INTRODUCTION

Despite the enormous improvements in numerical weather prediction over the last decade, the practical usefulness of weather forecasts is limited by their day-to-day variability in skill. This variability is partly associated with fluctuations in the predictability of the flow, a measure of how inevitable uncertainties in the initial state of a forecast amplify through the integration due to flow-dependent instabilies of the atmospheric circulation. A prediction of predictability would be an important additional component of the weather forecast, gving an *a priori* assessment of its reliability.

In recent years there has been much interest in the question of forecasting predictability. This has derived from a number of factors. Firstly, statistical skill prediction schemes have shown that in the short range about half the variance in daily forecast skill scores can be forecast. In addition, the development of adjoint models has made it possible to integrate linearly the analysis error covariance matrix. Finally, the rapid increase in computer power will soon allow the integration of relatively large ensembles of forecasts. Such ensembles, as well as providing an estimate of predictability in the range where initial uncertainties are developing nonlinearly, enable probability forecasts of different outcomes to be made.

A workshop held on 13-15 November 1991 was devoted to a study of recent developments in the field of predictability. It was a logical follow-up to a workshop on the same topic held at ECMWF in May 1988. The sense that this is a field where substantial developments are imminent can be sensed through the papers presented in the Proceedings.

it also can take into account model generated errors, and thus give a more complete prognostic equation for the variance-covariance structure of forecast errors. However, some knowledge of model generated errors is also required.

1.3 STATISTICAL ANALYSIS OF NWP FROM DIFFERENT MODELS

Recent results have shown that a first estimate of the skill of a forecast can be gained by a comparison of the forecasts coming from different meteorological centres. Each Centre's forecast can be seen as a trajectory, characterised by a different initial condition obtained using different analysis techniques, and by the use of a different model for the integration. A statistical analysis of these forecasts can give a "lower bound" of skill for the prediction of forecast skill. More sophisticated predictability studies should be able to provide at least this information.

1.4 THE TANGENT MODEL METHOD

Provided the error E is small, its evolution is satisfactorily given by the tangent model R. This seems to work satisfactory up to the 72h range. The evolution of the error covariance matrix M is straightforward:

$$M(t) = R\ M(0)\ R^T$$

This method is nothing but a Kalman filter in prediction mode. It avoids the generation of a random sample required for the Monte-Carlo technique. If one is interested only in a small set of n variables given by F = PE, where P is a local projection operator (*Opsteegh*, this volume), the covariance matrix for F is given by:

$$N(t) = P\ R\ M\ (0)\ R^T\ \ P^T$$

One can compute explicitly the matrix $R^T\ P^T$ by n adjoint integrations, and subsequently one can obtain the local covariance matrix N(t) by n integrations only.

Experiments with simple models (barotropic, three-level baroclinic) and homogeneous initial matrix M(0) have given 24h forecasts of standard deviations which are significantly better than "trivial" skill forecasts. The use of initial variances from operational OI has improved the results in the Northern Hemisphere. Additionally, this technique allows the computation of the eigenvectors and eigenvalues of $R^T\ P^T\ P\ R$, which give the locally fastest growing (so called optimal) modes. The largest eigenvalue is related to potential loss of skill, as has been demonstrated up to the 72h range.

The experiments indicate that tangent models can be used to predict error statistics up to 72 hours. A larger number of experiments are required in order to further validate the use of tangent models. The problem of the effect of physical parametrizations should also be studied.

In a more general context, this method can be used as a numerical tool to study the following questions, which are relevant for any method of skill forecasting.

i) What is the accuracy of the initial covariance matrix given by OI, and what is its required accuracy for the skill prediction?

ii) What is the model error?

iii) Do we use appropriate methods to generate samples in Monte Carlo forecasting, at least for the short range?

1.5 ENSEMBLE INTEGRATIONS

Ensemble integrations of the full nonlinear diabatic NWP model may provide a useful tool for short-range skill prediction, cross-validating and extending linear techniques. In cross-validation mode, Monte Carlo forecasts can be used to assess the range of applicability of the assumption of linearity. In addition they can be used to assess the extent to which physical processes provide important nonlinear feedbacks onto the synoptic flow. An example of when the latter is important might be during rapid cyclogenesis.

