AN EPS FOR THE SHORT AND EARLY MEDIUM RANGE

H Hersbach, R Mureau, J D Opsteegh
Koninklijk Nederlands Meteorologisch Instituut
De Bilt, Netherlands

J Barkmeijer
European Centre for Medium-Range Weather Forecasts
Reading, England

1. INTRODUCTION

The ultimate goal of an Ensemble Prediction System (EPS) is, given the initial probability density function (PDF) of the atmospheric circulation, to make a correct estimate of its temporal evolution. The initial perturbations of the ensemble members with respect to the analysis are to be chosen in such a way, that they grasp the relevant part of the PDF during the integration period of interest. Expectation values of meteorological quantities can then be estimated by an average of that quantity over the ensemble members.

At ECMWF, initial perturbations are based on singular vectors, which represent those error patterns that give maximum linear growth for a specific forecast period. In the linear regime, a sufficient number of evolved singular vectors spans the most relevant subspace of the PDF. The dimensionality and orientation of this subspace, highly depends on the optimalization of the parameter one is interested in. Optimalization of large scale structures, such as the geopotential at 500 hPa, will give rise to a different singular vector subspace than, for instance, precipitation. Also the horizontal and vertical domain in physical space is expected to result into different subspaces. If such a domain is a subset of a larger domain, it is expected that both relevant subspaces will have a large overlap, and that it is mainly the required dimension that will differ.

In the linear range, once the evolution of an independent set of initial perturbations is known, the evolution of the entire subset they span, is known. Beyond the linear range, this is not true anymore. The operational EPS at ECMWF is based on a subspace spanned by 25 singular vectors. In each direction two opposite initial perturbations are chosen and integrated 10 days forward in time. The area of interest are those large scale structures that give rise to a maximal error growth in the first 48 hours over the Northern Hemisphere. This is characterized by a total energy (TE) norm. This norm is also chosen for the initial time, and therefore, the flow-dependent and spatial structure of the initial error covariance is neglected. However, in the near future, (see *Barkmeijer* in this volume) this deficiency will be resolved, when the initial norm will be chosen such that it does represent the likelyhood of the initial error field.

In the medium range, the ECMWF EPS gives nice results (*Strauss and Lanzinger, 1996*), as is illustrated by reliability diagrams for large scale parameters, such as Z500, T850, and Z1000 and by Brier and ranked probability scores (RPS). However, there are indications that EPS is not optimal

for the short and early medium range. As was shown by *Talagrand et al. (1997)*, a poor-man EPS can be created that up to two days gives rise to better scores than the ECMWF EPS. An other undesired property is, that the fraction of ensemble members better (lower RMS and higher ACC) than the control is the first three days, very small.

In order to improve the behaviour of the ECMWF EPS up to the early medium range, KNMI and ECMWF have started a joint project. This project, which was initiated in January 1997, aims at a targeted ensemble prediction system (TEPS) which is concentrated on the European area only and which is optimized for a forecast period of 72 hours. Besides the usual fields, like Z500, T850, and Z1000, emphasis is being given to synoptic parameters like surface temperature and rainfall.

The use of such a targeted EPS has two advantages. First of all, by using a target area, only those singular vectors that are relevant for Europe, are selected. Therefore, it is expected that with a relatively small number of singular vectors, a larger part of the PDF can be covered.

Secondly, error growth is expected to be more or a less linear for the first few days. This means that one could fill the gaps between the members of the subspace they span by making linear combinations of the evolved perturbations. With this technique KNMI has a great deal experience. By linearly combining evolved singular vectors, a large ensemble at low cost can be created (*Barkmeijer et al., 1996*). The disadvantage of the KNMI approach is that only statistical statements can be made for large scale dynamics parameters, because the singular vectors are a result of a T42 Quasi Geostrophic tangent linear and adjoint model. The TEPS members are integrated with the full physics TL159L31 ECMWF model, and therefore contain much more information than the evolved singular vectors.

In this paper we will describe how the linear approximation can be extended for the TEPS members. Other questions, like why only a few ensemble members are better than the control up to day three, for instance, are not addressed in this paper.

The experimental setup of the TEPS system, as it is presently run at ECMWF, is described in section 2. In section 3 the nonlinear behavior of the ensemble is investigated, and an extension to the linear approach is proposed. Brier Skill Scores between EPS and TEPS are compared in section 4. In Section 5 conclusions are formulated.

2. SETUP

In order to be able to isolate the impact of using targeted singular vectors, the setup of TEPS was chosen to be as close as possible to that of the ECMWF EPS. The same tangent and linear model (T42 L31 PE) was used for the singular vector decomposition, the same rotation algorithm to create 50 initial perturbations was applied, and these were nonlinearly integrated up to day 5 with the same full physics model (TL159 L31).

The main difference is the targeting of the singular vectors and the optimization interval. These differences are displayed in Table 1.

For the targeted EPS only perturbations are selected which are relevant for a specific area at a specific optimization interval. In order to ensure that perturbations will not propagate out of this region too

	ECMWF EPS	TEPS
Optimization Interval	48h	72h
Initial Norm	Total Energy	Total Energy
Final Norm	Total Energy	Total Energy
Area	NH	Europe
Latitude	$30^oN - 90^oN$	$35^oN - 75^oN$
Longitude	$0^oE - 360^oE$	$40^oW - 30^oE$
Levels	1-31	9-31

Table 1: Comparison between the ECMWF and targeted (TEPS) singular vector computation. Both were calculated with the T42 L31 PE tangent linear and adjoint model of ECMWF.

	ECMWF EPS	TEPS test config.	TEPS present config.
Forecast Length	10 days	3 days	5 days
Type Perturbations	Rotated +/- SVs	Not Rotated +/- SVs	Rotated +/- SVs
Ensemble Size	2⋆25	2⋆10	2⋆25

Table 2: Specification of the ECMWF ensemble and the targeted ensemble. Both were integrated with the TL159 L31 full physics ECMWF model.

soon, it is important to choose an optimization interval that is well within the temporal interval of interest and a target area that is somewhat wider than the area of interest (see Table 1). In order to filter the occasional presence of stratospheric modes the upper 8 vertical levels were excluded from the final TE norm.

The targeted perturbations were only integrated up to day 5, rather than day 10 for the ECMWF EPS.

For TEPS, two data sets have been generated, a winter 96-97 set and an autumn 97 set. Each set contains about 20 ensembles.

Before this set was created, as a test, a number (20) of smaller experiments was performed. Here, only 20 members were integrated, only up to three days, and the initial perturbations were not rotated. The discussion in the next section is on the basis of these test ensembles.

3. NONLINEARITIES

3.1 Linear method (KEPS)

The KNMI ensemble prediction system (KEPS, *Barkmeijer et al.*, 1996) assumes that for the short range (up to 72h) error growth can be regarded as linear. In addition, it is assumed that the evolution of an initial error $\epsilon(0)$ from the operational T213 forecast can be approximated by the tangent linear T42QG model of *Marshall and Molteni* (1993):

$$\epsilon(T) = R(0,T)\epsilon(0), \tag{1}$$

where $R(0,T)$ is the resolvent of this tangent linear model. Therefore, once the evolution of the fasted growing singular vectors ν_i is known, the evolution of any error pattern within the subspace spanned by the initial $\nu_i(0)$, is known:

$$\epsilon(0) = \sum_{i=1}^{n} c_i \nu_i(0) \Rightarrow \epsilon(T) = \sum_{i=1}^{n} c_i \nu_i(T). \tag{2}$$

Any perturbation can be constructed in this subspace without having to perform an extra integration with the T213 model.

3.2 Investigation of nonlinear behaviour

For the Targeted Ensemble Prediction System (TEPS) it is tried to extend these ideas. By applying the same procedures to the non-linearly integrated singular vectors, the physics and resolution missed by the T42QG is recovered. The nonlinearity is assumed not to be too severe for the first few days, so that the linear assumption, and thus the creation of a large ensemble at low cost is still feasible. Let the trajectory of member i be given by $x_i(t)$, and of the control by $x^c(t)$. Then the initial perturbation $\epsilon_i(0) = x_i(0) - x^c(0)$ evolves according to:

$$\epsilon_i(T) = M(\epsilon_i(0)), \quad \text{where} \quad M(\epsilon(0)) = R_{T159}(x^c(0) + \epsilon(0)) - R_{T159}(x^c(0)). \tag{3}$$

Here R_{T159} is the resolvent of the T159 L31 ECMWF model.

The linear assumption is tested by comparing the evolution of two initially opposite perturbations throughout the entire integration period. Let $\epsilon_i^-(0) = -\epsilon_i^+(0)$ be the initial perturbation in the negative respectively positive direction of singular vector $\nu_i(0)$, then

$$M \text{ linear} \quad \Rightarrow \quad b_i(T) \equiv \frac{1}{2}(\epsilon_i^+(T) + \epsilon_i^-(T)) = \frac{1}{2}(M(\epsilon_i^+(0)) + M(-\epsilon_i^+(0)) \approx \frac{1}{2}M(0) = 0 \tag{4}$$

For the ECMWF EPS and for TEPS, in each direction two opposite perturbations are integrated. Therefore this requirement for linearity can easily be tested.

