# ENSEMBLE FORECASTING AT THE UK METEOROLOGICAL OFFICE

D S Richardson and M S J Harrison
Meteorological Office
Bracknell, UK

## 1. INTRODUCTION

Since their introduction in December 1988, ensembles of model integrations have provided an important input to the extended-range forecasts produced at the UK Met Office. Over the last 5 years a large database of ensemble forecasts has been built up. A brief summary of some general conclusions applicable to both medium- and extended-range ensemble forecasting will be given, based on results from these ensembles.

During 1993, 33-member ensembles have been produced by ECMWF for use in the medium range out to 10 days ahead. For two of these occasions 33-member ensembles have also been produced with the Met Office Unified Model (UM) using initial conditions generated by adding the ECMWF perturbations to Met Office analyses. The performance of both models' ensembles is assessed and contrasted.

One of the major problems in the utilisation of ensembles is the presentation of the large amounts of data produced by the ensemble. Some ideas on information presentation will be presented.

## 2. RESULTS FROM MET OFFICE MONTHLY ENSEMBLES

Each ensemble has nine members generated using the lagged average technique (LAF) where forecasts are run from successive 6-hourly analyses. The integrations are made using the operational Met Office global forecast model run at reduced resolution. Results from these ensembles have demonstrated the value of ensemble forecasts in the medium and extended ranges. The simple technique of using the ensemble mean as a best-estimate deterministic forecast does give on average more skilful forecasts than the alternatives of climatology or persistence. However use of the ensemble mean is based on the assumption that the ensemble distribution is Gaussian. In fact on many occasions the ensemble produces a range of possible outcomes. A typical example is shown in figure 1: here three distinct solutions are produced by the ensemble, each giving quite a different forecast for the UK. In such a case it is inappropriate to construct the ensemble mean, which could produce a very weak pattern giving no indication of the distinct solutions present in the ensemble distribution. Maximum benefit can only be derived by using all the information in the ensemble distribution: such a prediction may best be expressed explicitly as a probability forecast.

In the operational 30-day forecasts, predictions are made of temperature and rainfall for 10 districts of the UK. Linear specification equations are used to derive the temperature and rainfall forecasts in climatologically equally-likely categories (three categories for rainfall and five for temperature) from