Under conditions where linear and quasi-adiabatic assumptions hold, it will also be possible to assess from the linear models, the sample required to accurately predict error variances.

Providing the ensemble of the NWP model has adequate resolution, it may be possible to use short-range ensembles to give direct probability estimates of short-range weather elements, which linear models would be unable to provide directly. Such elements could include cloud cover, rainfall and other phenomena intrinsically nonlinearly related to the flow.

It is possible that if a high resolution NWP model was run many times over a relatively short period, a further medium- and extended-range ensemble, run with a lower-resolution model, could be chosen from the states at the end of the short range forecast ensemble e.g. by cluster analysis techniques. In this way ensemble requirements for the short, medium and extended range can be made consistent.

The use of linear models will continue to be important in constructing initial perturbations for ensemble predictions. Once reliable estimates of analysis error covariance are obtained, the optimal instabilities calculated from the linear models can be initialized by this covariance estimate.

1.6 RECOMMENDATIONS

- The covariance structure for the errors should be computed within OI.

- Despite the formidable cost of forecasting the error variance-covariance matrix using tangent linear models, we recommend the development of a low resolution KF. This would permit the performance of twin and real data experiments designed to compare the initial error covariance structures obtained from KF with those from OI and to obtain realistic initial error covariance fields.

- It is recommended that experimentation continues on the area of statistical skill prediction methods. It can be used as a reference when more sophisticated techniques become available.

- The linear and Monte Carlo techniques should be used in "twin" experiments with perfect models and a given initial matrix in order to prove either the validity of the adiabatic, linear hypothesis, or the method of sampling with fastest growing modes for the Monte Carlo forecasts.

- When computer power is adequate, ensemble forecasts should be run at a sufficiently high resolution ($\geq$ T106) to study the question of short-range predictability of both the synoptic-scale flow field and associated weather-related elements.

- In order to maintain a coordinated effort in predictability research, we recommend regular workshops and joint experimentation.

1. WORKING GROUP 1: TECHNIQUES FOR PREDICTION OF SHORT-RANGE FORECAST SKILL

1.1 INTRODUCTION

A valuable piece of additional information on short range forecast variables is provided by an estimate of their reliability. In order for a short-range reliability estimate to be more useful to a forecaster, it must be local in space. Estimates of local error variances can be produced with different techniques.

First of all statistical methods can be applied. These have been used and tested in some detail for large areas and their skill may form a reference for more advanced ways of predicting the reliability of a forecast. A promising way of computing the forecast error statistics follows from the assumption that errors evolve linearly. One can then use tangent models and their adjoints to integrate in time an initial covariance matrix for the errors. The derivation of a realistic estimate for the initial covariance matrix is as yet an unsolved problem. Probably the most advanced way of producing probabilistic forecasts is by Monte Carlo experimentation. Choosing a representative sample of initial states is again directly linked to the initial error distribution. Presently optimal modes for the growth rate of the kinetic energy are used but the approach needs validation.

In producing useful skill predictions, it is necessary to study in detail the dynamics of error growth. Experiments should be designed to understand error structures and their growth rates. Model errors must be distinguished from amplifying initial errors and their interactions must be studied. Finally the connection between errors in the short range and the medium and extended range must be clarified.

1.2 DEFINITION OF THE INITIAL ERROR FIELDS

Because the prediction problem considered here is the short-range forecast of skill, the accuracy of the actual forecast depends crucially on the accuracy of the initial conditions. Previous estimates of the errors of initial analyses have concentrated on the variances of the error fields. These estimates have been obtained from the comparison of analyses and observations, as a byproduct of optimal interpolation (OI) and from observing system simulation experiments (OSSE). For the purpose of predicting skill, however, the full covariance structure of the error fields is needed. In principle the covariance structure is also a byproduct of OI, however it is expensive to compute. Moreover, while this estimate is useful within OI, it is not widely believed to accurately reflect the true covariance structure of analysis error e.g. calling for the separability of the horizontal and vertical structure functions.