This will be done for the forecast starting on February 17th 1997, where a deep low came from the North Atlantic and swept over the British Isles at day 3. The control forecast for T=72h is given in the left top panel of Figure 1.

In order to keep the situation as transparent as possible, the initial perturbations were not rotated: $\epsilon_i^+(0) = \alpha\nu_i(0)$ and $\epsilon_i^-(0) = -\alpha\nu_i(0)$, where $\nu_i(0)$ is singular vector number i and $\alpha = 0.6$ is chosen such that the size of the initial perturbations is representative for the initial error.

The evolution of b_1 is given in the middle right panel of Figure 1. In the middle left panel,

$$a_i = \frac{1}{2}(\epsilon_i^+ - \epsilon_i^-) \tag{5}$$

for $i = 1$ is displayed. This field can be regarded as the best linear approximation of the error growth. By comparing b_1 with a_1 it is seen that at 72 hours, although very localized, the impact of the nonlinearity is of the same order of magnitude as the linearity. These nonlinearities already start to emerge after 48 hours (not displayed). Therefore, at first sight, the extension of the linear assumption seems impossible.

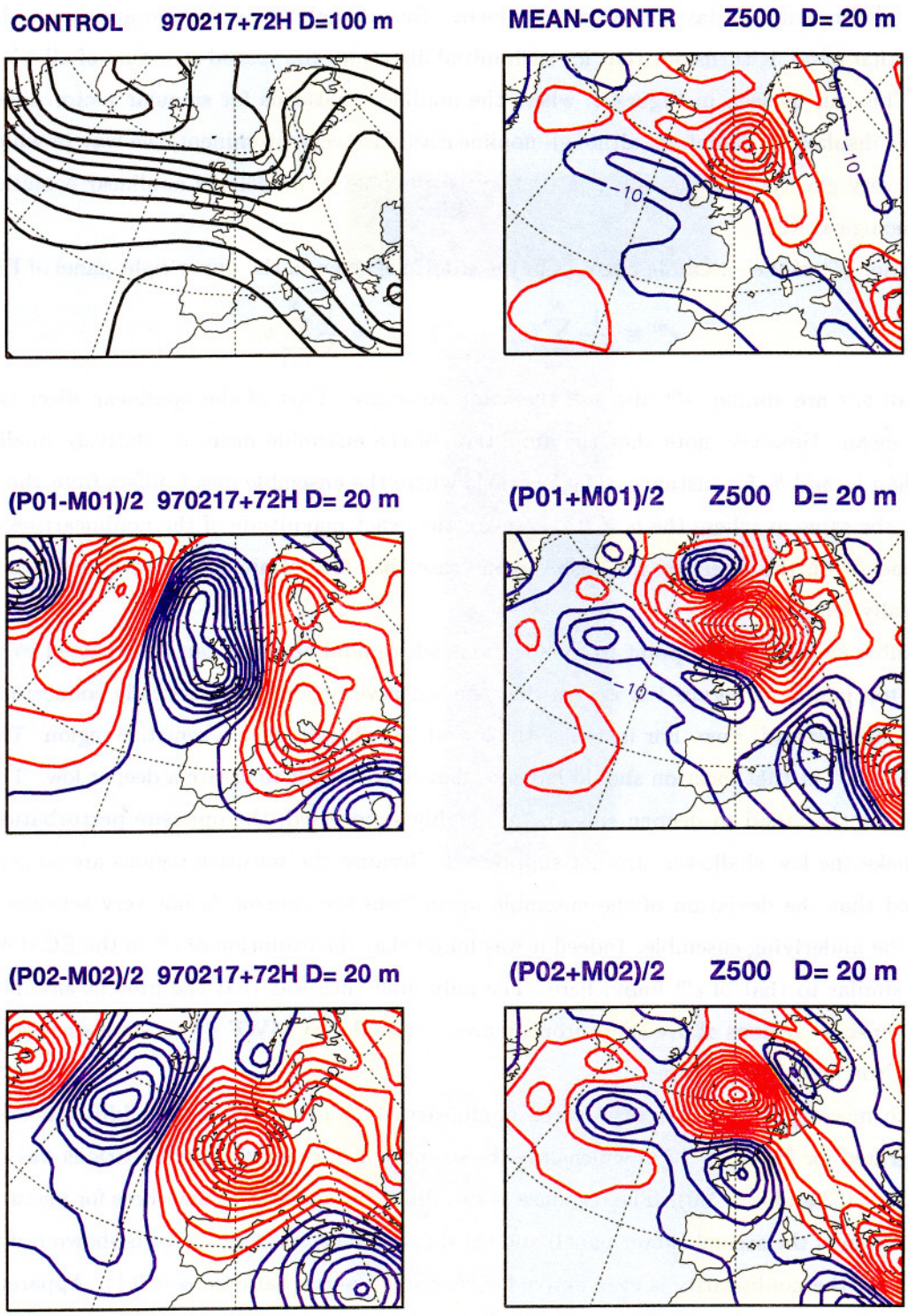

Figure 1: Patterns for a 3-day forecast starting at February 17, 1997, for the 500 hPa geopotential. The top panels show (left) the control forecast and (right) the ensemble mean with respect to the control. The middle panels show (left) the best linear approximation (5) and (right) nonlinear pattern (4) in the first singular vector direction. The lower panels show the same for the second singular vector direction.

In the lower panels of Figure 1, the same evolution for the second singular vector is displayed. Also here, the nonlinearity at day 3 is very prominent. Surprisingly, its structure appears to be very similar to that of b_1. It turns out that for each initial direction, the spatial structure of all b_i's is very similar. This can be seen in Figure 2, where the nonlinear patterns for singular vector direction 3 up to 8 are displayed. Therefore, although nonlinearities are very prominent, the way in which they emerge is very generic. This means that it may be possible to model the nonlinear evolution in a phenomenological way.

The deviation ϵ^m of the ensemble mean from the control is given in the upper right panel of Figure 1:

$$\epsilon^m \equiv \frac{1}{2N}\sum_{i=1}^{N}(x_i^+ + x_i^-) - x^c = \frac{1}{N}\sum_{i=1}^{N} b_i. \tag{6}$$

Because all b_i's are similar, ϵ^m also has the same structure. Part of the nonlinear effect is in the ensemble mean. However, note that the amplitude of the ensemble mean is relatively small, much smaller than b_1 and b_2 for instance. The locations where the ensemble mean differs from the control are about the same as where the $b_i \neq 0$. However, the exact magnitude of the nonlinearities considerably depends on the direction i. It is even the case that for certain b_i's certain sensitive regions are not really triggered.

The ensemble mean seems to point to the locations where nonlinearities can be expected to emerge. From the top panels of Figure 1 it is seen that the sensitive regions are physically connected to the control (= background) flow. For instance, the low at day three is such a sensitive region. The large magnitude of ϵ^m at that location should indicate that it is difficult to create a deeper low. Therefore perturbations that tend to deepen this low are highly suppressed, the opposite perturbations that tend to make the low shallower, are not suppressed. Because the sensitive regions are so generic, it is expected that the deviation of the ensemble mean from the control, is not very sensitive to the choice of the underlying ensemble. Indeed it was found that the evolution of ϵ^m in the ECMWF EPS was very similar to that of ϵ^m found here. The only difference was that the present ensemble only triggered sensitive regions above the European area, while the ECMWF ϵ^m exposed such regions for the whole Northern Hemisphere.

In order to investigate the structure of the nonlinearities more deeply, three additional ensembles were integrated for February 17 in which only the strength α of the initial perturbations was changed ($\alpha = 0.1$, $\alpha = 0.3$ and $\alpha = 1.0$). The response at two different locations at day three for perturbations in the direction of the second (lower panel) and third (top panel) singular vector is shown in Figure 3. It is seen that the nonlinearity is even active for very small perturbations ($\alpha = 0.1$). Apparently the scale on which these nonlinear effects emerge (which might be a property of the underlying physics?) is much smaller than the typical scale of the uncertainty in the initial state ($\alpha \sim 0.6$). In addition, it is seen that when perturbing in the same direction, the linearity of the system is much better obeyed than comparing two perturbations in opposite directions, that is,

$$M(c\epsilon) \approx cM(\epsilon), \quad \text{for} \quad c > 0. \tag{7}$$

For large perturbations a nonlinear saturation will obviously occur. But it should be pointed out

