Predicting uncertainty in forecasts of weather and climate

(Also published as ECMWF Technical Memorandum No. 294)

By **T.N. Palmer**

Research Department

November 1999

Abstract

The predictability of weather and climate forecasts is determined by the projection of uncertainties in both initial conditions and model formulation onto flow-dependent instabilities of the chaotic climate attractor. Since it is essential to be able to estimate the impact of such uncertainties on forecast accuracy, no weather or climate prediction can be considered complete without a forecast of the associated flow-dependent predictability. The problem of predicting uncertainty can be posed in terms of the Liouville equation for the growth of initial uncertainty, or a form of Fokker-Planck equation if model uncertainties are also taken into account. However, in practice, the problem is approached using ensembles of integrations of comprehensive weather and climate prediction models, with explicit perturbations to both initial conditions and model formulation; the resulting ensemble of forecasts can be interpreted as a probabilistic prediction.

Many of the difficulties in forecasting predictability arise from the large dimensionality of the climate system, and special techniques to generate ensemble perturbations have been developed. Special emphasis is placed on the use of singular-vector methods to determine the linearly unstable component of the initial probability density function. Methods to sample uncertainties in model formulation are also described. Practical ensemble prediction systems for prediction on timescales of days (weather forecasts), seasons (including predictions of El Niño) and decades (including climate change projections) are described, and examples of resulting probabilistic forecast products shown. Methods to evaluate the skill of these probabilistic forecasts are outlined. By using ensemble forecasts as input to a simple decision-model analysis, it is shown that probability forecasts of weather and climate have greater potential economic value than corresponding single deterministic forecasts with uncertain accuracy.

Table of contents

1. INTRODUCTION

1.1 Overview

A desirable if not necessary characteristic of any physical model is an ability to make falsifiable predictions. Such predictions are the life blood of meteorology and climate science. Predictions from vast computer models of the atmosphere, integrating the Navier-Stokes equations for a three dimensional multi-constituent multi-phase rotating fluid, and coupled to a representation of the land surface, are continually put to the test through the daily weather forecast (e.g. Bengtsson, 1999; see also http://www.ecmwf.int). On seasonal to interannual timescales, these same models, with 2-way coupling to similar mathematical representations of the global oceans, predict the development of phenomena such as El Niño, with consequences for seasonal rainfall and temperature patterns around much of the globe (e.g. *Stockdale et al*., 1998; see also http://www.iges.org/ellfb). Coupled ocean-atmosphere models are also widely used to make predictions of possible changes in climate over the next century as a result of anthropogenic influence on the composition of the atmosphere (e.g. *IPCC*, 1996).

However, there is little sense in making predictions without having some prior sense of the accuracy of those predictions (*Tennekes*, 1991); quantification of error is a basic tenet in experimental physics. Earth's climate is a prototypical chaotic system (*Lorenz*, 1993), implying that its evolution is sensitive to the specification of the initial state; however, an appreciation of the importance of quantifying the role that initial error plays in limiting the accuracy of weather predictions pre-dates the development of chaotic models (*Thompson*, 1957).

How would one go about making an a priori assessment of the accuracy of a weather forecast, or a prediction of El Niño? Of course, a 'climatological-mean' error can be derived by verifying a past set of predictions, and averaging the resulting forecast errors. However, such a crude estimate may not be particularly useful. Chaotic dynamics implies not only that a forecast is sensitive to initial error, but also that the rate of growth of initial error is itself a function of the initial state (see Section 2). Weather forecasters have a practical sense of this dependence of error growth on initial state; certain types of atmospheric flow are known to be rather stable and hence predictable, others to be unstable and unpredictable. As such, a key to predicting forecast uncertainty lies in the estimation of the effects of local instabilities in regions of phase space through which a forecast trajectory is likely to pass.

In addition to error in initial conditions, the accuracy of weather and climate forecasts are influenced by our ability to represent computationally the full equations of that govern climate. For example, there will be inevitable errors in representing circulations on scales comparable with or smaller than a model's truncation scale. These errors can

propagate upscale and influence weather and climate phenomena with characteristic size much larger than the truncation scale. Uncertainty in model formulation is certainly one of the most important factors which undermine confidence in climate forecasts - representation of cloud systems (in models which cannot resolve individual clouds) being a particular manifestation of this problem. As with initial error, uncertainties in model formulation impact on climatic circulation patterns through the projection of these uncertainties onto flow-dependent dynamical instabilities of the climate system.

In the body of this paper, results are shown from a number of numerical models of the climate system. It is useful to consider three types of model, distinguished by their degree of complexity. The first could be thought of as 'toy' models; they are used primarily to illustrate particular paradigms. Examples are the *Lorenz* (1963) model and the delayed oscillator model (see Sections 2 and 7). The second type of model could be described as 'intermediate'; it certainly has prognostic value, but is based on simplified equations of motion where terms which are second order in some small parameter are ignored. For many examples discussed in this paper, the so-called Rossby number $Ro = U/fL$ (e.g. *Gill*, 1982) is such a parameter. Here, U, L denote a typical horizontal velocity and length scale associated with a particular climatic or weather phenomenon, $f = 2\Omega\sin\phi$ is the Coriolis parameter, where Ω is the angular speed of the Earth and ϕ denotes latitude. Examples of intermediate models, are the atmospheric quasigeostrophic model (e.g. *Marshall and Molteni*, 1993: see Sections 3 and 7), and a simplified coupled ocean-atmosphere model of El Niño (e.g. *Zebiak and Cane*, 1987: see Sections 5 and 7). Intermediate models are generally truncated to have O(10^3) or less degrees of freedom, which makes numerical integration and stability analysis extremely tractable by modern computing standards.

The final type of model in the hierarchy of complexity are the comprehensive global climate and weather prediction models; these typically have O(10^6-10^7) degrees of freedom. At national (and international) meteorological and climate centres, quantitative weather and climate predictions are now almost universally based on output from these types of model. The models are formulated using finite (Galerkin) truncations of fluid-dynamic partial differential equations where (at most) only the hydrostatic assumption is applied to filter meteorologically-unimportant modes. A possible (and easily visualised) representation is in terms of grid points in physical space; a typical resolution would be about 100 km in the horizontal and 1km in the vertical (somewhat finer for weather prediction models, somewhat coarser for climate prediction models with longer integration times). These equations describe the local evolution of mass, energy, momentum and composition, with suitable source and sink terms. The most important atmospheric composition variable is water, represented in each of its different phases. Details of these equations can be found in many references (e.g. *Trenberth*, 1992). Such comprehensive models are integrated on supercomputers, with (at the time of writing) typical sustained speeds of O(10^{11}) floating point operations per second. In practice, the difference between the atmosphere component of weather and climate prediction models is not great - and in some instances there is no difference; however, weather prediction models do not generally have an interactive ocean, whilst climate models do. An example of this third type of comprehensive model, discussed below, is the European Centre for Medium-Range Weather Forecasts (ECMWF) weather and climate prediction model (*Bengtsson*, 1999).

In this paper, we consider two types of prediction. Following Lorenz (1975), we refer to initial value problems as 'predictions of the first kind'. By contrast, forecasts which are not dependent on initial conditions, for example predicting changes in the statistics of climate as a result of some prescribed imposed perturbation, would constitute a 'prediction of the second kind'. A weather forecast is clearly a prediction of the first kind; so is a forecast of El Niño, referred to as a climate prediction of the first kind. By contrast, estimating the effects on climate of a prescribed volcanic emission, prescribed variations in Earth's orbit (thought to cause ice ages) or prescribed anthropogenic changes in atmospheric composition, would constitute a climate prediction of the second kind.

1.2 Scope

This paper deals with the problem of forecasting uncertainty in weather and climate prediction from its theoretical basis, through an outline of practical methodologies, to an analysis of validation techniques including estimates of potential economic value. The author hopes that the mathematical description of these components will be of some help to readers wishing to gain some introduction to the quantitative methods used in the subject. However, at the least, the reader will be able to deduce that the topic of weather and climate prediction is quantitative and objective. (The days are over, of hanging out the seaweed, examining the size of molehills, or studying animal entrails for portents of coming tempests - that is, unless the computers are down!) On the other hand, readers not interested in the details of the mathematics should be able to appreciate many of the results given without dwelling on the equations at any length.