A more powerful method of producing the variance - covariance structure of the initial error distribution is to use the Kalman Filter (KF) technique. In addition to taking account of the advection of error covariance,

2. WORKING GROUP 2: TECHNIQUES FOR MEDIUM-RANGE ENSEMBLE FORECASTS

2.1 INTRODUCTION

In recent years the skill of deterministic forecasts provided by operational NWP centres has improved considerably in both tropics and extratropics. Forecasts of the Northern Hemisphere latitude flow usually provide useful information for 5 to 7 days; this time limit is shorter if one is concerned with weather elements in a particular region. However, depending on the user's needs, a time limit exists after which deterministic forecasts are no longer useful. This time limit is a function of the forecast skill which shows both seasonal variations and case-to-case variations related to the instability of the atmospheric flow during the forecast period.

It is therefore of the greatest importance that all users should be provided with an estimate of the forecast reliability, in order to make the best possible use of the NWP products. This estimate can take two forms:

i) a prediction of the probable range of skill scores which are usually adopted to quantify the forecast quality;

ii) a probabilistic forecast of weather and circulation parameters (or fields), from which the forecaster can have a direct estimate of the possible future states of the atmosphere.

So far, operational centres have performed medium-range predictions with a single deterministic forecast, complemented by experimental predictions of forecast skill derived from different techniques. We believe that in the future ensemble forecasting will be an essential method to provide reasonable probabilistic forecasts.

However, ensemble forecasting is a computationally expensive technique. Therefore, despite the continuous increase in computer power, a conflict will arise between the need of improving the resolution and sophistication of the operational model and the need to provide ensembles which are large enough to produce realistic probabilistic estimates. Two major approaches are possible:

i) continue to provide a 'control' forecast with the best possible model available, and run ensembles with models of lower resolution or simplified physical parametrization;

ii) run ensembles in which all forecasts are performed with the same version of the forecast model.

The approach (i) is currently in place in rudimentary form in several centres. However, even if the large-scale flow can be obtained with relatively simple models, an estimate of the reliability of near-surface weather parameters predicted by the control forecast may not be obtained if the perturbed forecasts are run with a model of much lower resolution or simpler physics so that the second approach should be given serious consideration.

2.2 REGIME TRANSITIONS

As recognised in the conclusions of a number of recent meetings on medium and extended-range predictions, transitions between large-scale flow regimes are probably the most difficult phenomena to predict. Therefore, in the experimental phase aimed at defining an "optimal" system for ensemble forecasting, particular emphasis should be put on evaluating the ability of the ensemble forecast to provide at least a good probabilistic estimate of the occurrence of regime transitions. The problem is made difficult by our limited knowledge of the dynamical processes that cause these transitions. However, there are strong indications that quasi-stochastic forcing by baroclinic waves is the most important (though not the only) process for regime transition in mid latitudes. The level of eddy kinetic energy in a model is a clear indicator of the ability of the model to simulate baroclinic waves correctly. This may be an important factor to take into account when choosing a model for ensemble forecasting. It is likely that models with insufficient eddy kinetic energy cannot represent regime transitions properly. Therefore, perturbed forecasts with such models may have significant different statistical properties (as far as regimes are concerned) from more advanced models used for the 'control forecast'. If lower resolution models are used for the perturbed forecasts, care should be taken that the physical parametrization is properly tuned to give realistic energy balances.

2.3 FORECAST OF THE SKILL

An a priori estimation of the skill of the forecast can considerably enhance the usefulness of the forecast. This is particularly true in the medium range, where quality of the forecast varies substantially from day to day and from region to region. For this reason, we support the concept that a forecast should be accompanied by an estimation of its skill.

Different measures of "skill" are in use, some of which would not pass a formal definition of skill (following A H Murphy's work). While a word of caution about measures of skill and trivial components to skill forecasts remains in effect, we do not believe that measurement of skill is a major hang-up in improving forecasts of a priori estimates of skill.

Since the Predictability Workshop of 1988, considerable experience has been gathered in this area and new potential methods of prediction have been suggested. Possible predictors of skill include:

i) Forecast agreement or forecast spread among an ensemble of forecasts. This is one of the most effective predictors, but requires a careful selection of ensemble members to cover realistically the range of possible realizations. Two examples of such ensembles that have been shown to be more successful than classical Monte Carlo ensembles are:

- Selection of members which have fastest growing perturbations;
- Use of ensembles of forecasts from different operational centres.

ii) Regime indices which are associated with atmospheric instability properties. Initial optimistic results, using for example a PNA index, have not been found to be consistent throughout the years.

iii) Amplitude of the forecast anomaly.

iv) Measures of atmospheric persistence and stability such as forecast persistence, the time tendency of the rotational kinetic energy or the local Lyapunov exponents.

v) Short-term forecast errors.