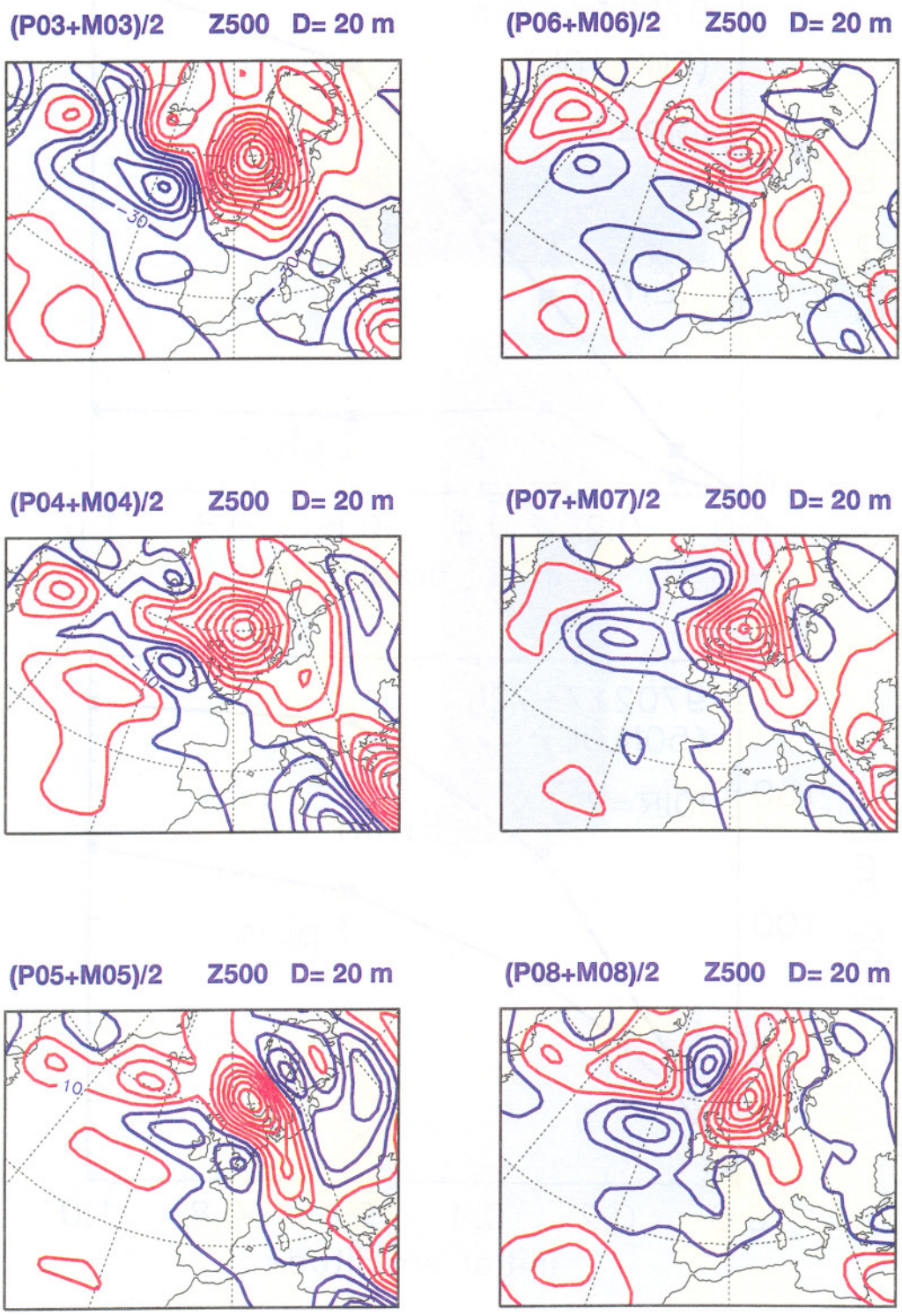

Figure 2: Nonlinear patterns for a 3-day forecast starting at February 17, 1997, for the 500 hPa geopotential. Patterns (4) in the 3th up to 8th singular vector direction are plotted.

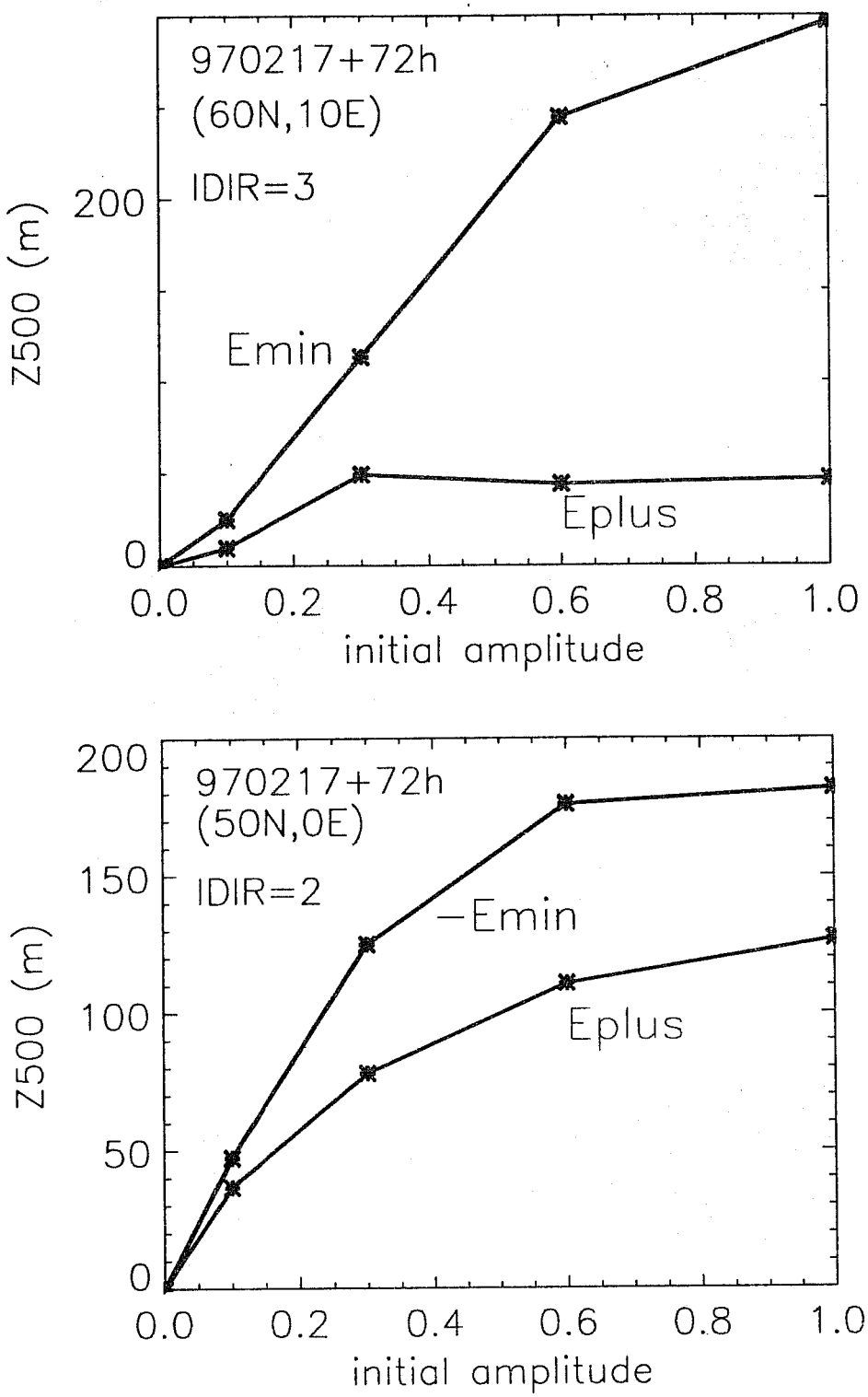

Figure 3: Response of the 500 hPa geopotential for a 3-day forecast starting at February 17, 1997. The top panel shows the response at (60N, 10E) for initial perturbations of 0.1, 0.3, 0.6 and 1.0 times the 3th singular vector (Eplus), and -0.1, -0.3, -0.6 and -1.0 this vector (Emin). The lower panel shows the same at (50N,0E) and in the 2nd singular vector direction.

that this saturation is a gradually growing form of a nonlinearity, such as $y = ax - bx^2$, while the nonlinearity around the origin has a more abrupt nature, like $y = ax + b|x|$ (which is supported by the observation that the nonlinear indicators b_i appear to be more or a less linear to the strength α, rather than quadratic).

The above described observations were made for one case, namely 17 February 1997, and for one parameter (geopotential at 500 hPa). However, similar results were found for all cases considered and for different parameters, such as the geopotential at 1000 hPa and the temperature at 850 hPa.

3.3 Nonlinear extension of KEPS

In the previous subsection it was observed that a linear expansion of initial perturbations given by (2) is for a nonlinearly integrated ensemble, like TEPS, not valid, even for the short range. However, it was seen that the nonlinearities are localized and emerge in a generic way. This enables the opportunity to construct an empirical nonlinear extension to (2).