In Section 2, we consider how to forecast uncertainty in a prediction of the first kind, assuming a perfect deterministic forecast model. The evolution equation for the probability density function (pdf) of the climate state vector is the Liouville equation; an example of its solution is given for illustration. However, application to the real climate system is severely hampered by two fundamental problems. The first is directly associated with the dimensionality of the climate equations; as mentioned above, current numerical weather prediction models comprise O(10^7) individual scalar variables. The second problem (not unrelated to the first) is that, in practice, the initial pdf is not itself well known.

To amplify on this last remark, a description of current (variational) meteorological data assimilation schemes is described in Section 3. These schemes are used to determine initial conditions for weather and climate forecasts, given a set of atmospheric and oceanic observations whose density is heterogeneous in both space and time. Such data assimilation schemes are based on minimising a cost function which combines these observations with a background estimate of the initial state provided by a short-range model forecast from an earlier set of initial conditions. In principle, given Gaussian error statistics, the Hessian or second derivative of the cost function determines the initial pdf. In practice, there are significant shortcomings in our ability to estimate this pdf.

The number of degrees of freedom in comprehensive climate and weather prediction models is not determined by any scientific constraint (there is no obvious 'gap' in the energy spectrum of atmospheric motions), but rather by the degree of complexity than can be accommodated using current computer technology. As such, there are inevitably processes occurring in the atmosphere and oceans which are partially resolved or unresolved and must be represented by some parametrised closure approximation. Examples are associated with cloud formation and dissipation, and momentum transfer to the solid earth by topography. However, there is a fundamental indeterminacy in the formulation of these parametrisations since there is no meaningful scale separation between resolved and unresolved scales in the climate system. Section 4 describes two recent attempts to represent the pdf associated with this uncertainty in the computational representation of the equations of motion of climate: the multi-model ensemble, and stochastic parametrisation.

A theoretical framework for describing error growth is developed in Section 5. Two common measures of perturbation amplification used in different branches of physics and mathematics are normal mode growth and Lyapunov exponent growth. Neither is well suited to describing error growth in the climate system. Firstly, because of the advective nonlinearity in the governing equations of motion, the linearised dynamical operators are not normal; as such, over finite times, perturbation growth need not be bounded by the fastest eigenmode growth. Also, dominant Lyapunov or eigenmode growth in a comprehensive multi-scale model may refer to fast instabilities (such as convective instabilities) whose spatial scales are much smaller than those describing weather or climate phenomena. To address these problems, we discuss in Section 5 a general formulation of perturbation growth in the linearised approximation, in terms of a singular value decomposition of the linearised dynamics (building on the developments in Section 3). Examples of singular vectors for weather and climate prediction problems are shown, and their fundamental non-modality is discussed. Because of the nonlinearity of the underlying dynamics, the appropriate

singular values vary on the attractor; this variation describes why forecast error can fluctuate for fixed initial error. The variation of singular values on the attractor is also relevant for understanding the amplification of model error by flow dependent instabilities. The relationship between singular values, eigevalues and Lyapunov exponents is discussed.

Section 6 discusses some applications of the singular vector analysis. In one application ('chaotic control of the observing system') singular vectors are used to determine locations where additional 'targeted' observations might significantly improve a forecast's initial state.

Section 7 describes the basis behind attempts to predict uncertainty in daily, seasonal and climate change forecasts using ensembles of atmosphere or coupled ocean-atmosphere model integrations. In practice such ensembles are interpreted in probabilistic form. If the ensemble of forecast phase-space trajectories evolve though a relatively stable part of the climate attractor, then resulting probability forecasts will be relatively sharp. Conversely, if the ensemble passes through a particularly unstable part of the attractor, then the corresponding forecast probability may be little different from a long-term climatological frequency.

The question of how to validate probability forecasts is discussed in Section 8. Two particular techniques are described. The first is based on a root mean square distance between the probability forecast of a dichotomous event and the corresponding verification. This measure allows one to formulate the notion of reliability of probability forecasts. The second quantity measures the so-called hit and false alarm rate of the forecast of a dichotomous event, assuming that the event is forecast if the predicted probability exceeds some prescribed probability threshold.

A fundamental question when assessing probability forecasts is whether a useful level of skill has been attained. Obviously, different users have different criteria for judging usefulness. For some, probability forecasts might be deemed useless unless they are sharp and quasi-deterministic. For others, who might be looking to accrue benefit over a long time, forecast probabilities which are only marginally different from climatological frequencies, may be useful. To assess this issue more quantitatively, a simple cost/loss decision model is applied in Section 9 based on the hit and false alarm rates discussed in Section 8. It is shown, that the (potential) economic value of probability weather forecasts for a variety of users, is higher than the corresponding value from single, deterministic forecasts.

Concluding remarks are made in Section 10.

2. The Liouville equation

The evolution equations in a climate or weather prediction model are conventionally treated as deterministic. These (N dimensional) equations, based on spatially-truncated momentum, energy, mass and composition conservation equations will be written schematically as

$$\dot{X} = F[X] \tag{1}$$

where X describes an instantaneous state of the climate system in N-dimensional phase space. Eq. (1) is fundamentally nonlinear and deterministic in the sense that, for any initial state X_a, the equation determines a unique forecast state X_f. (As described in Section 3 below, information from meteorological observations are combined with a prior background state through a process called data analysis and assimilation. In meteorology, the initial state is often referred to as the initial ‘analysis’ - hence the subscript ‘a’.)

The meteorological and oceanic observing network is sparse over many parts of the world, and the observations themselves are obviously subject to measurement error. The resulting uncertainty in the initial state can be represented by the pdf $\rho(X, t_a)$; given a volume V of phase space, then $\int_V \rho(X, t_a) dV$ is the probability that the true

initial state X_{true} at time t_a lies in V. If V is bounded by an isopleth of ρ (i.e. co-moving in phase space), then, from the determinism of Eq. (1), the probability that X_{true} lies in V is time invariant. Hence, (similar to the mass continuity equation in physical space), the evolution of ρ is given by the Liouville conservation equation (introduced in a meteorological context by *Gleeson*, 1966, and *Epstein*, 1969)

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial X}(\dot{X}\rho) \equiv L\rho \tag{2}$$

where $\dot{X}$ is given by Eq. (1). In the second term of Eq. (2), there is an implied summation over all the components of X.

Fig. 1 illustrates schematically the evolution of an isopleth of $\rho(X, t_a)$. For simplicity we assume the initial pdf is isotropic (e.g. by applying a suitable coordinate transformation). In the early part of the forecast, the isopleth evolves in a way consistent with linearised dynamics; the *N*-ball at initial time has evolved to an *N*-ellipsoid at forecast time t_1. For weather scales of $0(10^3)$ km, this linear phase lasts for about 1-2 days into the forecast. Beyond this time, the isopleth starts to deform nonlinearly. The third schematic shows the isopleth at a forecast range in which errors are growing nonlinearly. Predictability is finally lost when the forecast pdf $\rho(X, t_a)$ has evolved irreversibly to the invariant distribution ?inv of the attractor. This is shown schematically in Fig. 1 using the *Lorenz* (1963) attractor - a 'toy-model' surrogate of the real climate attractor (*Palmer*, 1993a).

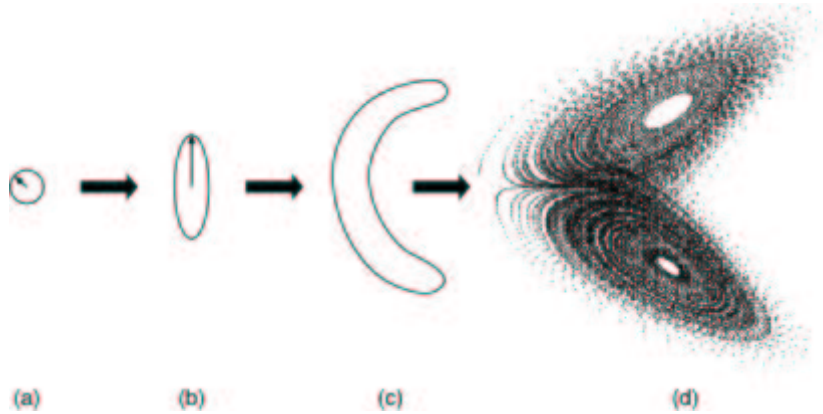

Figure 1: Schematic evolution of an isopleth of the probability density function (pdf) of initial and forecast error in N-dimensional phase space. (a) At initial time, (b) during the linearised stage of evolution. A (singular) vector pointing along the major axis of the pdf ellipsoid is shown in (b), and its pre-image at initial time is shown in (a). (c) The evolution of the isopleth during the nonlinear phase is shown in (c); there is still predictability, though the pdf is no longer Gaussian. (d) Total loss of predictability, occurring when the forecast pdf is indistinguishable from the attractor's invariant pdf.