A different approach designed to predict short-term local errors using the adjoint of the linear tangent model has been tested at KNMI and the French Weather Service.

Quasi-operational regression schemes based on some of these predictors have been run at ECMWF (*Palmer and Molteni*, 1991), SMHI (*Akesson*, this volume) and NMC (*McCalla and Kalnay*, 1988; *Kalnay and Toth*, this volume). The current level of correlation between predictors and observed regional skill is mediocre, between 0.2 and 0.5 for RMS and slightly higher for ACC, occasionally reaching 0.7.

It may not be possible to reach a level much higher than 0.7 ever for perfect models (*Barker*, 1991; *Mureau et al.*, this volume). This is because forecast spread and forecast errors are measures of the same type of variable, but the forecast error consists of only a single realization and therefore is a much noisier estimate than the ensemble spread, based on N realizations.

2.4 TECHNIQUES FOR CREATING FORECAST ENSEMBLES

A simple way of creating the initial perturbations for ensemble forecasting is the classic MC approach: to add random noise to the initial state. Such perturbations do not project well on the relevant meteorological scales and would be mostly removed in NWP models by the initialization, thus making the effective amplitude of the perturbation uncertain. Several additional techniques listed in the working group reports of the Workshop on Predictability in 1988 are:

1) random combination of analyses

2) analyses from other weather centres

3) analyses by different observing system scenarios

4) use of estimated analysis errors in the distribution of perturbations

5) unstable modes of the atmosphere

6) time lagging

7) perturbation in the physics

Two of these techniques have been applied quasi-operationally with varying success like the time lagging (at NMC, ECMWF and Met Office) and the analyses from other weather centres (at NMC). The use of unstable modes of the atmosphere is currently being tested at ECMWF and have proved to be successful in specifying the sensitive areas and structures of the atmosphere. This technique is described (*Mureau, Molteni, Palmer*, this volume) in these Proceedings.

Four more techniques have been recently tested:

8) Recent first guess errors: for the ensemble forecast from today's analysis (a linear combination of) scaled 6 hour forecast errors of the days immediately preceding the analysis are taken. Such perturbations are short range forecast errors and are therefore typical for errors and are to a good approximation typical for the errors of today's flow regime. These perturbations were used at ECMWF as a benchmark for the optimal mode perturbations.

9) Scaled lagged analysis forecasts: in this NEWLAF technique, perturbations for today's ensemble forecast are generated from scaled down errors from a set of 6, 12, 18 hour etc. short-range forecasts verifying on today's analysis. This will give 8 perturbations if we use the short-range forecast from the last 2 days. These perturbations can be added to or subtracted from the initial state of the ensemble forecast, generating 16 ensemble members. This technique has been tested at NMC with good results and is described in more detail in these Proceedings (*Kalnay and Toth*, this volume).

10) Spatially coherent noise: in the simple MC approach random noise is added to the IC. Several researchers have proposed to add "noise" that has the right spatial structures (*Schubert and Suarez*, 1990). For instance, the noise can be made red so as to mimic the climatological variance spectrum. Also mixed coherence (velocity-temperature) and vertical coherence can be taken into account.

11) Selection of MC members based on a one time step tendency: a large set of perturbations (about 1000) is generated from a scaled random linear combination of archived analyses. The model is then integrated for one time step after which only those perturbations are retained which are growing fast in terms of rotational kinetic energy. Preliminary tests at NMC indicate that the method may efficiently select the perturbations with the fastest one day error growth.

2.5 POST-PROCESSING AND PRESENTING ENSEMBLE FORECAST RESULTS

To both the operational forecaster and the end users, the provision of ensemble and probability forecasts will provide information of a type not previously encountered. As with any new technique, thorough assessment of all aspects of ensemble predictions will be required prior to general acceptance of ensembles within the meteorological communities. In addition new forms of presentation will be required to provide users with ready access to essential information available within ensembles. The following lists are offered in order to identify aspects of ensemble predictions which will require verification, and to suggest displays that need to be presented to the developer and eventually the forecaster.