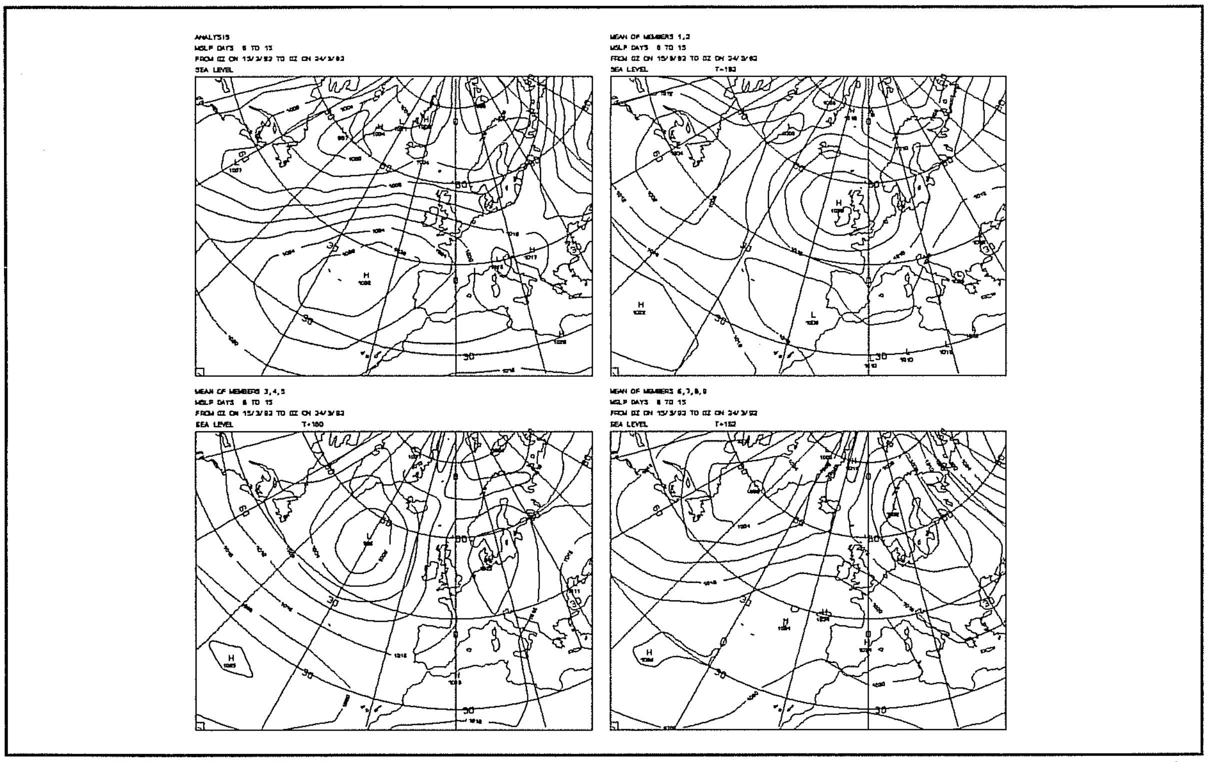

**Figure 1**. Example of range of solutions produced by an ensemble forecast. PMSL fields for forecast days 6-15 for verifying analysis (upper left) and 3 distinct clusters formed from the 9 ensemble members.

forecast sea-level pressure fields (and, for temperature only, current local SSTs) . A forecast probability distribution can be built up from the temperature categories forecast by each ensemble member. The average skill of such probability forecasts has been shown to be higher than that of the ensemble mean forecast (taken as a 100% forecast probability for the ensemble mean category). Spread of the ensemble was found to be a poor predictor of skill. In particular some low spread forecasts give a misleading indication of high confidence in cases where a regime transition occurs, but all ensemble members fail to predict this. A well known example is the difficulty many forecast models have in predicting the development of a block more than a few days ahead. This highlights the need of a good forecast model in the medium and extended range to have realistic variability, ie to be able to simulate the full range of possible atmospheric evolutions. The new Unified model has been seen to be much better in this respect than the previous Met Office model. This matter is particularly important when trying to give an accurate representation of a forecast probability distribution in a ensemble forecast.

## 3. PERTURBATION OF MODEL FORMULATION

Ensemble forecasts reflect the sensitivity of integrations to small perturbations of the initial conditions. The effect of the exact formulation of the model on the ensemble must also be considered. Even minor changes to tunable parameters in a model can have a large impact on not just an individual forecast but on an ensemble of forecasts. This can be illustrated by a case study; this case was chosen as an example of a difficult forecast of the development of a block over Europe (this is one of the cases examined for the 33-member ensembles in section 4). The operational 30-day 9-member ensemble was taken as the control ensemble. Three other ensembles were then run from the same initial conditions but for each ensemble one minor change was made to the model formulation. There is no a priori reason for choosing either the original or the modified values for the perturbed parameters, and in fact both have been used in the UM at some point.

The skill of each ensemble mean forecast is compared in figure 2. At days 1-5 all four are similar but differences develop by days 6-10 and these become marked from days 11-15 onwards. In this case study one of the ensembles retains high skill throughout the forecast period. This should not be taken as evidence that this formulation is best, but the variation between the ensembles does illustrate that minor formulation changes can have marked effects on an ensemble prediction. The variance across each ensemble also changes with model formulation. At days 6-10 the location of the main centres of variation are similar but the magnitudes differ: at later periods both location and magnitude vary to some extent. The range of solutions over all 36 integrations is greater than across any individual ensemble. Comparison of this 36-member ensemble, produced by a combination of LAF and model perturbation, with a 33-member UM ensemble initialised using the ECMWF perturbations shows that there is a similar range of solutions produced by both ensembles. This case study has demonstrated the potential impact of minor changes to model formulation on an ensemble prediction. Formulation changes may produce as great an effect on an ensemble distribution as perturbations to initial conditions. Both may need to be used (perhaps combining different models) to produce a representative distribution. Perturbing model formulation could be used as a simple complementary method of generating an ensemble.