As mentioned in the introduction, the growth of the pdf through the forecast range is a function of the initial state. This can be seen by considering a small perturbation δx to the initial state X_a. From Eq. (1), the evolution equation for δx is given by

$$\delta \dot{x} = J\delta x \tag{3}$$

where the Jacobian is defined as

$$J = dF/dX \tag{4}$$

Since $F[X]$ is at least quadratic in X, then J is at least linearly dependent on X. This dependency is illustrated in Fig. 2 showing the growth of an initial isopleth of an idealised pdf at three different positions on the *Lorenz* (1963) attractor. In the first position, there is little growth, and hence large local predictability. In the second position there is some growth as the pdf evolves towards the lower middle half of the attractor. In the third position, initial growth is large, and the resulting predictability is correspondingly small.

Figure 2: Phase-space evolution of an ensemble of initial points on the *Lorenz* (1963) attractor, for three different sets of initial conditions. Predictability is a function of initial state.

The nonlinear phase of pdf evolution can be much longer than the linear phase. For example, Smith et al. (1999) have studied the evolution of an initial pdf on the *Lorenz* (1963) attractor using a Monte Carlo process. The initial pdf was obtained by adding some notional prescribed 'observation' error to points on the attractor. The initial pdf is sharp, consistent with a small 'observation' error, and initially spreads out in a way consistent with linear theory. The pdf resharpens as it enters the region of phase space where small perturbations decay with time (cf Fig. 2), and then bifurcates, leading to a highly non-normal distribution. The existence of such bimodal behaviour indicates that it may not be sufficient to describe forecast uncertainty in terms of a simple ‘error bar'.

As shown in *Ehrendorfer* (1994a), the Liouville equation can be formally solved to give the value of ρ at a given point X in phase space at forecast time t. Specifically

$$\rho(X,t) = \rho(X',t_a)/\exp\left\{\int_{t_a}^{t} tr[J(t')dt']\right\} \tag{5}$$

where 'tr' denotes the trace operation. The point in this equation corresponds to that initial point, which, under the action of Eq. (1) evolves to the given point X at time t.

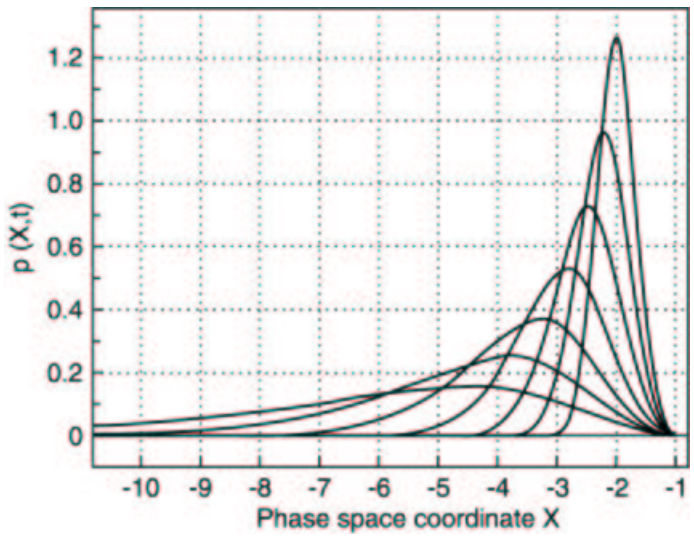

Figure 3: An analytical solution to the Liouville equation for an initial Gaussian pdf (shown peaked on the right-hand side of the figure) evolved using the Riccati equation (see text). From *Ehrendorfer* (1994a).

Using the identity $\det \exp A = \mathrm{tr} \exp A$, then Eq. (5), can be written as

$$\rho(X,t) = \rho(X',t_a)/\det M(t,t_a) \tag{6}$$

where

$$M(t,t_a) = \exp\int_{t_a}^{t} J(t')dt' \tag{7}$$

is the so-called forward tangent propagator, mapping a perturbation $\delta x(t_a)$, along the nonlinear trajectory from X to X' to

$$\delta x(t) - M(t,t_a)\delta x(t_a) \tag{8}$$

A simple example which illustrates this solution to the Liouville equation is given in Fig. 3 , for a 1 dimensional Riccati equation (*Ehrendorfer*, 1994a)

$$\dot{X} - aX^2 + bX + c \tag{9}$$

where $b^2 > 4ac$, based on an initial Gaussian pdf. The pdf evolves away from the unstable equilibrium point at $X = -1$ and therefore reflects the dynamical properties of Eq. (9). Within the integration period, this pdf has evolved to the nonlinear phase.

The forward tangent propagator plays an important role in meteorological data assimilation systems; see Section 3 below. However, even though the forward tangent propagator may exist as a piece of computer code, this does not mean that the Liouville equation can be readily solved for the weather prediction problem. Firstly, the determinant of the forward tangent propagator is determined by the product of all its singular values (see Section 5). For a comprehensive weather prediction model, a determination of the full set of O(10^7) singular values is currently impossible. Secondly, the inversion of Eq. (1) to find an initial state X', given a forecast state X, is itself problematic. Even on timescales of a day or so, decaying phase-space directions (as determined by the existence of small singular values of the propagator, see Section 5) will lead to the inversion being poorly conditioned (*Reynolds and Palmer*, 1998). Thirdly, a particular type of weather at a particular location is not related 1-1 with a state X of the climate system. For example, to estimate the probability of it raining in London two days from now, we would have to apply Eq. (6) and the inversion to find X', to each state on the climate attractor, for which it is raining in London.

An alternative to using the solution form (6) is to integrate the partial differential equation (2) by randomly sampling the initial pdf, and integrating each sampled point using (1); the Monte-Carlo solution. However, the problem of dimensionality continues to be a significant issue. If phasespace is N dimensional, then, even in the linear phase, O(N^2) integrations will be needed to determine the forecast error covariance matrix. In the nonlinear phase, many more integrations are needed to determine the pdf, as it begins to wrap itself around the attractor. Ehrendorfer (1994b) has shown that even for a 3-dimensional dynamical system, a Monte-Carlo sampling of O(10^2) points can be insufficient to determine the pdf within the nonlinear range.

Yet another method of solution of the Liouville equation is possible, writing Eq. (2) in terms of an infinite hierarchy of equations for the moments of ρ, and applying some closure to this set of moments (*Epstein*, 1969). This method is certainly useful for evolving the pdf within the linearised phase, and indeed forms the basis of the so-called Kalman filter approaches to data assimilation (see Section 3 below). *Ehrendorfer* (1994b) has shown that in the nonlinear phase, substantial errors in estimating the first and second moments of ρ can arise from neglecting third and higher order moments. A more sophisticated approach to closure is to use arguments from turbulence theory (see Section 5) to seek scaling relations between moments (*Frisch*, 1995). *Nicolis and Nicolis* (1998) have studied an approach in which high order moments are expressed as time-independent functionals of low-order moments, based on a study of dynamical systems which showed that subsets of moments vary on a timescale given by the dominant eigenvalues of the Liouville operator L, defined in Eq. (2). In general, however, this method of moment decomposition has not yet been studied in the context of realistic weather and climate systems.

In conclusion, whilst a formal analytic solution can be found to the problem of predicting the forecast pdf, there are practical problems associated with the dimension of the underlying dynamical system. However, the issue of dimensionality affects the problem in other, more insidious, ways. These are discussed in the next two sections.

3. The probability density function of initial error

In order to discuss how the pdf of initial error can be estimated in weather and climate prediction, it is necessary to outline the method by which observations are used to determine the initial conditions for a deterministic weather or climate forecast.