2.5.1 Fields

(In order to present a condensed view of the outcome of the ensemble)

a. ensemble mean fields (the induced smoothing will provide forecast fields with unpredictable scales filtered out.)
   - geopotentials/MSL pressure
   - temperature/thickness
   - upper air winds

b. fields of variances
   - geopotentials/MSL pressure
   - temperature/thickness

c. clustering of ensemble solutions into a few distinct regimes

d. identification of extreme events (e.g. blockings in the mid latitudes and hurricanes in the tropics).

e. fields of higher moments, if warranted by the number of ensemble members (e.g. fields of the skewness of the pressure distribution).

2.5.2 Frequency distributions

(In order to verify the realism and skill of the ensemble forecast system, explore the structure of the ensemble and provide a probabilistic display for the end user.)

a. free atmospheric winds, temperature and geopotentials

b. humidity, clouds and precipitation

c. probability charts of relevant parameters

To facilitate the interpretation of the distribution of solutions of the ensemble we recommend that new displays be tried, e.g. interactive processing on workstations.

With the ensemble forecasting we will be looking at model output here in ways never seen before. This will force us to identify - "systematic errors" such as impossibly high or low variance, assigning non-zero probability to weather events never observed etc. Strategies may have to be developed to correct such problems. Also, feedback onto model development may results from these new diagnostics.

2.6 RECOMMENDATIONS

The recommendations of the previous Workshop (ECMWF, 1988) are still remarkably apt. As noted above, we see the evolution of two broad "forecast methods" to replace the single deterministic forecast currently produced, namely a deterministic forecast with an accompanying measure of skill or a direct "probabilistic" forecast. The production of such forecasts depends separately and jointly on methods of predicting forecast skill and on the use of ensembles.

- Continued and vigorous research into methods of predicting forecast skill must be pursued. As noted in the previous Workshop report, there are many methods which must be investigated. Although considerable progress has been made, at this time there is no clearly best approach.

- Research into the production of optimum ensembles of forecasts is vital. The basic problem is to properly sample the space of initial conditions so that the resulting forecast ensemble correctly reflects the probability distribution to be associated with the forecast. Random MC perturbations neither reflect the probability distribution of the observation errors nor are they likely to sufficiently sample the initial condition space, while the selection of the "fastest growing" perturbations may result in a distribution of forecasts which could overestimate the forecast error.

- More advanced analysis systems should provide error covariance information that should be used to improve the realism of the ensemble perturbations.

- Since the type of forecast information provided to the users will radically change, it is important to investigate methods of presenting this new form of information. In particular "probabilistic" forecasts could comprise, for instance, an ensemble mean together with the standard deviation of the forecasts, a set of outcomes which account for specified amounts of the variance or much more detailed information consisting of frequency distributions and other statistics. The useful production and interpretation of such information is an area which requires considerable investigation.

- Ensemble forecasting will require a major research effort on the best methods to translate the results into guidance that provides a succinct summary of the probabilistic information to the users - this may be in the form of ensemble average fields, fields of standard deviations, clusters of probable realizations, etc. It will be necessary not only to develop these techniques at ECMWF (and other centres), but also to devote considerable effort to verify their usefulness. In this area ECMWF and interested users should work closely and in a complementary fashion. There probably should be a trial phase when the development and verification should be done centrally.

- One of the strengths of meteorological research is in its cooperative efforts to investigate important issues. Coordinated experimentation should be encouraged by choosing a suitable common set of forecast dates for international experimentation, by involving the WGNE and other interested bodies in the promotion of this coordinated efforts and in the sponsoring of workshops and meetings to bring together active researchers in the field. In particular, no technique to create ensemble members or to forecast the skill has yet been shown to be clearly superior. Therefore, coordinated comparisons by different centres testing methods for the same areas and the same periods should be performed in order to find the most promising techniques or combination of techniques.

- A proper evaluation of the performance of ensemble forecasts requires a detailed knowledge of the systematic errors of the forecast model. These systematic errors consist not only of a shift of the time mean state but also of different distributions of variances and modifications of the teleconnection pattern and regime frequencies. All these errors can seriously affect the reliability of probabilistic forecasts observed from ensembles and serious effort should be made to evaluate them in a proper way.