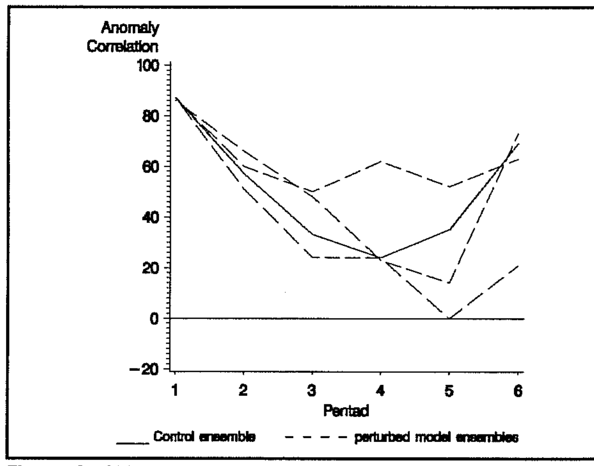


Figure 2. Skill of model formulation perturbation ensembles. Anomaly correlation of 500 hPa height fields for each ensemble mean.

## 4. 33-MEMBER ENSEMBLES

The performance of ensembles in the medium range has been considered in 6 case studies. For each case the ECMWF 33-member ensemble is available and for two of the cases the Unified Model ensemble has been run with the initial conditions generated by applying the ECMWF perturbations to a Unified Model analysis. The initial perturbations are generated by the singular vector technique (see paper by TN Palmer, this volume). The Unified Model forecasts have been run at a lower resolution (2.5° lat x 3.75° long) than the T63 ECMWF forecasts. Results from the two cases where ensembles from both models were available are presented below. Some general remarks based on results from all 6 cases are also given.

### 4.1 Case 1: 24/1/93

The performance of each ensemble as measured by anomaly correlation of 500hPa height fields is