In meteorology and oceanography, data assimilation is a means of obtaining a forecast initial state which in some well-defined sense optimally combines the available observations for a particular time with an independent background state (*Daley*, 1991). This background state is usually a short-range forecast (e.g. 6 hour) from an estimate

of the initial state valid at an earlier time, and this carries forward information from observations from earlier times. A very simple example of the basic notion can be illustrated by considering two different independent estimates, s_o and s_b, of a scalar s. Suppose that the errors associated with these two estimates are random, unbiased and normally distributed, with standard deviations σ_o and σ_b, respectively. Then the maximum-likelihood estimate of s is the state s_a which minimises the cost function

$$J(s) = \frac{(s - s_b)^2}{2\sigma_b^2} + \frac{(s - s_o)^2}{2\sigma_o^2} \tag{10}$$

The least-squares solution

$$s_a = s_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}[s_o - s_b] \tag{11}$$

is easily found. The error associated with s_a is normally distributed with variance given by

$$\frac{\partial^2 J}{\partial s^2} = \sigma_b^{-2} + \sigma_o^{-2} = \sigma_a^{-2} \tag{12}$$

The data assimilation technique used in weather prediction (e.g. at ECMWF) is a multi-dimensional generalisation of this technique (*Courtier et al.*, 1994, 1998). The analysed state X_a of the atmospheric state vector is found by minimising the cost function

$$J(X) = \frac{1}{2}(X - X_b)^T B^{-1}(X - X_b) + \frac{1}{2}(HX - Y)^T O^{-1}(HX - Y) \tag{13}$$

where X_b is the background state, B and O are covariance matrices for the pdfs of background error and observation error respectively, H is the so-called observation operator, and Y denotes the vector of available observations. For example, if Y includes a radiance measurement taken by an infrared radiometer onboard a satellite orbiting the earth then HX includes an estimate of the infrared radiance that would be emitted by a model atmosphere as represented by the state vector X. Similarly, if Y includes a surface pressure measurement taken at some point ρ on the earth's surface, then HX includes the surface pressure at p given X. Since X is finite dimensional, the operator H inevitably involves an interpolation to p. Similar to Eq. (12), the Hessian of J is given by (*Fisher and Courtier*, 1995)

$$\nabla\nabla J = B^{-1} + H^T O^{-1} H \equiv A^{-1} \tag{14}$$

We refer to A as the analysis error covariance matrix.

In the current ECMWF operational data assimilation system, the background error covariance matrix B is not dependent on the present state of the atmospheric circulation. This is believed to introduce considerable imprecision in the estimate of the initial pdf as given by (14). This estimate can be improved; within the linearised regime (cf. Fig. 1), the forecast error covariance matrix F implied by Eqs. (6) and (8) can be written

$$F(t) = M(t, t_a) A(t_a) M^T(t_a, t) \tag{15}$$

where M is the tangent propagator along the trajectory between the initial state X_a and the forecast state X at time t. Since the time between consecutive analyses (typically 6 hours) is broadly within this linearised regime, then a flow-dependent estimate of the background error covariance matrix at time t_a can be obtained by propagating the analysis error covariance matrix from the earlier analysis time t_{a-1}, i.e.

$$B(t_a) = M(t_a, t_{a-1})A(t_{a-1})M^T(t_{a-1}, t_a) \tag{16}$$

The propagator M, and its transpose M^T are essential components of 4-dimensional data assimilation (*Courtier et al.*, 1994) where observations are assimilated over a time window. Using M, a perturbation δx can be evaluated at the same time that an observation is taken. Given the dimension of comprehensive weather prediction models, M is not known in matrix form, and is represented in operator form (cf. Eq. (8)). Similarly the transpose M^T is also represented in operator form M^* (see Eq. (18) below) and is known as the adjoint (tangent) propagator.

However, Eq. (16) is computationally intractable for numerical weather prediction, requiring $O(10^{14})$ individual linearised integrations of M for a complete specification of the propagated matrix MAM^T. Three possible solutions have been proposed. The first is essentially a Monte Carlo solution, whereby a random sampling of A is evolved using M (*Evensen*, 1994; *Andersson and Fisher*, 1999). The second proposal involves solving the propagation Eq. (16) with an intermediate complexity model (*Ehrendorfer*, 1999). The final proposal (the so-called reduced-rank Kalman filter; *Fisher*, 1998) is to propagate A explicitly only in the appropriate unstable subspace defined by the dominant flow-dependent local instabilities of the attractor. Broadly, speaking, the proposal is to have the best possible knowledge of the initial state in that part of phase space from which forecast errors are most likely to grow. At present these three different proposals are being evaluated.

Since the notion of local flow-dependent instability features strongly in later sections of this paper, it is worth outlining some more detail on how these instabilities can be estimated. First consider a Euclidean inner product <..,..> so that for any perturbations δx, δy,

$$< \delta x, \delta y > = \sum \delta x_i \delta y_i \tag{17}$$

In terms of <..,..> the adjoint tangent propagator M^* is defined by

$$\delta y(t_a) = M^*(t_a, t)\delta y(t) \tag{18}$$

where

$$< \delta y, M\delta x > = < M^*\delta y, \delta x > \tag{19}$$

for an arbitrary pair of perturbations $\delta x(t_a)$, $\delta y(t)$.

The analysis error covariance matrix A defines a secondary inner product

$$(\ldots, \ldots) = < \ldots, A^{-1} \ldots > \tag{20}$$

Here A^{-1} is the covariant form of an analysis error covariance metric, g_{A-1} (*Palmer et al.*, 1998). Hence the perturbation $\delta x(t)$, which has maximum Euclidean amplitude at t and unit g_{A-1} norm at initial time t_a is given by

$$\max_{x(t_a)\neq 0} \frac{< \delta x(t), \delta x(t) >}{< \delta x(t_a), A^{-1}\delta x(t_a) >} = \max_{x(t_a)\neq 0} \frac{< \delta x(t_a), M^*M\delta x(t_a) >}{< \delta x(t_a), A^{-1}\delta x(t_a) >} \tag{21}$$

This is equivalent to finding the dominant eigenvector of the generalised eigenvector equation

$$M^{*}M\delta x(t_{\mathrm{a}}) \;=\; \lambda A^{-1}\delta x(t_{\mathrm{a}}) \tag{22}$$

Formally, by taking the square root of A, Eq. (22) can be transformed to a singular vector equation which can be solved using a Lanczos algorithm (*Strang*, 1986). More generally, Eq. (22) is solved using a generalised Davidson algorithm (*Barkmeijer et al.*, 1998). We refer to the solutions $\delta x(t_{\mathrm{a}})$ in Eq. (22) as g_{A-1}-singular vectors of M.

The set of dominant g_{A-1} singular vectors of M (with largest singular values $\sigma_i \;=\; \sqrt{\lambda_i}$) defines an unstable subspace in the tangent space at t_{a}. It comprises the set of most rapidly-growing directions defined locally on phase space, relative to a basic-state trajectory between t_{a} and t, subject to the constraint that the initial perturbations are normalised with respect to the initial pdf. At forecast time t, these singular vectors have evolved into the major axes of the forecast error ellipsoid, or, equivalently, into the eigenvectors of the forecast error covariance matrix. The first two parts of Fig. 1 show schematically a dominant g_{A-1} singular vector at initial and forecast time. Further discussion of these singular vectors, and their relation to more familiar forms of perturbation growth (such as normal mode and Lyapunov exponent growth) are discussed in Section 5.

4. Representing uncertainty in model formulation

So far, we have assumed the 'classical' chaotic paradigm, that loss of predictability occurs only because of inevitable uncertainty in initial conditions. However, there are also inevitable uncertainties in our ability to represent computationally the governing equations of climate. These uncertainties can contribute both to random and systematic error in model performance. In practice, as discussed below, it is not easy to separate the predictability problem into a component associated with initial error and a component associated with model error.

As mentioned, weather and climate models have a resolution of O(100 km) in the horizontal. This immediately raises the problem of closure - how to represent the effect of partially resolved or unresolved processes onto the resolved state vector X. The effects of topography and cloud systems are examples. In weather and climate prediction models, Eq. (1) is generally expressed as

$$\dot{X} \;=\; G[X] + P[X;\alpha(X)] \tag{23}$$

where $G[X]$ represents terms in the equations of motion associated directly with resolved scales, and $P[X;\alpha]$ stands for some parametrised representation of unresolved processes. Conceptually, a parametrisation is usually based on the notion of a statistical ensemble of sub-grid scale processes within a grid box, in some secular equilibrium with the grid-box mean flow. This allows Eq. (23) to be written

$$\dot{X}_j \;=\; G_j[X] + P[X_j;\alpha(X_j)] \tag{24}$$

where X_j and G_j represent the projection of X and G into the subspace associated with a single grid box x_j in physical space. Borrowing ideas from statistical mechanics, a familiar parametrisation might involve the diffusive approximation, where α would be a diffusion coefficient depending on the Richardson number at x_j. However, whilst diffusive closures do in fact play a role for example in representing the effects of the turbulent planetary boundary in the lower kilometre of the atmosphere, they are certainly insufficient, and, in some circumstances, may be fundamentally flawed. To see this, it is enough to concentrate on one relevant process - atmospheric convection.