Given an initial perturbation $\epsilon(0)$, chosen within the subspace spanned by N initial singular vectors $\nu_i(0)$:

$$\epsilon(0) = \sum_{i=1}^{N} c_i \nu_i(0), \tag{8}$$

the question is how to predict its temporal evolution

$$\epsilon(T) = M(\epsilon(0)) = \epsilon^T(c_1, \cdots, c_N), \tag{9}$$

on the basis of the $2N$ nonlinearly integrated members

$$\epsilon_i^{\pm}(T) = \epsilon^T(0, \cdots, \pm\alpha, 0, \cdots), \tag{10}$$

without having to perform an extra nonlinear integration.

In the previous subsection it was argued that the nonlinearity is a very special type of a nonlinearity that has a rather abrupt nature. In one direction the linear growth is different from the linear growth in the opposite direction. For instance, an initial perturbation that will try to deepen a low at a certain forecast time, could grow less than the opposite perturbations that will tend to make the low shallower. Therefore, the set of initial perturbations is divided into two disjunct regions (I and II). The regions are each mirror images, that is, when $\epsilon \in I$ than $-\epsilon \in II$. Therefore in each direction i a choice must be made which $\epsilon_i^{\pm}(0)$ belongs to region I (which will be renamed by ϵ_i^I), and which to region II (ϵ_i^{II}). In the case of the low, it is clear that all perturbations that tend to weaken the low will belong to one region (say I), the remaining N will belong to II. In a more general case, it can be decided that the in absolute value fasted growing perturbation is chosen to be in I, the other in II.

The second assumption is that inside the two regions the evolution can still be regarded as linear. When an initial perturbation is in region I for instance, its evolution is expected to be a linear combination of the members ϵ_i^I. In general, let

$$\epsilon(0) = \sum_{i=1}^{N} a_i \epsilon_i^I(0) = \sum_{i=1}^{N} -a_i \epsilon_i^{II}(0), \tag{11}$$

where $a_i = \pm c_i/\alpha$ (depending on $\epsilon_i^+ \in I$ or II), then

$$\epsilon \in I \quad \Rightarrow \quad \epsilon(T) = \sum_{i=1}^{N} a_i \epsilon_i^I(T) \tag{12}$$

$$\epsilon \in II \quad \Rightarrow \quad \epsilon(T) = \sum_{i=1}^{N} -a_i \epsilon_i^{II}(T) \tag{13}$$

The question now is: when is ϵ chosen to be situated in I and when in II. It is clear that perturbation $\epsilon = (0, \cdots, a_i, 0, \cdots)$ belongs to I if $a_i > 0$ and to II when $a_i < 0$. For an initial perturbation which is a combination of more than one ϵ_i^I, such an identification is much less clear. However, it is possible to find the intersection between both regions by demanding that prescription (11-13) is continuous. For $\epsilon \in I \cap II$ one should have:

$$\sum_{i=1}^{N} a_i \epsilon_i^I(T) \quad = \quad \sum_{i=1}^{N} -a_i \epsilon_i^{II}(T) \tag{14}$$

$$\Rightarrow \qquad d \equiv \sum_{i=1}^{N} a_i b_i(T) = 0. \tag{15}$$

Therefore the intersection between I and II is formed by the plane that is normal to the vector $b_1, \cdots, b_N$. For an arbitrary perturbation ϵ, the sign of inner product d can be exploited to identify to which region the perturbation belongs, because all ϵ with the same sign of d must belong to the same region. For $\epsilon = \epsilon_i^I$ one has $d = b_i(T)$. In order to guarantee that these perturbations all belong to I it is essential that the sign of all $b_i(T)$ is equal. In the cases considered it was indeed found that at the locations that are sensitive to the development of nonlinear behavior, the sign of b_i was quite robust. Therefore one can summarize the condition to:

$$\text{sign}(d(\epsilon)) = \text{sign}(e^m) \quad \Rightarrow \quad \epsilon \in I \tag{16}$$

$$\text{sign}(d(\epsilon)) = -\text{sign}(e^m) \quad \Rightarrow \quad \epsilon \in II \tag{17}$$

This completes the definition of the nonlinear extension to (2). It is to be regarded as a first approximation to describe the nonlinear behavior of the system in the short and intermediate range. It fills the gap between the lines shown in Figure 3. Note that the classification of the initial perturbations in region I and II depends on the final time and location one is interested in. For instance, for the location of a low at one forecast time, it could be the case that both ϵ_1^+ and ϵ_2^- are classified to belong to the same region, while the position of a ridge at a different location and forecast time would suggest that both should belong to different regions. This paradox however is not a conflict, because due to the continuity of the integrated ensemble members in space and time, (11-17) will result into an unambiguous continuous prescription. The classification into two separate regions is only a conceptual aid.

The piecewise linear prescription (11-17) only differs from the linear expansion (2) at locations where the nonlinear effects are important. In regions where $b_i \approx 0$ prescription (11-17) reduces to the linear one (2).

3.3 Large nonlinear ensemble at low cost

With the nonlinear extension defined in the previous subsection, an arbitrary ensemble member can be created at low cost. Consider a large ensemble $\gg N$ of initial perturbations $\epsilon(0) = (c_1, \ldots, c_N)$ drawn from a Gaussian distribution:

$$< c_i >= 0, \quad < c_i c_j >= (\sigma^2/N)\,\delta_{ij}. \tag{18}$$

Here the brackets denote the expectation value. Then it can be derived that the expectation value of $\epsilon(T)$ is given by:

$$< \epsilon(T) > \;=\; (\sigma/\alpha) \left(\sqrt{\frac{2}{\pi}\frac{1}{N}\sum_{i=1}^{N} b_i^2} \right) \mathrm{sign}(\epsilon^m) \tag{19}$$

$$< \epsilon^2(T) > \;=\; (\sigma/\alpha)^2 \frac{1}{2N} \sum_{i=1}^{N} \left(\epsilon_i^{+2} + \epsilon_i^{-2} \right). \tag{20}$$

First of all it is seen that when $\sigma = \alpha$ is chosen, the spread of such an ensemble is the same as the spread one receives by only regarding the 2N nonlinearly integrated members. The ensemble mean, however, is somewhat different from the ensemble mean defined by (6). The structure and sign of $< \epsilon^2(T) >$ is equal to that of ϵ^m, its magnitude, however, will be different. Expression (20) will stress the strongest b_i more than ϵ^m does. In the special case all b_i are equal (only then the nonlinearity is in the ensemble mean), one finds $< \epsilon^2(T) >= \sqrt{2/\pi}\epsilon^m$, that is, about 20% weaker than the ensemble mean based on the 2N members.

The above mentioned expressions are exact. This means one does not actually need to create a large ensemble, in order to receive these results. However, if one is interested in more detailed information about the PDF, the creation of such an ensemble is desirable. In Figure 4 the one-dimensional restriction of the PDF of the 500 hPa geopotential at (60.47N, 5.63E) is shown for the three day forecast starting at 17 February 1997. This location corresponds to the location of the low (see left top panel of Figure 1). The black curve is the prediction of the PDF on the basis the linear method. Using (2) for an ensemble of 10,000 members and the T42 L31 ECMWF linearly evolved (adiabatic) singular vectors, the final PDF is Gaussian and rather broad. The grey curve is on the basis of an ensemble consisting of the same initial perturbations, but now evolved according to (11-17). The difference is striking. It is seen that the deepening of the low is very unlikely. The verifying analysis at day three is given by the vertical solid line. The low indeed appeared to be less deep than was predicted by the control. Therefore, in this case, the nonlinear ensemble based on (11-17) gives more information than the KEPS ensemble.

4. SCORES

An objective way to test the surplus value of TEPS with respect to the ECMWF EPS, is to test: Ranked Probability Scores, Brier Skill Scores, Reliability and Talagrand diagrams. In this section some preliminary results on the performance of TEPS are given. Scores were calculated on the basis of the 50-member 5-day ensembles.

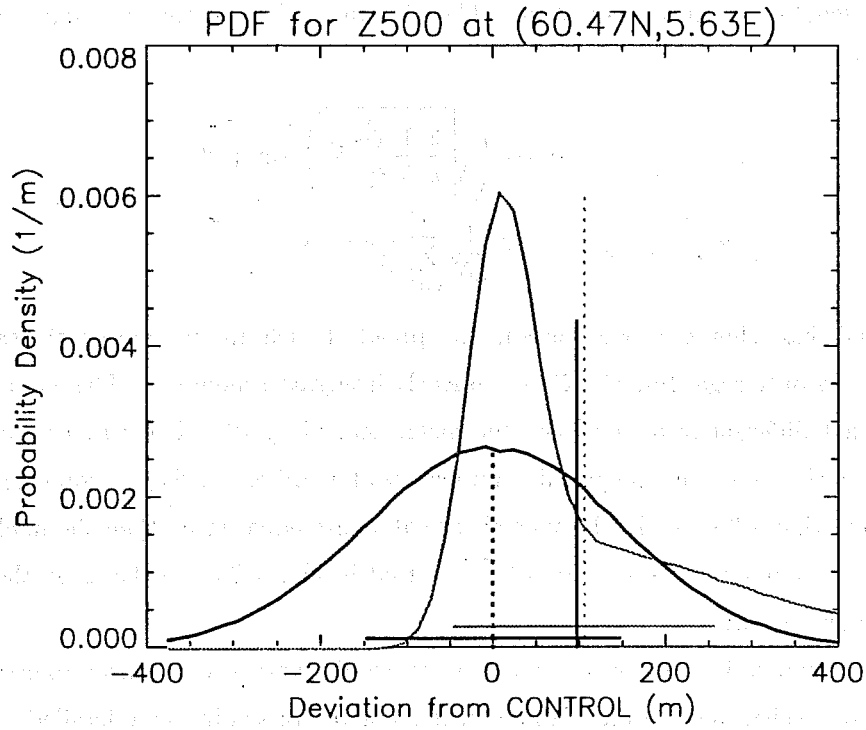

Figure 4: One-dimensional Probability Density Function of the geopotential at 500 hPa at (60.47N, 5.63E) for a 3-day forecast, starting at February 17, 1997. The broad Gaussian curve is a result of the linear prescription (2); the sharp, asymmetric curve is a result from nonlinear prescription (11-17). The solid vertical line indicates the verifying analysis