- Two main options for future investment of computer power are possible, namely: (1) a single deterministic model with very high resolution (and a suite of methods to estimate its skill, which can include running ensembles of less expensive models); 2) ensembles of somewhat less expensive models which have equally probable members. Whereas so far only option (1) has been followed, option (2) also deserves serious consideration.

3. WORKING GROUP 3: STRATEGIES FOR EXTENDED-RANGE PREDICTION

3.1 INTRODUCTION

Despite their potential practical importance, dynamical extended-range predictions this far have not had a major impact on operational practice, even in advanced NWP centres. Prediction of atmospheric behaviour beyond the deterministic predictability limit for instantaneous fields is still very much an open research area. Given the importance of distinct atmospheric weather regimes within the same season a major problem is how to assess their predictability (both theoretical and practical) on time scales of the same order of regimes' lifespan. The ability of a model to represent regimes and the transitions from one to another is essential, not only in a statistical sense (climatology of the model vs real climatology, i.e. systematic errors), but also as an initial value problem. Our usual perspective of model behaviour and its seasonal dependence must be changed to account for the existence of regions of the atmosphere phase-space (more than one, possibly several) where residence is more probable, be it because they are more persistent or because they are visited more often, or both. Model skill appears largely dependent on the regime prevailing during the forecast and regime-to-regime transitions become the main stumbling block for forecasting models.

Such a profoundly modified perspective requires new strategies and new approaches, and both the theoretical and the empirical avenues should be pursued. It is perhaps of significance that even at this very workshop on Predictability, little was said along the line of theoretical predictability studies and how our concepts in this field should be modified by the existence of distinct weather regimes. Most studies concentrated on a more empirical approach, of the type "try and see". This could very well reflect the objective difficulties to be faced by the development of scientific understanding at this particular point in time and is probably suggestive of the need for a strong concerted effort of the atmospheric scientific community in this particular area.

3.2 BOUNDARY VERSUS INTERNAL AND INITIAL CONDITIONS VARIABILITY SOURCES

In the extended range, forecasts are more successful when they have to predict situations in which the initial large-scale pattern persists, which probably also reflects the fact that the large-scale variability of models is often weaker than observed. This implies that some sources of variability are either missing or misrepresented in the model. It is important to identify these sources.

The first source is the variability of the boundary conditions (i.e. essentially the sea surface temperature, but also the soil moisture and snow cover which are prognostic variables strongly influenced by the initial conditions). On the month-to-season time scale, the part of the atmospheric variability explained by the boundary forcings might be comparatively smaller, but since it is perhaps the only part we can predict beyond day 20-30, it is significant. However, experiments based on the ENSO phenomenon show that the use of climatological SSTs may be detrimental even in the extratropics. It is not yet well established if the day-to-day fluctuations of SST may improve the extra-tropical skill compared to simple persistence of SST anomalies: this day-to-day fluctuation would require a very accurate calculation of the surface fluxes in the model. If a model is improved by the use of varying SSTs, there is the problem of atmosphere-ocean coupling. But as a first step, it would be useful to measure the impact of SSTs and other lower boundary forcings on atmospheric variance. It would be useful also to determine the time-range beyond which, on average, a forecast with persistent SST anomalies is outperformed by a forecast with actual SST anomalies. Many of these quantities have been already estimated with GCM's, but we require up-to-date estimates with the most recent models, since they may depend on model formulation, e.g. horizontal resolution.

Even with the same boundary conditions, synoptic weather systems of integrations which slightly differ in the initial conditions become decorrelated after about 20 days. This is reflected in the forecast spread in Monte Carlo-type simulations. However, in such ensemble integrations, limiting the perturbation to the initial situation relies on the hypothesis that the model describes perfectly the atmospheric evolution, or at least that the main source of the error is the amplification of forecast error. This hypothesis is true for the short range, questionable for the medium range, and false for the extended range (with present models): the spread of individual integrations and the skill of their average have little correlation. Moreover, the verification of probability forecasts suggests that the spread underestimates the uncertainty on the forecast.

There is a need for theoretical and numerical studies on the forecast uncertainty in a range of models. One way could consist of performing Monte Carlo forecasts by varying an empirical (tunable) model parameter (e.g. diffusion coefficient) or by using different parametrization schemes or different models from different forecasting centres. Another more difficult, but more promising way could consist of introducing into the prognostic equations stochastic terms representing internal model errors. To do so one first needs to assess how this error source would contribute to the total error budget (see *Boer* in these Proceedings).