On average, the atmosphere is stable to convective overturning. However, on local space and time scales, especially in tropical regions where incoming solar radiation strongly heats the surface of the earth, conditions exist for con-

vective instability (*Emanuel*, 1994). Often, such instability is released through overturning circulations whose horizontal scales are small compared with the smallest resolved scale of a global weather or climate model. One manifestation of such small-scale convective instability is the cumulus cloud beloved by glider pilots. These clouds can develop into precipitating cumulonimbus clouds; typically a few kilometres in horizontal extent, but still much smaller than the smallest resolved scale in a weather or climate model. However, at the other end of the spectrum of convectively-driven circulations is the organised mesoscale convective complex (*Moncrieff*, 1992), feared by aviators in general. Such complexes have horizontal scales of perhaps 50 km. They can be simulated explicitly in regional models with O(1 km) resolution, but such resolution is not practicable for global weather and climate models.

The form of parametrisation given in Eq. (24) is appropriate for describing cumulus and simple cumulonimbus. For example, in a contemporary convective parametrisation (e.g. *Betts and Miller*, 1986) if the resolved-scale vertical temperature gradient at x_j is convectively unstable, then over some prescribed timescale (given by α) P will operate to relax X_j back to stability. On the other hand, the existence of organised mesoscale convective complexes poses a problem for parametrisations of this form. In particular, the basic assumption of a quasi-equilibrium of subgridscale convectively-forced motions (with the implication that the kinetic energy released by overturning circulations is dissipated on sub-grid scales, rather than injected into the large scale) cannot be fully justified.

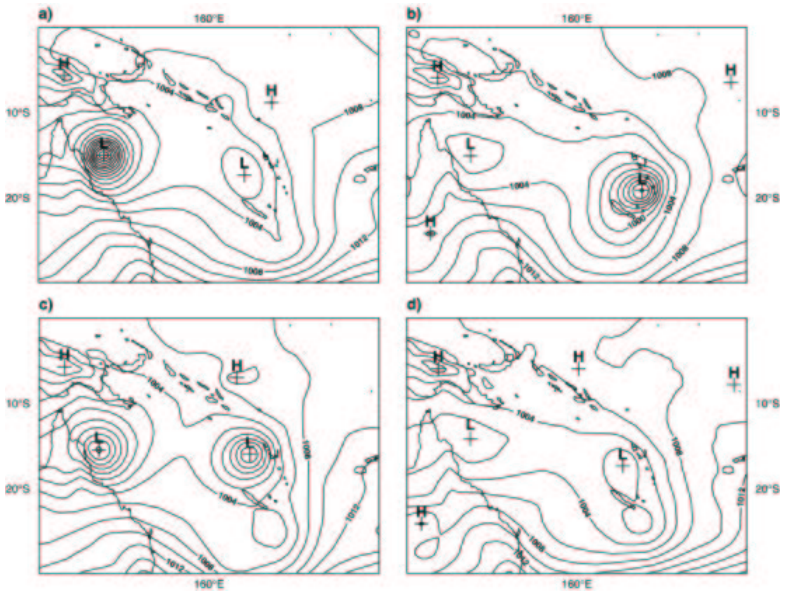

Figure 4: Four 2-day integrations of the ECMWF model from identical starting conditions but different realisations of the stochastic parametrisation scheme represented by Eq. (25) with parameter settings as given in Buizza *et al.* (1999). The field shown is sea-level pressure over parts of Australia and the west Pacific. The depressions in the pressure field represent potential tropical cyclones.

One means of addressing the erroneous assumption of deterministic locality in Eq. (24) would be to add to Eq. (24) a stochastic energy source term $S(X_j;\alpha)$. Recognising this, *Buizza et al.* (1999a) have proposed the simple stochastic form

$$\dot{X}_j = G_j[X] + \beta P[X_j;\alpha(X_j)] \tag{25}$$

where β is a stochastic variable drawn from a uniform pdf with values between 0.5 and 1.5. A particular realisation of β would be used over more than one model gridpoint and timestep, thus making the overall parametrisation both

stochastic and nonlocal. Stochastic representation of sub-grid processes is a technique already utilised in turbulent flow simulations (*Mason and Thompson*, 1992).

An example of the impact of the random effect of the stochastic parametrization represented by Eq. (25) is given in Fig. 4 , which shows sea-level pressure over part of Australia and the west Pacific from four 2-day integrations of the ECMWF model. The integrations have identical starting conditions, but different realisations of the pdf represented by β . The figure shows two tropical cyclones. The intensity of the cyclones can be seen to be very sensitive to the realisation of the stochastic parametrisation. In Fig. 4 (a) the western cyclone is intense; in Fig. 4 (b) the eastern cyclone is intense; in Fig. 4 (c) they are both intense; in Fig. 4 (d) neither is intense. This rather extreme example clearly shows the difficulty in predicting tropical cyclone development, and its sensitivity to sub-grid scale parametrisation.

It is possible that a systematic misrepresentation of small-scale variability (such as convectively organised mesoscale systems) can have a systematic effect on larger-scale climatic variability. There is evidence (L. Ferranti, C. Jakob, personal communication) that the parametrisation of organised convection is fundamental in determining the strength of the so- called Madden-Julian oscillation - the dominant mode of intraseasonal variability in the tropics (*Madden and Julian*, 1994), and poorly simulated in many climate models. Furthemore, as suggested by *Moore and Kleeman* (1996), the (initial-time) singular vectors of El Niño (see Section 5 below), have a strong projection onto the Madden-Julian oscillation. Hence, by a series of nonlinear processes, it is possible to infer that the climatology of El Niño may be linked to the climatology of organised mesoscale convection in the tropical Pacific. Quantifying these links between the small scale and large scale is an area currently under investigation at ECMWF.

The purpose of the discussion above is to point out that although current parametrisations have been enormously successful in representing subgrid processes, there are inevitable uncertainties in the representation of such processes. In Section 2 we represented the evolution of the initial pdf given the deterministic Eq. (1) in terms of a Liouville equation. In the idealised case where model uncertainties are represented by an additive Gaussian white noise with zero mean and variance Γ , then Eq. (2) becomes a Fokker–Planck equation (e.g. *Hasselmann*, 1976; *Moss and McClintock*, 1989).

$$\frac{\partial \rho(X)}{\partial t} + \frac{\partial}{\partial x}\left[F[X]\rho(X) - \frac{\Gamma}{2}\frac{\partial \rho(X)}{\partial X}\right] = 0 \tag{26}$$

However, in practice, a realistic state-dependent stochastic forcing (e.g. in Eq. (25)) would be too complex for this simple representation to be directly relevant.

For ensemble forecasting (discussed in Section 7), there are two other commonly-used techniques for representing model uncertainty. The first is the multi-model ensemble (*Harrison et al.*, 1999; *Palmer et al.*, 2000). This can be achieved by incorporating within the ensemble a number of quasi-independent models (e.g. constructed from the stock of climate models developed by different groups around the world; *Gates*, 1992). In the second technique, the values α of the parameters used in the deterministic Eq. (23) of one particular model, are perturbed in some stochastic manner (cf. *Houtekamer et al.*, 1996).

On the other hand, these latter techniques should be seen as conceptually distinct from the type of stochastic physics scheme described schematically in Eq. (25). In multi-model ensembles, and ensembles with stochastic α , the model perturbations account for the fact that the expected value of the pdf of sub-grid processes is not itself well known. (Hence, for example, there are many different atmospheric convection parametrisation schemes in use around the world; the Betts-Miller scheme described above is but one of these. The existence of this ensemble of convection schemes is an indication that the expectation value of the pdf of the effects of subgrid convection is not known with complete confidence!) By contrast, the stochastic physics scheme described in Eq. (25) is an attempt to account for the fact that in circumstances of convective organisation, the pdf of sub-grid processes is not espe-

cially sharp around the mean. This would argue for the combined use of multi-model ensembles and stochastic physics parametrisation.

As discussed in Section 3, the initial state for a weather prediction is determined by assimilating observations onto a background field obtained from a short-range integration from an earlier initial condition. As such, model error clearly plays a role in the determination of the initial pdf. In general, it is difficult to separate cleanly the role of model error and observation error, even though information from the observations and the background state are input into separate terms in the cost function (13). For example, as mentioned in Section 3, local observations are assimilated by comparing with HX, where H is an observation operator which includes an interpolation from the grid point to the observation point. This interpolation will necessarily be affected by random and systematic errors in the representation of unresolved or poorly-resolved scales.