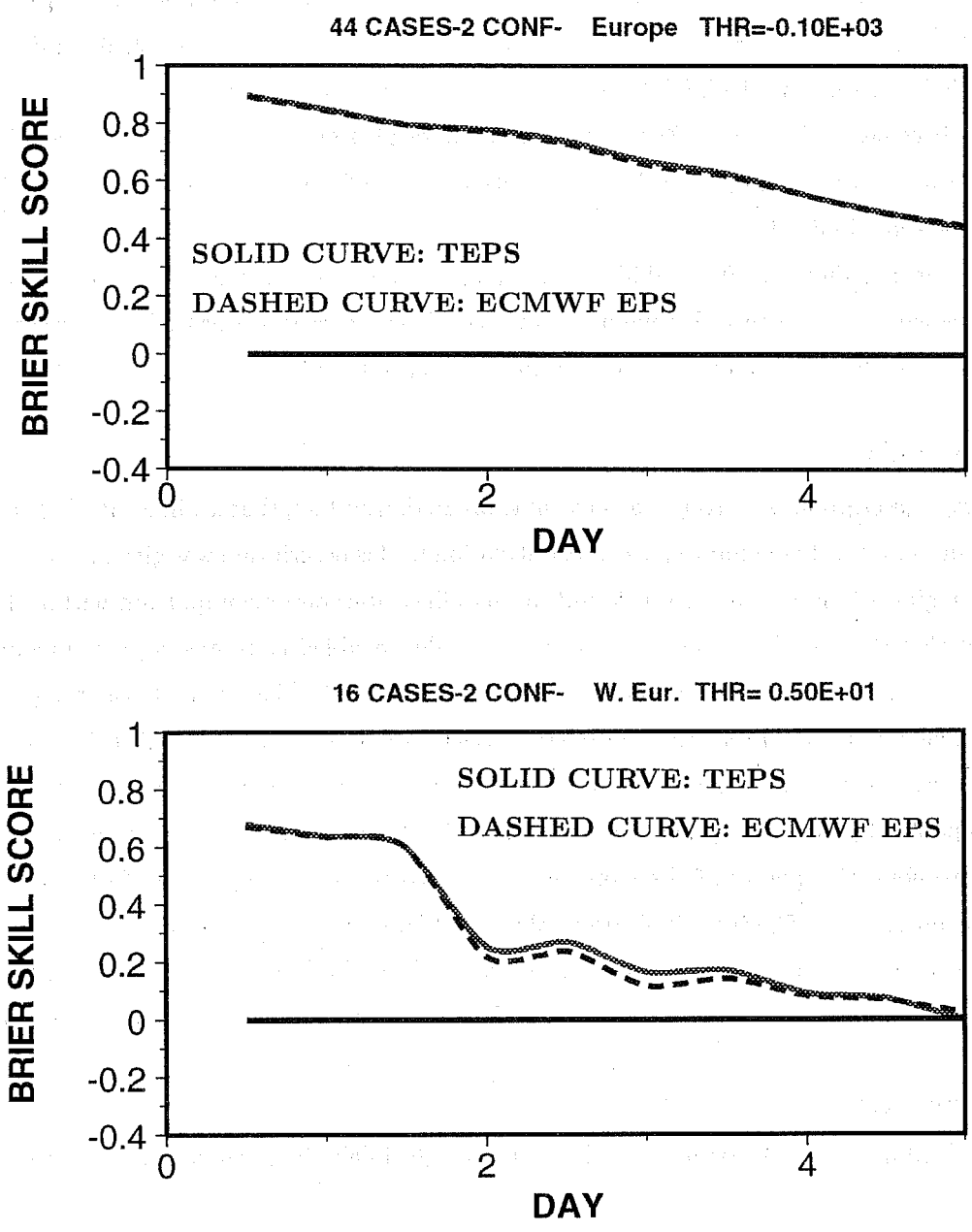

Figure 5: Brier Skill Scores for TEPS (solid lines) and the ECMWF EPS (dashed lines). The top panel shows the BSS for the geopotential at 1000 hPa averaged over 44 cases for Europe (30°N-75°N, 20°W-45°E) in the winter of 96/97 and autumn of 97. The lower panel shows the BSS for Large Scale Precipitation averaged over 16 cases in the autumn of 97 for Western Europe (34.5°N-70.5°N, 21°W-21°E). The threshold for the top panel is -100m and 5mm/12h for the lower panel.

In the top panel of Figure 5 the Brier Skill Score for a deviation more than 100 m below the 1000 hPa geopotential climate, averaged over Europe is given. The dashed curve refers to the ECMWF ensemble, the solid curve to TEPS. They are based on 44 cases in the winter of 96/97 and autumn 97. From this it is seen that the TEPS ensemble only leads to a small improvement. At day 5, the performance becomes a bit worse, which is obviously the result from the fact that the perturbations propagate out of the target area. For the geopotential at 500 hPa and the temperature at 850 hPa similar results were obtained.

In the lower panel of Figure 5, Brier Skill Scores for more than 5 mm Large Scale Precipitation per 12 hours, averaged over Western Europe are displayed. These curves are based on 16 cases in the autumn of 97. For this parameter, the positive impact is meaningful.

5. CONCLUSIONS

In this report, a description was given of an ensemble prediction system, that is aimed at the European area for the short and early medium range. Most attention in this description was given to the validity of the linear regime of the system. It was found that nonlinearities become important within the first few days but that the way they emerge is very generic. This enabled us to give a phenomenological description that can be regarded as a first approximation to an extension of the linear range.

It was seen that for the standard EPS parameters, Z500, T850 and Z1000, the TEPS system does not lead to a real improvement of scores. The reason for this may be, that the singular vectors used (TE), are capable of giving a very good description of the for these parameters important parts of the PDF. Therefore, the quality of the scores is not limited by the quality of the EPS, but by the quality of the initial field. There are indications that TEPS does lead to an improvement for weather parameters, like precipitation. This is what was expected. In addition it is expected that TEPS will give rise to a better performance in extreme cases. Research in this direction is being performed.

6. REFERENCES

Barkmeijer, J., Mureau, R., Oortwijn, J., and J.D. Opsteegh, 1996: Research on regional predictability at KNMI.

Marshall, J., and F. Molteni, 1993: Toward a Dynamical Understanding of Planetary-scale Flow Regimes. J. Atmos. Sci.,50,1792-1818.

Strauss, B., and A. Lanzinger, 1996: Verification of the Ensemble Prediction System (EPS). ECMWF Newsletter, 72,9-15.

Talagrand, O., Vautard. R. and B. Strauss, 1997: Some results on the evaluation of the ECMWF Ensemble Prediction System. ECMWF Scientific Advisory Committee, 26th Session, item 9.

THE RELATIVE EFFECTS OF MODEL AND ANALYSIS DIFFERENCES ON ECMWF AND UKMO OPERATIONAL FORECASTS

D S Richardson
ECMWF

Summary

A set of 25 forecasts of the ECMWF model initialised from UKMO analyses has been run. These forecasts have been compared with the corresponding EPS control and UKMO operational forecasts to assess the relative effects of using different models and analyses. Analysis effects are found to predominate throughout the forecast (results are available to day 6). The mean of the EPS control and UKMO operational forecasts has substantially lower rms error than either contributing forecast; it is demonstrated that most of this improvement can be achieved by combining ECMWF forecasts from the two analyses, although a small additional benefit from the two-centre mean remains which may be due to the use of different models. It is further shown that a single ECMWF forecast run from a mean of the ECMWF and UKMO analyses has on average the same or even lower error than the mean of forecasts from the two different analyses.