3.3 DEVELOPMENT AND APPLICATION OF DIAGNOSTIC TOOLS FOR PREDICTABILITY STUDIES

Two major issues in this area require urgent attention. Firstly, there is a need to specify agreed upon definitions for identifying regimes and/or circulation patterns; blocking (and zonal situations) are but one (regional) way to look at this problem and more effort should be devoted in trying to reconcile objective ("abstract") regime definitions (e.g. clusters in some suitably chosen low-order phase-space) with more synoptically-based concepts, like blocking and cut-off-lows. Secondly, we require improved physical understanding of the phenomena in question in order to suggest ways of improving models.

Note that even if models reproduce regimes, this does not necessarily mean they do so for the right physical reasons. Therefore, detailed studies of the underlying dynamical processes in real vs model atmospheres should be performed. One commonly used tool is the use of isentropic potential vorticity formulation (IPV). Diagnostics in terms of IPV allows, for instance, the distinction between diabatic forcing and advection. For the validation of GCMs, the statistical approach, as well as case studies are needed. A key question is whether GCM's reproduce correctly the interactions between mean flow, large-scale low-frequency patterns and baroclinic transients. Another question is the possible importance of tropical/extratropical interactions.

3.4 MAJOR PROBLEM AREAS IN ERF

3.4.1 Extracting the predictable component of the signal

Assuming that a degree of predictability exists for some atmospheric variables beyond the instantaneous deterministic predictability limit, the problem remains to extract this predictable component of the forecast signal from the background noise of the unpredictable structures. The zero order approach is, for example, to filter the forecast with a time-mean operator, which also acts as a spatial filter. There is, however, no guarantee that this is the only (or the best) way to accomplish the task. In other words, present day signal-extracting procedures need to be extended and further methods need to be developed to identify and possibly separate out those subspaces of the atmospheric attractor where the projection of the atmospheric trajectory is (more) predictable, at the extended range, then the complete trajectory.

To accomplish this, it may not be sufficient to consider the static components of the fields to be predicted (e.g. the leading EOFs), but it might turn out to be necessary to incorporate some properties describing life cycles, oscillation patterns and some effects of transients.

Particularly necessary to this identification work are observational studies, as well as theoretical developments. Essential to the former is an adequate database, and current efforts to secure support for reanalysis projects (e.g. ECMWF and NMC reanalysis plans) should be further pursued.

Extended range forecasting is a problem strongly affected by the concept of limit of deterministic predictability. Thus estimates of error growth rates (for example in terms of the spread of ensemble forecasts, of the spectrum of Lyapunov exponents, etc) may not suffice. Scale dependent error budgets and/or measures of the rate at which errors fill the more predictable part of the phase space should be studied, both theoretically and by means of numerical experimentation.

3.4.2 Regime transitions

The prediction of regime transitions is possibly the major unsolved problem in extended-range and medium-range forecasting. Predicting the onset and demise of blocking episodes is a typical example. Others include forecasting major shifts in phase and amplitude of the planetary scale waves and the development and demise of cut-off lows. There are several difficulties related to this general problem. One is the exact definition of regime(s), as already discussed above. Both maintenance of regimes and regime transitions are thought to be phenomena to describe which scale-interaction processes might be important, if not essential. Notable examples are the effects of transient synoptic eddies on blocking and the role of stochastic baroclinic forcing on cluster-to-cluster transitions. An area of controversy is the importance of model resolution in simulating low-frequency variability (LFV) when faster, smaller-scale transient systems (and therefore sensitive to model resolution) are important components of the LFV generating processes.

The starting regime should actually be split into two sets of states: states that led to the transition, and states that do not. Those two sets of states are difficult to distinguish (which is why we classified them as being in the same regime). This situation, where spatially similar states have differing evolutions, leads to error growth. Simplified models, such as tangent models, can and should be analyzed to determine the structures distinguishing these two sets of states. Filtered tangent models, which focus on large-scale, slowly evolving phenomena, are likely to be the most useful. In particular, tangent models may allow one to explicitly compute covariance evolution, at least in the short range, thus permitting calculation of the rate and mode of separation between the two sets of states.

