded as being of different "types".

It is hoped that we will soon be able to offer better guidance on what data is likely to be available in the silos and what is not. For the moment you can assume that data created recently is more likely to be in the silos than older data.

Finally, if you are going to store a large amount of data in ecfile over a period we may be able to assist if we know what you are doing. In a recent case a "family" was setup for a user so that his files were "cached" before being written to tape. This led to a marked reduction in the number of tape mounts and increased the efficiency of access for everyone.

To sum up then, if you plan to read or write large amounts of data over a period please discuss your requirements, in advance with User Support. Do not assume that ECMWF can automatically "cope". The pressures that will be placed on the current ECFILE/MARS/CFS system by more compute power being available when the Fujitsu arrives will make efficiency of access even more critical in the coming months. Also a new Data Handling system is due to be introduced soon and there will be a transition period during which there will be even more CFS activity.

Although this article has concentrated on the current ecfile service we also expect much of this to apply to the new data handling system.

John Greenaway

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 J Goas	Mr D Birman	Mr J Goas
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	Mr S Kruizinga	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 M Leitao	Mr C M Fernandes	Mrs I Barros Ferreira
Switzerland	Mr H Müller	Mrs C Raeber	Mr M Schönbächler
Finland	Dr M Alestalo	Mr T Hopeakoski	Mr P Nurmi
Sweden	Mr L Moen	Mr S Orrhagen	Mr O Åkesson
Turkey	Mr M Örmeci	Mr M Örmeci	Mr M Örmeci
United Kingdom	Dr S R Mattingley	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 1996

July 3 - 4	Council	*44th*
Sept 2 - 6	Seminar - Data assimilation	
Sept 9 - 11	Workshop - Nonlinear aspects of data assimilation	
Sept 30 - Oct 2	Scientific Advisory Committee	*25th*
Oct 7 - 8	Computing Representatives meeting	
Oct 8 - 10	Technical Advisory Committee	*23rd*
Oct 15 - 16	Finance Committee	*57th*
Nov 4 - 7	Workshop - New Insights and Approaches to Convective Parametrization	
Nov 27 - 28	Council	*45th*
Dec 2 - 6	Workshop - 7th ECMWF Workshop on the Use of Parallel Processors in Meteorology	

ECMWF Publications

Technical Memoranda

No. 223 (not yet published)

No. 224 Das, S., M. Miller, P. Viterbo: Impact of prognostic cloud scheme and subgrid scale orography on the simulation of the Asian summer monsoon. December 1995.

Technical Reports

No. 77 Lanzinger, A.: ECMWF forecasts of the floods of January 1995. December 1995

Computer Bulletins

No. B8.2/1(1) Supporting incoming/outgoing magnetic tapes at ECMWF.

No. B0.1/1(5) ECMWF Computer Division Management and Personnel list.

Miscellaneous Publications:

International TOVS Working Group. Technical Proceedings of the Eighth International TOVS Study Conference. Queenstown, New Zealand, 5-11 April 1995.

Forecast and Verification Charts:

Forecast and Verification Charts to February 1996