1. Introduction

The aim of this study is to examine the contributions of analysis and model effects to the differences observed between operational forecasts from ECMWF and UKMO. The objective is to assess whether the forecast differences are due mainly to running from different analysed initial states or whether it is the use of different models which is the main cause. An ensemble mean of the two forecasts has on average lower rms error than either individual forecast and again the relative contribution of analysis and model differences to this improvement are examined. If analysis effects are found to predominate then it may be that a major portion of the benefits obtained from running joint ensembles using both ECMWF and UKMO models and analyses may be due to the use of two analyses. A further objective is to study the effect of running a single forecast from an average of the ECMWF and UKMO analyses - it may be expected that the combined analysis will be a better estimate of the true atmospheric state and hence that a single forecast from this analysis would be better than a forecast run from either of the independent analyses.

2. Method and data

Because the ECMWF and UKMO models do not use the same horizontal representation, vertical resolution, orography or even the same primary variables, it is not possible to run the ECMWF model from exactly the UKMO analysis - some degree of interpolation (and hence error) is inevitable. To attempt to minimize interpolation errors the following method is adopted to construct an analysis close to the UKMO analysis. The difference between UKMO and ECMWF upper air fields (u, v, T, and specific humidity) is constructed from data available on 15 standard pressure levels. This difference is then interpolated to ECMWF model levels and the resulting perturbation is added to the original ECMWF analysis. A forecast run using the ECMF model from this perturbed analysis is referred to as being from the UKMO analysis.

Three sets of forecasts are compared to assess analysis and model impacts: the UKMO operational forecast, the EPS control forecast (run at T_L159) and forecasts run with the ECMWF model (again at T_L159) from the UKMO analysis. A set of 25 cases is studied consisting of forecasts initialized every 5 days from December 1996 to April 1997. Verification is concentrated on the rms error and rms difference between forecasts of the 500 hPa height field over the Northern and Southern Hemispheres (latitudes 20° - 80°) and is against the ECMWF analysis unless stated otherwise. The UKMO operational forecast is only run out to T+144 hence verification is limited to this forecast range.

Differences between the ECMWF forecast from the UKMO analysis (ECuka) and the EPS control are classed as analysis effects while those between the UKMO forecast and ECuka are taken to represent model errors. It should be borne in mind that since ECuka is initialized from an approximation to the UKMO analysis, effects counted as "analysis effects" are in fact a lower bound on the influence of different analyses and that "model effects" include any residual analysis effects not captured by the interpolation procedure.

3. Results

There is a clear difference in the mean scores over all cases of the EPS control and UKMO operational forecasts (Figure 1). The mean scores of the ECuka forecasts are remarkably close to those of the UKMO forecasts. On a case by case basis the ECuka forecast is generally closer to that of the UKMO although by Day 5 there are notable differences in skill on some occasions (Figure 2).

The relative size of analysis and model impacts is assessed by consideration of the rms difference between forecasts and also by correlation between difference fields; the proportion of the total variance of the UK-EPS control difference is given by the square of the correlation. The unexpectedly large differences between the UK and ECuka forecasts at initial time are a result of the multiple interpolation of the initial difference from pressure levels to model levels and back to pressure levels and give some indication of the inherent uncertainties in the interpolation between analyses. Throughout the forecast range the analysis effects are substantially greater than the model effects, the difference being especially marked in the Southern Hemisphere (Figure 3). Analysis effects account for around 50% or more of the total variance in the Northern Hemisphere, and for 60-90% in the Southern Hemisphere. The correlation between model effects and the total difference is surprisingly constant throughout the forecast range at about 50% in the Northern Hemisphere (explaining 25% of the total variance) and 35% in the Southern Hemisphere. Note that if systematic errors and differences in variance between forecast sets are ignored then the expected correlation asymptotes to 50% (i.e. purely random differences would give this level of correlation), thus the possibility that random rather than purely model effects make a significant contribution to the "model effects" cannot be excluded.

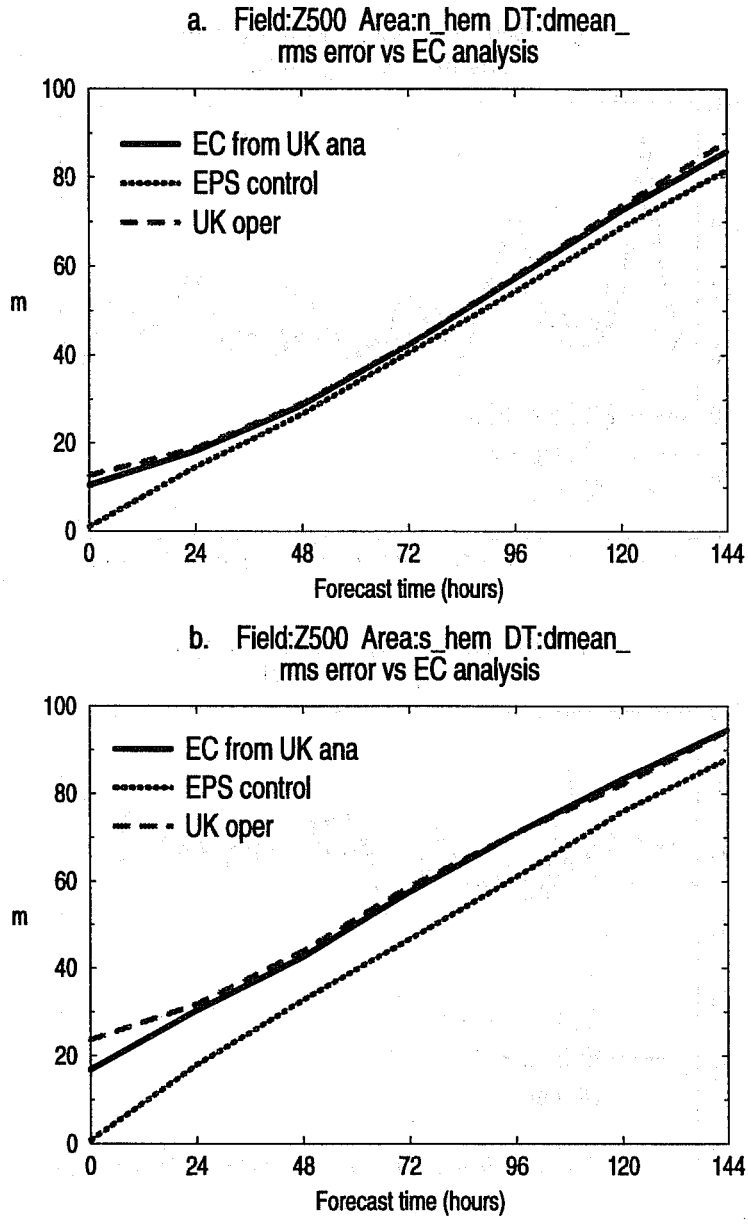

Figure 1: Mean rms error over 25 cases for 500 hPa height over a) Northern Hemisphere and b) Southern Hemisphere for EPS control, UKMO operational and ECMWF from UK analysis.

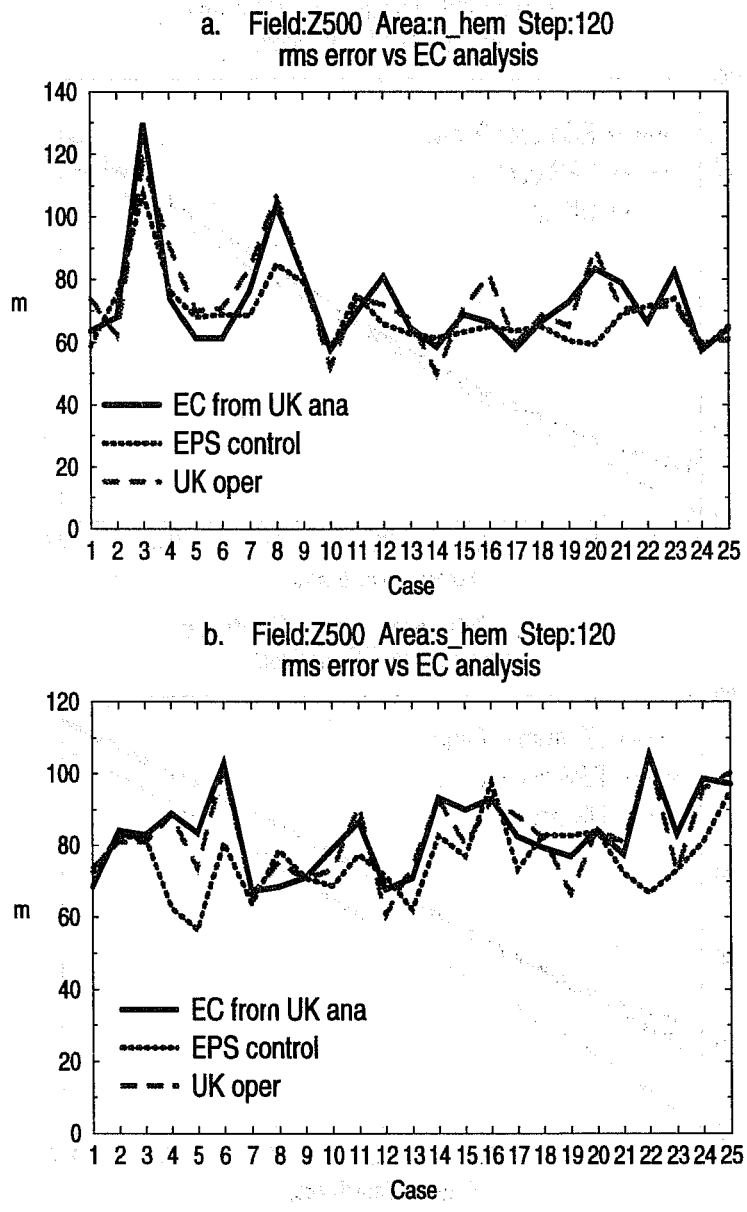

Figure 2: rms error at day 5 for 25 cases for 500 hPa height over a) Northern Hemisphere and b) Southern Hemisphere for EPS control, UKMO operational and ECMWF from UK analysis.