5. Error growth in the linear and nonlinear phase

5.1 Singular vectors, eigenvectors and Lyapunov vectors

As discussed in Section 3, the initial phase-space directions which evolve into the major axes of the ellipsoid of forecast error pdf are determined by the dominant g_{A-1} singular vectors of the forward tangent propagator M. Singular vectors of M have been studied extensively in weather prediction models using simpler metrics than g_{A-1} (*Lacarra and Talagrand*, 1988; *Molteni and Palmer*, 1993; *Buizza and Palmer*, 1995). One simple choice is the energy metric g_E. The g_E singular vectors of M are the dominant eigenvectors of M^*M where '*' now denotes the adjoint defined with respect to the inner product $< \ldots, E \ldots >$ where $< \delta x, E \delta x >$ is the total energy of the perturbation δx. An explicit expression for E is given in *Buizza and Palmer* (1995).

An example of a g_E singular vector is given in Fig. 5 , calculated using the forward and adjoint tangent propagators of the ECMWF weather prediction model, optimised over a three-day period, with starting conditions on 9 January 1993. The streamfunction (inverse Laplacian of vorticity) associated with the singular vector is shown at three levels in the atmosphere, at initial and optimisation time. The singular vector has some qualitative resemblance to an idealised baroclinically unstable eigenmode of the atmosphere (*Charney*, 1947; *Eady*, 1949); for example the disturbance amplifies as it propagates through the region in the west Atlantic where north-south surface temperature gradients are largest, and the disturbance shows evidence of westward tilting phase with height, consistent with a northward flux of heat (*Gill*, 1982).

However, the singular vector also clearly illustrates non-modal characteristics. At initial time the disturbance is localised over the west Atlantic, at optimisation time the disturbance has propagated downstream to Europe. At initial time, maximum disturbance amplitude is located in the lower troposphere, whilst at final time amplitude is largest in the upper troposphere at the level of maximum winds. Finally, the horizontal scale of the initial disturbance is noticeably smaller at initial time than at optimisation time.

A simple way of understanding the non-modal vertical structure of the singular vector is to consider a steady zonally symmetric basic state flow u_0, which is slowly varying in the vertical. In such a flow, the wave action $E/(\omega - k u_0)$ of a small-amplitude disturbance with energy E, zonal wavenumber k and frequency ω will be conserved as it propagates vertically on the background flow (e.g. *Gill*, 1982). Optimal energy growth will therefore. tend to be associated with propagation away from a region of small intrinsic frequency (near the so-called baroclinic steering level usually located in the-lower troposphere), to a region of large intrinsic frequency (such as would occur at the jet stream level in the upper troposphere).

The upscale evolution is consistent with an inverse energy cascade characteristic of two dimensional turbulence (see below). At first sight it might appear paradoxical that such a nonlinear property can be emulated by a linearised

calculation. However, this non-modal characteristic of the linear perturbation is a consequence of the non-normality (i.e. $MM^* \neq M^*M$) of the tangent propagator (*Farrell and Ioannou*, 1996). In turn non-normality derives its existence from the advective nonlinearity in Eq. (1). Hence, in some sense, the cascade properties associated with nonlinear advection in Eq. (1) are partially reflected in the non-normality of M.

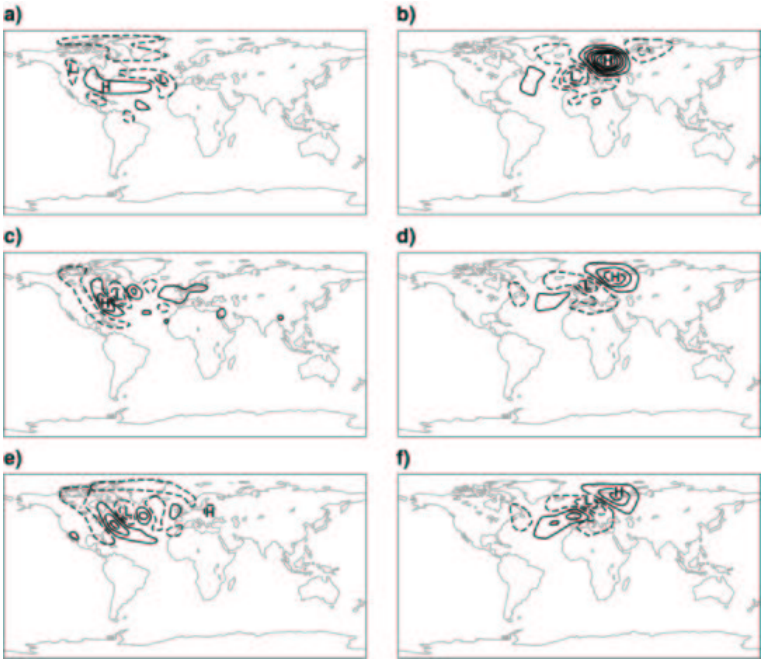

Figure 5: Streamfunction of the dominant singular atmospheric singular vector calculated using a primitive equation numerical weather prediction model for a 3-day trajectory portion made from initial conditions of 9 January 1993 at: (a) and (d) 200 hPa; (b) and (e) 700 hPa; (c) and (f) 850 hPa. The quantities in (a) - (c) are at initial time, in (d) - (f) at final time. The contour interval at optimisation time is 20 times larger than at initial time. From *Buizza and Palmer* (1995).

The ability of singular vectors to describe upscale evolution can be illustrated in simpler 'toy models'. *Hansen* (1998) has studied singular vectors of the two-scale dynamical system

$$\begin{aligned} \dot{x}_i &= -x_{i-2}x_{i-1} + x_{i-1}x_{i+1} - x_i + F - \frac{c}{b}\sum_{j=1}^{N} y_{j,i} \\ \dot{y}_{j,i} &= -cby_{j+1,i}y_{j+2,i} + cby_{j-1,i}y_{j+1,i} - cy_{j,i} + \frac{c}{b}x_i \end{aligned} \tag{27}$$

of *Lorenz* (1996). Here x_i are taken as large-scale variables, $y_{j,i}$ as small-scale variables. For each large-scale variable, there are N small-scale variables. Hansen maximises

$$\frac{< M_{\mathrm{L95}}\delta x, \mathrm{P}M_{\mathrm{L95}}\delta x >}{< \delta x, \delta x >} \tag{28}$$

where M_{L95} is the forward tangent propagator of Eqs. (27), and P projects the entire state vector onto the subspace of the large-scale x variables. For particular choices of the parameters b, c and F, the dominant singular vector which maximises amplitude in the large-scale variables at optimisation time, is comprised almost entirely of small-

scale components at initial time, emulating the upscale cascade exhibited in this nonlinear model.

It is also possible to compute singular vectors of the tangent propagator from intermediate coupled ocean-atmosphere models (e.g. *Moore and Kleeman*, 1996; *Chen et al.* 1997; *Xue et al.* 1997). As with the atmosphere-only singular vectors, the growth characteristics are extremely non-modal. The surface temperatures of the evolved singular vectors at optimisation time is characteristic of the El Niño phenomenon itself, whilst at initial time the singular vector has little spatial correlation with the El Niño pattern. The singular values are strongly dependent on time of year, being largest from (boreal) spring to autumn, and weakest from (boreal) autumn to spring. As will be argued below, the seasonal cycle in these coupled ocean-atmosphere singular vectors may be vital in understanding how the climate responds to external forcing perturbations.

Let us consider the relationship between singular vector growth and eigenvector growth when Eq. (1) is linearised about a stationary solution. In this case, normalised eigenvectors ξ_i of J with eigenvalues p_i give rise to modal solutions $\xi_i \exp[\mu_i(t-t_a)]$.) (The propagator M will also have eigenvectors ξ_i but with eigenvalues $\exp[\mu_i(t-t_a)]$.) Irrespective of normality, eigenvectors η_j and eigenvalues θ_j of the adjoint Jacobian satisfy the biorthogonality condition

$$(\mu_i - \theta_i^{cc}) < \xi_i, \eta_i > \; = 0 \tag{29}$$

where ‘cc' denotes the complex conjugate.