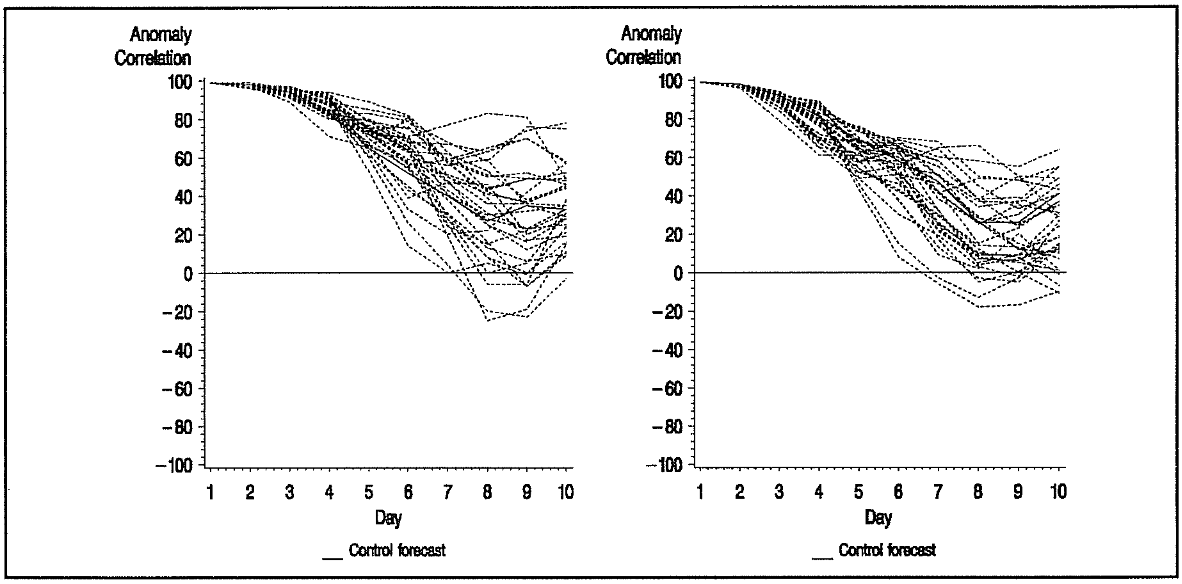


**Figure 3.** Daily anomaly correlation of 500 hPa height fields over N Atlantic and Europe for ECMWF (left) and UM(right) ensemble members. Forecast initialised on 24 Jan 1993.

shown in figure 3. In both ensembles the range of skill across ensemble members is relatively small to day 4. For later times this range increases, several members of each ensemble retaining high skill while others lose all skill. Initially the ECMWF ensemble has higher skill but beyond day 4 the same range of scores is seen for both sets of forecasts. To investigate the grouping of the forecasts in phase space cluster analysis has been applied to all 66 forecasts and the verifying analysis taken together for day 5 and day 10 (figure 4). At day 5 the ECMWF and UM forecasts cluster quite distinctly indicating that the two models are producing different sets of solutions. However by day 10 there is one main cluster which contains equal number of ECMWF and UM forecasts. This grouping into a single cluster indicates that in this case the two ensembles lie in the same region of phase space. The benefit of combining the two ensembles is in

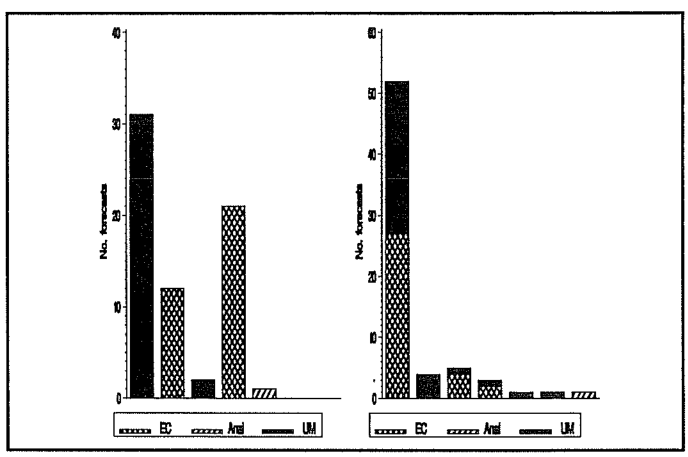

**Figure 4.** Clustering of 500hPa height fields over Europe for all 66 forecasts initialised on 24 Jan 1993 and verifying analysis for day 5 (left) and day 10 (right).

the better representation of forecast probability distribution given by the doubled ensemble size. In this case the analysis remains distinct from the forecast clusters at both times: even with large ensembles there is no guarantee that the phase space evolution of the atmosphere will be correctly captured.

### 4.2 Case 2: 7/3/93

In this case the skill of both ensembles is higher than in the previous case and both have a much smaller range of scores (figure 5). Again the ECMWF forecasts have a noticeable advantage in the