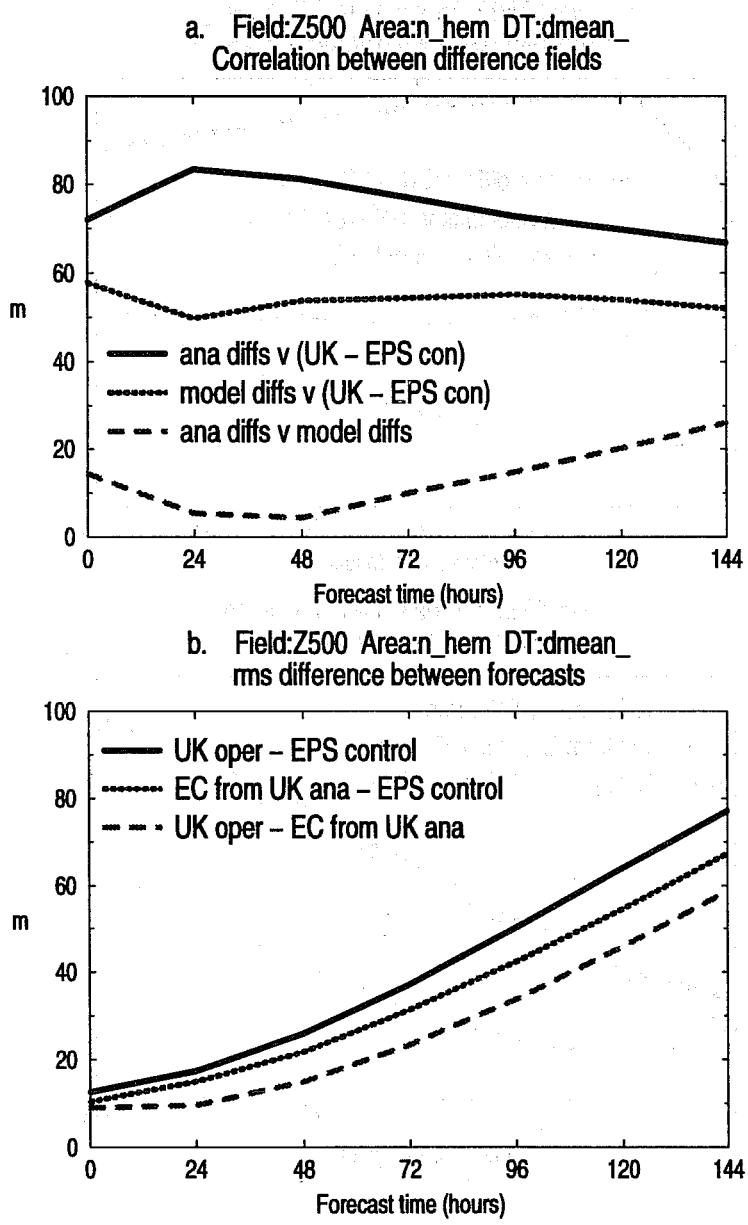

Figure 3: Mean difference scores over 25 cases for 500 hPa height over Northern Hemisphere. a) correlation between difference fields: ana diffs = (EC from UK analysis) - (EPS control); model diffs = (UKMO operational) - (EC from UK analysis). b) rms differences between forecasts.

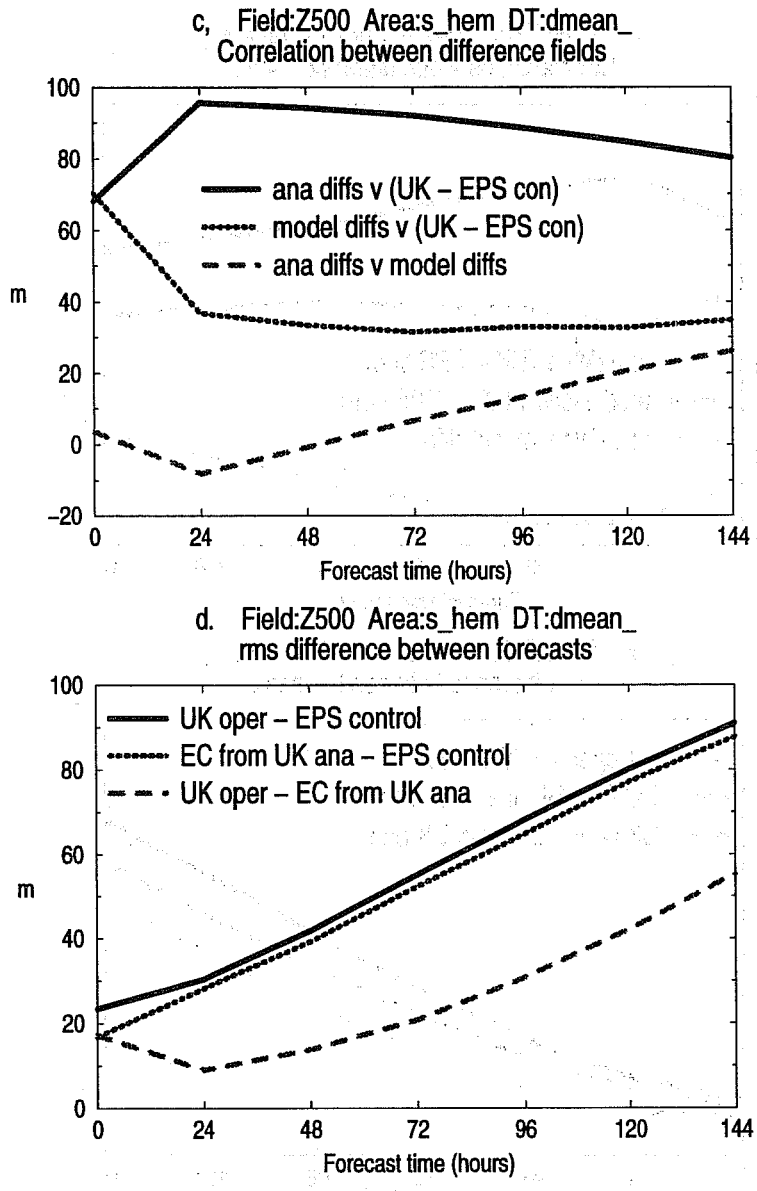

Figure 3 cont.: c), d) as Figure 3 a), b), but for Southern Hemisphere.

Looking at individual cases, analysis differences are greater than model differences on 80% of occasions at day 5 for the Northern Hemisphere and in all cases for the Southern Hemisphere.

An indication of the benefit of combining forecasts from different centres can be seen from the skill of an ensemble mean forecast made by averaging the UKMO and EPS control forecasts. The average rms error of this mean forecasts is substantially lower than that of either contributing forecast (Figure 4 - here forecasts are verified against corresponding analyses, the mean forecasts being scored against an average of the two analyses). The contribution of the different analyses to this improved performance can be seen from an ensemble mean of the ECuka and EPS control forecasts. It is apparent that a major proportion of the improvement can be attributed to analysis effects although there is still a small additional effect which may be due to the use of different models. In contrast to the improvement throughout the forecast range obtained by averaging forecasts from different analyses, an ensemble mean formed from the EPS control and the first perturbed EPS member has on average a larger rms error than the EP control at least until day 4 or 5. This raises the possibility that an ensemble incorporating information from different analyses may be able to provide an improvement over the current EPS, especially at the earlier forecast ranges.

Having demonstrated that a substantial reduction in rms error may be achieved by averaging forecasts run from different analyses, the question arises as to how much of this improvement may be achieved by running a single forecast from an average of the two analyses. To investigate this question, a further set of 25 ECMWF forecasts was run, initialised by adding half of the UKMO-ECMWF analysis difference to the ECMWF analysis. The average rms error of these forecasts is less than that of the EPS control forecasts at all forecast times but the reduction is not as large as that obtained by averaging the EPS control and UKMO forecasts. However a proportion of the reduction in error of the mean forecast is due to a reduction of variance in the resulting field; to take account of this, the rms errors have been normalised by dividing by the corresponding sample variance for each forecast time (area-averaged variance across all 25 cases). Following this normalisation, the error of the forecasts from the mean analysis is closer to that of the ensemble mean forecast, although the ensemble mean forecast still demonstrates an additional benefit, particularly in the Northern Hemisphere (Figure 5). Also shown in Figure 5 is the normalised error of the mean of the EPS control and ECuka forecasts: comparing this with the error of the forecast from the mean analysis gives a better indication of the impact of the purely analysis-related effects of running from a single averaged analysis compared to averaging forecasts run from each contributing analysis. On average it can be seen that the single forecast is as good as or even better than the ensemble mean forecast once the reduction in variance of the ensemble mean is taken into account.