If an initial disturbance is written in terms of the eigenmodes ξ_i, i.e.

$$x(t) \; = \sum_i c_i \xi_i \exp\{\mu_i(t-t_a)\} \tag{30}$$

then from Eq. (29)

$$c_i \; = \frac{< \eta_i x(t_a) >}{< \eta_i, \xi_i >} \tag{31}$$

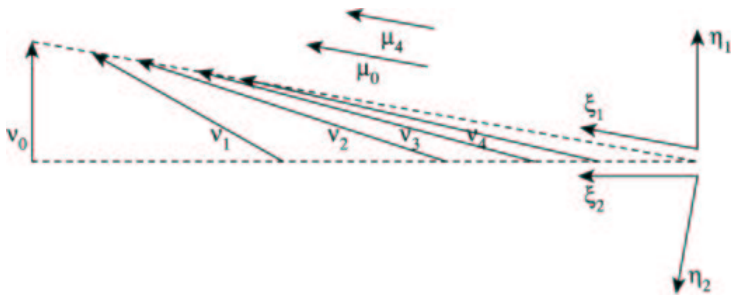

Figure 6: This diagram illustrates schematically the crucial difference between eigenvector and singular vector growth, and the relationship between singular vectors and adjoint eigenvectors. See text for details. From *Buizza and Palmer* (1995).

From Eq. (30), the fastest growing eigenmode ξ_1 will ultimately contribute most to the growth of $x(t)$. In order to maximise the contribution of the first eigenmode, c_i should be as large as possible. From Eq. (31), this occurs when $x(t_a)$ is parallel to η_1 (for a non self-adjoint operator, $\xi_i \neq \eta_1$.) Hence, in order to maximise the asymptotic amplitude of $x(t)$, the initial perturbation should not project onto the fastest-growing eigenmode, but onto its ad-

joint.

Fig. 6 illustrates schematically the crucial difference between eigenmode and singular vector growth. An idealised 2D system has two very non-orthogonal decaying eigenmodes ξ_1 and ξ_2; we take ξ_1 to have the larger real eigenvalue component. The adjoint eigenmodes η_1 and η_2 are also shown, consistent with the biorthogonality condition (29). A normalised vector v_0 is shown parallel to η_1; the sequence of vectors $v_1, \ldots\ldots, v_4$ show the time evolution of v_0. The projection of v_4 onto ξ_1 is much larger than that of a second vector μ_4 which was initially normalised and aligned along ξ_1.

Another familiar measure of perturbation growth is given by the Lyapunov exponent. The system's Lyapunov exponents can be defined as

$$\bar{\lambda}_i = \lim_{t \to \infty} \lim_{d_i(0) \to 0} \frac{1}{t} \ln \frac{d_i(t)}{d_i(0)} \tag{32}$$

where d_i is the eigenvalue of the i-th eigenvector of the operator

$$\lim_{t \to \infty} \frac{1}{t} [MM^{\mathrm{T}}]^{1/2t} \tag{33}$$

These eigenvectors, or Lyapunov vectors, correspond to (instantaneous) realisations of evolved singular vectors, for long optimisation times.

Although Lyapunov exponents are normally associated with mean growth over the attractor, local Lyapunov exponents can be defined as the growth rates

$$\lambda_i = \frac{\mathrm{d}}{\mathrm{d}t} \ln d_i(t) \tag{34}$$

Just as singular vector growth can exceed eigenmode growth for stationary flows, so it can also exceed local Lyapunov growth for transient flows (*Trevisan and Legnani*, 1995).

In fact in a multi-scale system, the concept of Lyapunov exponent growth is not a useful one. For example, in Eq. (27), the dominant Lyapunov exponent is determined (for $c > 1$) by the growth of the small scale y variables. However, if b is sufficiently large, then their effect on the predictability of the x variables will be small compared with the effect of initial uncertainty in the large-scale x variables themselves. In the atmosphere, Lyapunov exponent growth would be associated with small-scale convective instability, rather than large-scale more slowly growing baroclinic instability (which determines the structure of extratropical weather systems). As such, the predictability of these weather systems can be much longer than a dominant atmospheric Lyapunov timescale.

One way over overcoming this problem with the Lyapunov vector is through the so-called breeding-vector modification (Toth *and Kalnay*, 1993, 1997). A bred vector is obtained by integrating the model twice over a time ?t from initial conditions differing by a small random perturbation. The resulting difference field is rescaled to the initial perturbation amplitude, and the procedure repeated. The essential difference between a bred vector and a Lyapunov vector is that the time Δt is chosen to be sufficiently long that instabilities on scales much smaller than weather, which would in practice determine the atmosphere's dominant Lyapunov exponent, are damped by nonlinear saturation.

5.2 Error dynamics and scale cascades

A fundamental characteristic of error growth in climate and weather prediction is the upscale cascade of error. This characterises the essential element of the 'butterfly effect' paradigm, as much as does amplitude growth (i.e. initial conditions are sensitive not only to small amplitude error, but also small-scale error). In this section, we briefly describe estimates of predictability arising from cascade processes, using from simple scaling arguments.

Let k denote a characteristic wavenumber in such a system, and $E(k)$ denote the energy kinetic energy per unit wavenumber, at wavenumber k. Assume that observations cannot resolve wavenumbers higher than k_r. Hence the initial pdf will certainly be non-zero in the phasespace sub-manifold associated with wavenumbers higher than kr. We present below a heuristic argument for determining the time it takes this component of initial error to infect the meteorological scales $k_m \ll k_r$. Following *Lorenz* (1969) and *Lilly* (1973), let us assume that the time it takes before complete uncertainty at wavenumber $2k$ to strongly infect wavenumber k, is proportional to the 'eddy turnover time' $\tau(k) = k^{-3/2}[E(k)]^{-1/2}$. The time $\Omega(N)$ taken for uncertainty to propagate from wavenumber $2^N k_m$ to wavenumber k_m is therefore given by

$$\Omega(N) = \sum_{n=0}^{N-1} \tau(2^n k_m) \tag{35}$$

In the case of a two-dimensional isotropic homogeneous turbulence in the inertial subrange between some large-scale (e.g. baroclinic) forcing scale and dissipation scale, then $E(k) \sim k^{-3}$, τ is independent of k, and $\Omega(N) \sim N$ which diverges as $N \to \infty$. Hence, although uncertainty at small scales (within the inertial range) can infect the larger meteorological scale, in theory long-range predictions are possible providing the initial data has sufficient small-scale resolution. This two dimensional paradigm is appropriate to weather systems where quasi-geostrophic scaling is appropriate (Charney, 1971).

By contrast, for homogeneous isotropic 3-dimensional flow $E(k) \sim k^{-5/3}$ and $\tau \sim k^{-2/3}$, the familiar Kolmogorov inertial range (*Frisch*, 1995). In this case $\Omega(N)$ tends to a finite limit as $N \to \infty$, that is

$$\Omega(\infty) \sqsupset 2.7\tau(k_m) \tag{36}$$

This remarkable result says that the predictability time for a large-scale system is on the order of an eddy turn-over time, a few days. In the limit of infinite Reynolds number, this is consistent with the hypothesised lack of existence of a unique smooth solution of the Navier-Stokes equations in the Euler limit (*Frisch*, 1995), since it implies that uncertainty on arbitrarily small scales can destroy predictability on large scales in finite time.

However, for the quasi-2D large-scale weather and climate systems discussed in this paper, this 3D paradigm does not appear to be directly relevant; for scales O(100 km) or larger, horizontal scales dominate the vertical scale (the aspect ratio is small). However, there is evidence for a $k^{-5/3}$ spectrum on scales of $\sim 100 - 800$ km (*Nastrom and Gage*, 1985; *Gage and Nastrom*, 1986); *Cho et al.*, 1999). This is likely to be associated with a 2D inverse energy cascade (*Kraichnan*, 1971) possibly forced by organised convective activity (*Lilly*, 1983), of the type discussed in section 4. This suggests that in the atmosphere there are two sets of 2D inertial ranges, an k^{-3} enstrophy cascading range associated with forcing from weather systems at large scale, and a smaller-scale $k^{-5/3}$ inverse energy cascading range associated with forcing from mesoscale variability on small scales (*Lilly*, 1983, 1989). As such, uncertainties on these mesoscales may influence the predictability of large-scale weather systems within a few days, and reinforces the need for a stochastic representation of such effects in global weather and climate models (see Section 4).

6. APPLICATIONS OF SINGULAR VECTORS

In this section we describe some applications of the singular vector analysis described above.

6.1 Data assimilation

As described in Section 3, the approximation of treating the background error covariance matrix as flow independent can lead to significant imprecision in defining the initial pdf. However, the integration of Eq. (16) is currently computationally impossible using comprehensive weather prediction models. As discussed above, a reduced-rank solution has been developed (*Fisher*, 1998) in which the initial pdf is evolved explicitly in the space of dominant singular vectors using a g_{A-1} metric.