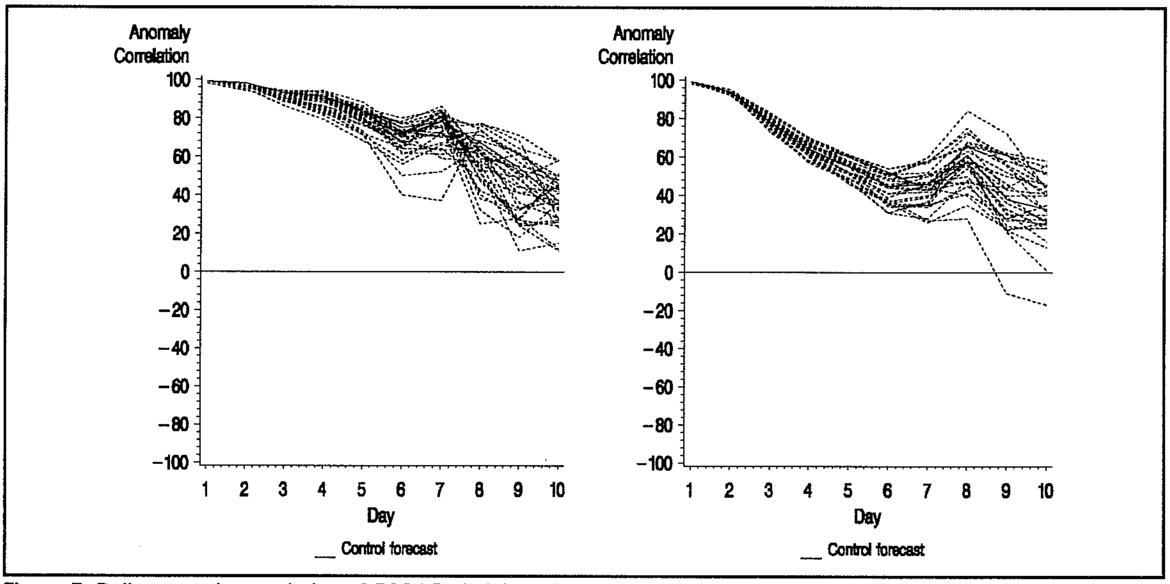


**Figure 5.** Daily anomaly correlation of 500 hPa height fields over N Atlantic and Europe for ECMWF (left) and UM(right) ensemble members. Forecast initialised 7 Feb 1993.

first few days, probably due to the higher resolution. But again by day 10 the range of results is very similar. As would be expected from the differences in skill early in the forecast, cluster analysis separates the two ensembles quite distinctly at day 5 (figure 6). But at day 10, somewhat surprisingly the two models are still clustered distinctly (figure 6). The verifying analysis is grouped with the ECMWF members at day 5 and with the UM at day 10. This time the two models are exploring different regions of phase space. Neither model on its own would show fully the complete range of possible developments but, taken together, the joint distribution does cover the analysed state (but not necessarily the full range of possible solutions). The reason for the differences between ECMWF and UM ensembles could be either differences in the ECMWF and Met Office analyses to which the perturbations were applied or differences in the two models (as seen in section 3 even minor changes in a model can produce distinct ensemble forecasts). Further work is needed to investigate this more fully.

### 4.3 General remarks

The two case studies presented above have shown a number of features common to all 6 cases. Each ensemble has at least some members which are skilful (anomaly correlation greater than 0.5) out to day 10. The skill of any individual forecast can vary greatly from day to day and a particular member may switch from being one of the least skilful on one day to being among the most skilful on the following day. There is no simple relationship between the skill of the control forecast and that of the ensemble members - the control can be anywhere in the ensemble distribution. There is a large variation from case to case in the spread and skill of the ensemble. There are cases where all members are highly skilful out to 10 days, however in the majority of cases a wide range of skill is seen beyond day 5.

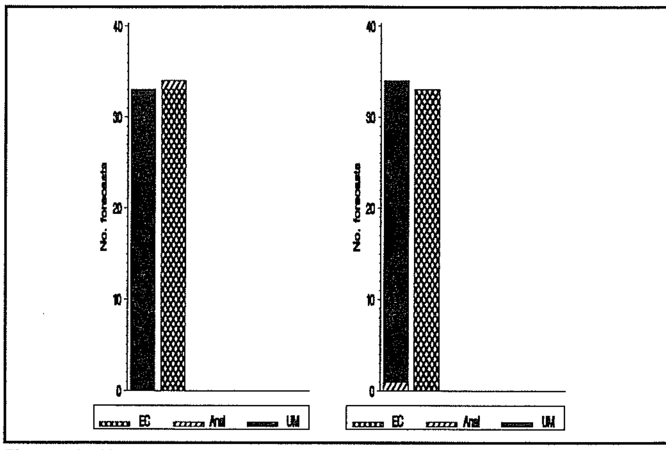


**Figure 6** Clustering of 500hPa height fields over Europe for all 66 forecasts initialised on 7 Feb 1993 and verifying analysis for day 5 (left) and day 10 (right).