The potential benefit of running an EPS perturbed about a mean of the ECMWF and UKMO analyses will be investigated.

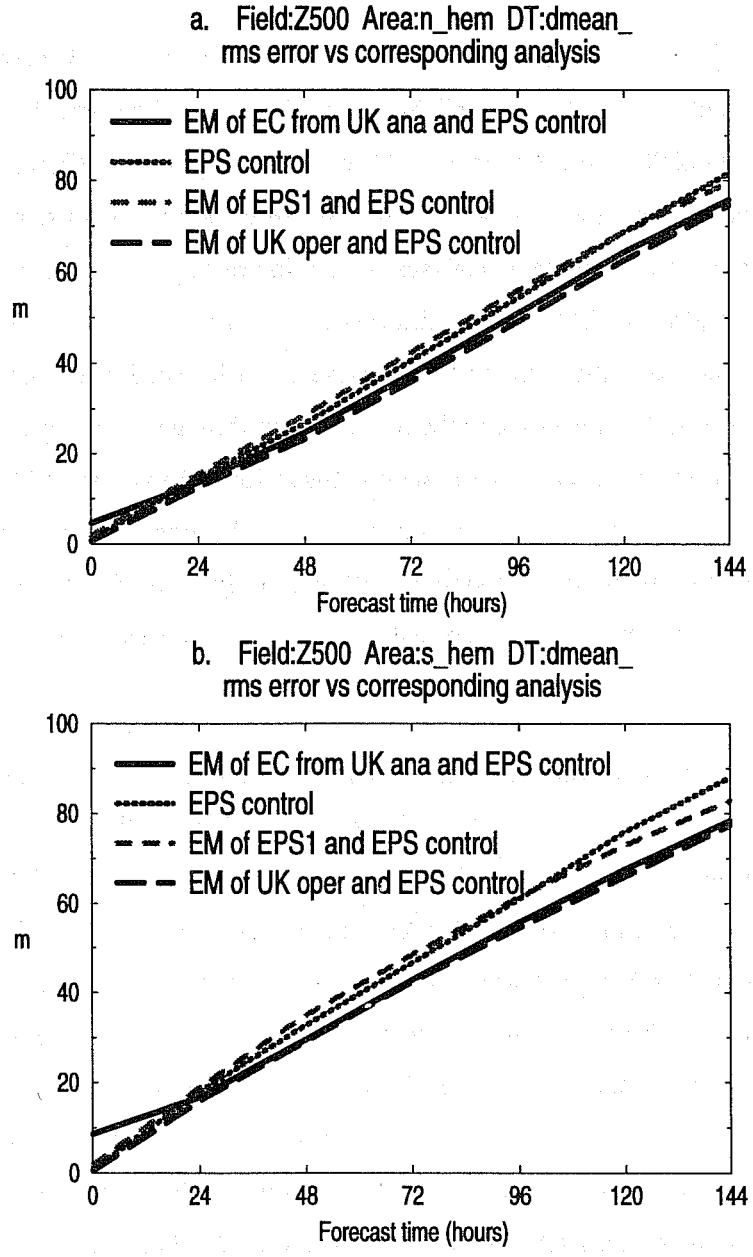

Figure 4: Mean rms error over 25 cases for 500 hPa height over a) Northern Hemisphere and b) Southern Hemisphere for EPS control, mean of ECMWF from UK analysis and EPS control, mean of UKMO operational and EPS control, mean of EPS perturbed member 1 and EPS control.

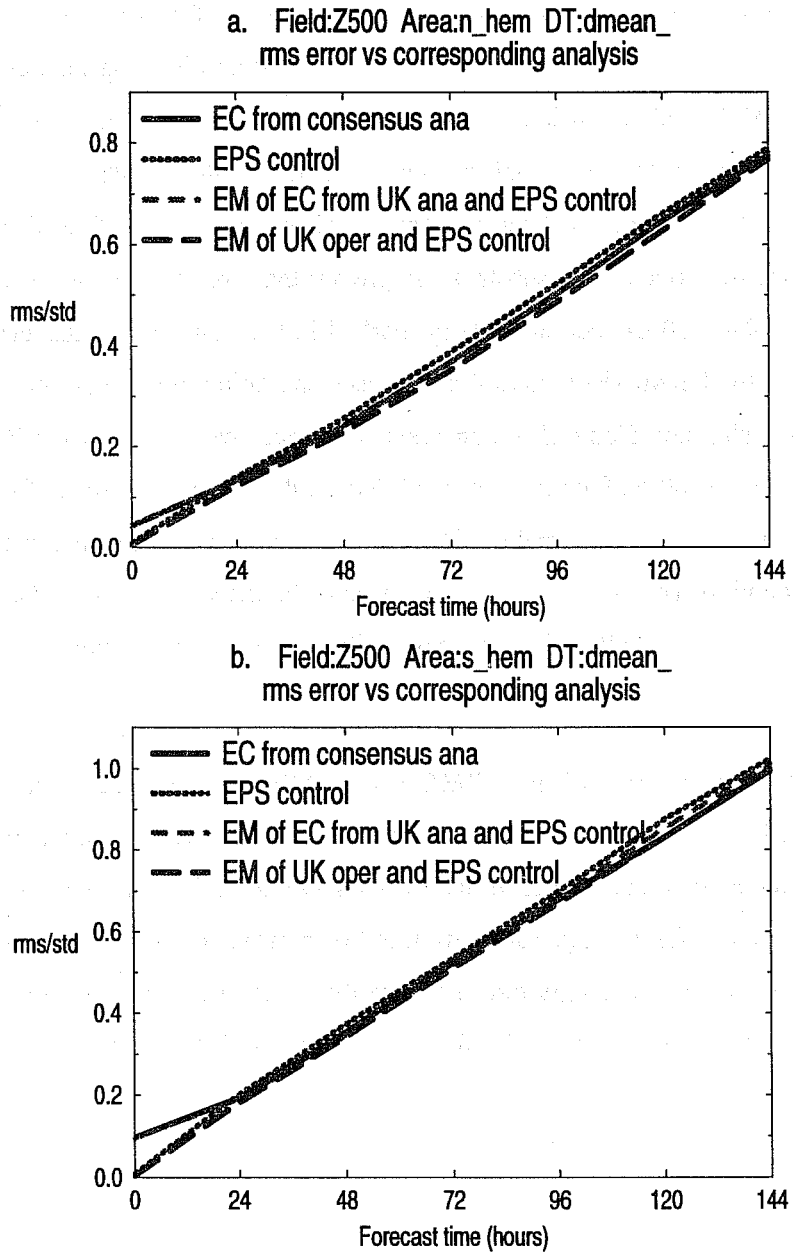

Figure 5: Mean rms error normalised by sample variance over 25 cases for 500 hPa height over a) Northern Hemisphere and b) Southern Hemisphere for EPS control, ECMWF from UK analysis, mean of ECMWF from UK analysis and EPS control, mean of UKMO operational and EPS control.

4. Conclusions

The relative effects of using different analyses and different models on forecast errors have been assessed using a set of 25 cases covering December 96 to April 97. A perturbation approximating the difference between UKMO and ECMWF analyses was constructed for each case and added to the ECMWF analysis. Forecasts with the ECMWF model from this analysis were compared with the EPS control forecast (analysis differences) and with the UKMO operational forecast (model differences). The effect of running one model from the different analyses was found to be substantially greater than that of running different models from the same analysis. The analysis effect was particularly marked in the Southern Hemisphere, accounting for 60-90% of the variance of the UKMO-EPS control differences and being larger than the model effect in all cases at day 5. Even in the Northern Hemisphere, analysis differences were greater in 80% of cases at day 5 and on average accounted for 50-60% of the total variance. Model differences accounted for only around 25% and 15% of total variance in Northern and Southern Hemispheres respectively. The analysis effects should be considered as a lower bound to the true influence of analysis differences since interpolation to the UKMO analysis is not perfect and the "model effects" may include differences due to this imperfect representation of the UKMO analysis.

ECMWF forecasts run from an average of the UKMO and ECMWF analyses have on average lower rms errors than either the UKMO or EPS control forecasts although not as low as an ensemble mean of the UKMO and EPS controls. However, after compensating for the reduced variance of the meaned forecasts, the forecast from the mean analysis accounts for the major proportion of the improvement of the ensemble mean forecast. Considering analysis effects alone, the single forecast from the mean analysis is as good as, or even better than, the mean of the EPS control and the ECMWF forecast from the UKMO analysis.