Several questions also arose about the poorly represented transitions between regimes. There are theoretical and practical questions. For instance, are these transitions only due to random occurrence of particular small-scale transients, or are there some variables evolving smoothly able to favour these transitions? More practical questions could be: how do we recognize a transition? The former question, again, can be formulated in terms of IPV and scale interactions. The latter question relies mostly on the definition of indices for weather regimes.

Of concern is the role of scale interaction processes and tropical/extratropical interactions in regime transitions. The former, as mentioned above, has implications with regard to resolution and, thereby, to the maximum number of ensemble members practically possible with given computer resources. The latter is important with regard to tropical predictability and, therefore, to properly specifying SSTs and the physics of communicating the effects to the atmosphere (see next sub-section).

3.4.3 Tropical-Extratropical Interactions

It is well known that the impact of the tropics-extratropics interaction is at times very large already beyond the short range. The relaxation technique, by which the flow in a given area is relaxed towards the analysis, has been successfully used on establishing the importance of this mechanism. In fact, when the extratropical fields were relaxed towards observed data in extended range experiments, (*Ferranti et al.*, 1990) noticeable improvement in skill scores for the forecast of tropical flow, in particular for its rotational component has been found. A similar improvement in extratropical skill scores has been shown by relaxing tropical fields to the analysis in period of a strong 30-60 day oscillation. This technique could be helpful to study the role of other sources of internal variability in maintaining the low-frequency component of the atmospheric flow.

3.4.4 Interannual variability of intraseasonal blocking frequency

The performance of GCM's in predicting blocking events is normally judged as being satisfactory if models are able to generate and maintain anomalous circulations in a relatively simple statistical sense, i.e. in terms of seasonal distributions of blocking frequency as functions of longitude. With the advent of complex and more ambitious coupled ocean-atmosphere models, and with the increasing use of observed SSTs in extended range and multi-year integrations, blocking verification techniques should be improved to take into account the initial value problem aspect of blocking predictions. It is anticipated that the correct interannual variability of seasonal blocking frequency may become not only an essential requirement for extended-range and seasonal forecasting, but also an important factor is assessing climatology (and climate drift) of GCM's used for a variety of purposes, up to the climatic time-scale.

3.5 COORDINATION OF RESEARCH: NEED FOR COMPLEMENTARITY AND FOR INTERCOMPARISONS

To comprehensively assess the limits of predictive skill in treatment of regime transitions it is recommended that extended range forecast experiments be conducted on a fairly large set of cases (about 20-30) selected on the basis of regime characteristics. A set of cases, in fact, has been suggested in response to a recommendation of the Workshop on Numerical Long-Range Prediction in June 1990 in Boulder (*Baumhefner et al.*, BAMS, November 1991). The idea set forth at that workshop was for as many groups as possible to evaluate the capabilities and limitations of prediction models, as well as strategies of ensemble forecasting, on common cases. The expectation is that intercomparison of results and analysis for the reasons for similarities and differences will contribute considerably to understanding predictability as it relates to regime.

Therefore, there is a clear need of coordinated work between the different teams working on these problems. Exchange of data and formulations has to be arranged and made easier. The application of different diagnostic methods on common datasets, as well as the application of individual methods on multiple datasets should be encouraged.

3.6 RECOMMENDATIONS

- Extended Range Forecasts experimental work should be continued.

- Determine the time-range beyond which, on average, a forecast with persistent SST anomalies is outperformed by a forecast with actual SST anomalies. Up-to-date estimates are required with the most recent models.

- Specify agreed upon definitions for identifying regimes and/or circulation patterns. More effort should be devoted to trying to reconcile objective regime definitions with more synoptically-based concepts.

- Scale dependent error budgets and/or measures of the rate at which errors fill the more predictable part of the phase space should be studied, both theoretically and by means of numerical experimentation.

- Simplified models, such as tangent models, should be analyzed to determine the instability structures distinguishing possible states in the phase space.

- Blocking verification techniques should be improved to take into account the initial value problem aspect of blocking predictions.

- Experiments should be conducted on a fairly large set of cases (about 20-30) selected on the basis of regime characteristics.

- Exchange of initial conditions among research centres for these cases should be encouraged.