### 4.4 Temperature and rainfall

Forecasts of temperature and rainfall for days 1-5 and 6-10 have been derived for all members of the 6 ECMWF ensemble. This enables probability distributions to be built up of forecast categories of temperature and rainfall. The skill of these probability forecasts has been assessed using the ranked probability score, standardised against climatology. Predictions of temperature were found to be generally skilful; rainfall forecasts were particularly skilful for the first 5 days but less so in the later period. Nevertheless the conclusion was that useful information can be gained from the probability distributions forecast by the ensemble.

## 5. ENSEMBLE DISPLAYS

Results presented in previous sections have demonstrated that the full potential of ensemble forecasts can only be realised by explicitly using the probability information from the ensemble distribution. The presentation of information from the ensemble is an important and difficult issue. The amount of information already available is vast and if (as seems likely) even larger ensemble are needed the information available from an ensemble will increase even further. The ensemble mean is not an appropriate field, yet it is not feasible for a forecaster to look carefully at all ensemble members individually. It is clear that some measures to summarise the data must be used; these may present the ensemble information as probability distributions. Some examples of possible displays can be found elsewhere in this volume. However much work remains to be done in the area of presentation of information to forecasters and end users, each of whom are likely to need the information presented in different ways.

As a first step towards this a table of possible probability displays is shown in figure 7. It is not exhaustive but indicates a range of methods which present aspects of the ensemble forecast distribution in readily assimilable formats. The displays in the table have been stratified according to time and space: information for a single time (or period mean) or for a range of times; and for either a single point (or a region) or for a number of locations. Each of the displays listed would show information from a number of forecasts rather than an individual forecast.

| | **Single point** | **Multiple points** |
|---|---|---|
| **Single time** | histograms<br>- anomalies or values<br>cumulative prob distribution<br>prob of > or < x<br>moments<br>pie charts<br>interactive point selection | spatial histograms<br>spatial cpd's<br>regime membership<br>maps of<br>- most probable<br>- extremes<br>-prob > or < x, etc |
| **Multiple times** | histogram sequence<br>plumes<br>extremes<br>time averages<br>abrupt changes<br>overlapping ensembles<br>period totals | maps of extremes<br>plumes<br>total rainfall<br>event timing<br>storm tracks<br>spread<br>ranges |

**Figure 7.** Possible displays of ensemble information stratified according to temporal and spatial range.

## 6. CONCLUSIONS

Results from five years of ensemble forecasting at the UK Met Office have shown that ensemble predictions are valuable forecasting tools. However, simple summary measures such as ensemble spread and ensemble mean fields can be misleading. To achieve maximum benefit from an ensemble it may be necessary to use the ensemble probability distribution more directly. Evidence is that skilful probability forecasts can be made from ensembles in the medium and extended range. Results from a case study show that perturbation of model formulation can have a large effect on an ensemble forecast. Model perturbation may provide a simple way of increasing ensemble size and may also be necessary to produce an ensemble which can simulate a wide range of possible atmospheric developments.

Evidence from a limited set of 10 day 33-member ensembles shows the evolution of the ensemble distribution can be complex. The two case studies of UM and ECMWF 33-member ensembles indicates that combination of the two ensembles may produce a more complete picture than is available from either ensemble separately. The amount of information produced by large ensembles presents a challenge to develop methods of presentation of this information to forecasters and end users.