6.2 Chaotic control of the observing system

The second application can be thought of as an example of the principle of chaotic control. Suppose we could supplement the standard observing network with additional observations, and that, in principle, these observations could be made anywhere in the atmosphere, where should they be made? These observations would have the biggest impact on forecast error if they were made in the region of the dominant g_{A-1} singular vector (*Palmer et al.*, 1998).

Consider, for example, the singular vector calculation shown in Fig. 5 . The actual period corresponded to a time of storminess over the North Atlantic (the oil tanker Braer ran aground on the Shetland Islands, leading to fears of a major environmental disaster). As can be seen the evolved dominant singular vector is located over the north-east Atlantic. In order to improve the quality of a 3-day weather forecast over this region of the north-east Atlantic (and assuming that g_E can be taken as a reasonable approximation for g_{A-1}), the location of the initial singular vector suggests that additional observations should be taken near the eastern seaboard of the US, particularly in the mid-lower troposphere.

However, since singular vectors are flow dependent, a fixed-location observation at the site of this particular singular vector would have been highly sub-optimal a few days later. On the other hand, new pilotless aircraft technology (e.g. the aerosonde; *Holland et al.*, 1992) enables such flow-dependent targeted observations. to be made. Essentially, the aerosonde can be programmed to fly to the principal locations of the dominant singular vectors, measure vertical profiles of wind and temperature, and transmit the data to satellite. The additional data would be assimilated in the determination of the initial state of the atmosphere. Real experiments testing this possibility are currently in the design phase.

Such types of observation are not yet possible on an operational basis. However, the basic concept discussed in this sub-section was tested during the Fronts and Atlantic Storm Track Experiment (FASTEX; *Joly et al.*, 1997), during the winter of 1997. Aircraft were flown to the locations of the dominant g_E singular vectors over the Atlantic, and special dropsondes were released. These dropsondes measured wind and temperature profiles. Sometimes these extra targeted data had a substantial impact on weather forecast quality over Europe (*Montani et al.*, 1999). It should be noted that for such applications, the singular vectors are targeted to produce maximum growth in a region defined by a projection operator P (e.g. P = 1 for European grid points, P = 0 otherwise). The corresponding dominant singular vectors maximise

$$\frac{< M\delta x(t_a), \mathrm{P}M\delta x(t_a) >}{< \delta x(t_a), A^{-1}\delta x(t_a) >} \tag{37}$$

6.3 The response to external forcing: paleoclimate and anthropogenic climate change

The singular vector analysis described above is not only relevant in describing the system's response to initial errors, but also in describing the system's response to some external forcing. Consider the problem of the linearised response of the climate system to a weak imposed external forcing $f(t)$. As mentioned in the introduction, this forcing could also represent a prescribed perturbation e.g. radiative forcing due to orbital variations or volcanic eruption, or a doubling of atmospheric CO_2. Adding f to Eq. (4) we have

$$\dot{\delta}x = J\delta x + f(t) \tag{38}$$

Using the tangent propagator M, the solution to (38) can be written over the finite interval $[t_a, t]$ as

$$\delta x(t) = M(t, t_a)\delta x(t_a) + \int_{t_a}^{t} M(t, t')f(t')\mathrm{d}t' \tag{39}$$

If f is time independent over $[t_a, t]$, and setting $\delta x(t_a) = 0$,

$$\delta x(t) = \mathrm{M}(t, t_a)f \tag{40}$$

where

$$M = \int_{t_a}^{t} M(t, t')\mathrm{d}t' \tag{41}$$

The system's response to f depends on the projection of f onto local phase-space instabilities. To see this more explicitly, perform a singular vector decomposition on M, so that

$$\mathrm{M} = U\Sigma V^{*} \tag{42}$$

where U is the matrix whose columns contain the left singular vectors u_i of M, V is the matrix whose columns contain the right singular vectors v_i of M, and Σ is a diagonal matrix whose real elements are the singular values of M.

If we expand f as

$$f = \sum_{i=1}^{N} \alpha_i \mathrm{v}_i \tag{43}$$

then, from Eq. (40)

$$x(t) = \sum_{i=1}^{N} \sigma_i \alpha_i u_i \tag{44}$$

Now if f is normalised, then the largest possible response is obtained if f is equal to the leading right singular vector of M, so that $\alpha_i = \delta i 1$. However, this is a rather contrived situation. If, by contrast, f projects uniformly onto the so that $\alpha_i = 1/N$, then the largest contribution to $x(t)$ will come from the dominant (left) singular vec-

tor u_i. The state vector will be relatively insensitive to f in regions of the attractor where σ_1, is small, and relatively sensitive in regions of phase space where σ_1, is large.

Fig. 7 shows an example of the variable impact of a small fixed imposed forcing on the *Lorenz* (1963) model, with equations

$$\begin{aligned} \dot{X} &= -\sigma X + \sigma Y + f \\ \dot{Y} &= -XZ + rX - Y + f \\ \dot{Z} &= XY - BZ \end{aligned} \tag{45}$$

The leading singular value σ_1 (for a finite length trajectory of 0.12 non-dimensional Lorenz time units) is shown in Fig. 7 (a), clearly showing that the attractor instability is localised near to the origin (consistent with the ensemble integrations shown in Fig. 2). The scalar plotted in Fig. 7 (b) is $[\sum_{i=1}^{3} \sigma_i^2 \alpha_i^2]^{1/2}$; it shows clearly that the Lorenz system is most sensitive to the imposed forcing in the region where the dominant singular value is large.

From this it is easy to see that the response to a forcing with zero time mean, need not itself be zero (consider for example, in Eq. (45), the forcing $f = \varepsilon(Z - \bar{Z})$, where $\bar{Z}$ is the time-average value of the Z variable in the unperturbed Lorenz equations. When $Z \sim 0$ the state vector will be in a relatively sensitive region and the (negative) forcing will be relatively effective. By contrast, when $Z > \bar{Z}$, the state vector will be in relatively insensitive region and the positive forcing will be relatively ineffective.

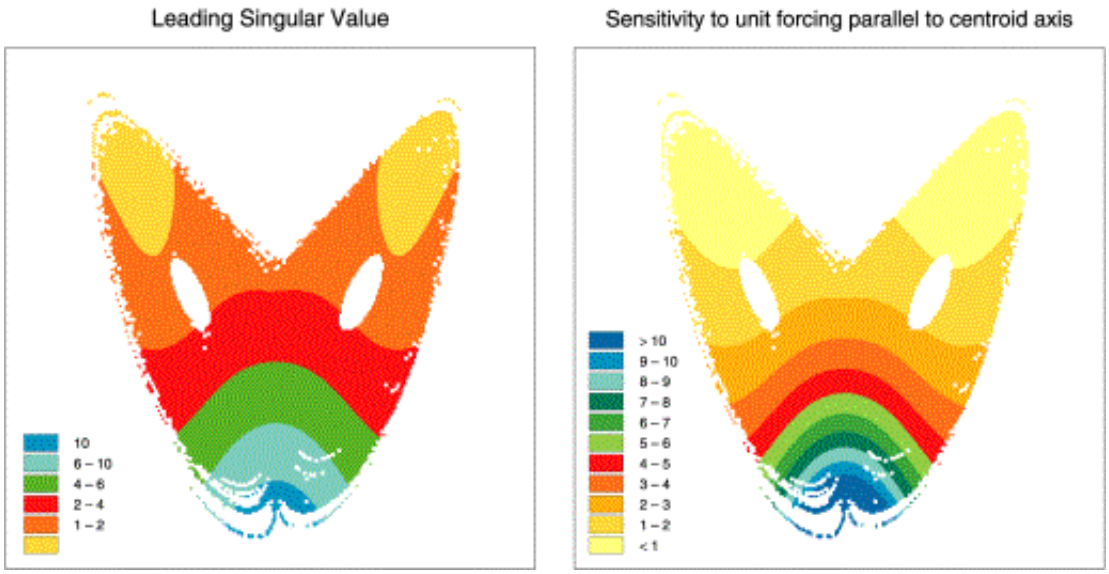

Figure 7: (a) The dominant singular value of the *Lorenz* (1963) model based on trajectory lengths of 0.12 Lorenz time units. (b) The amplitude of the local response to a weak constant forcing applied to the Lorenz model (see Eq. (45)).

The importance of this variable response has been demonstrated in the context of paleoclimate simulation. *Clement and Cane* (1999) and *Cane and Clement* (1999) have studied the response of the intermediate-complexity *Zebiak and Cane* (1987) coupled ocean-atmosphere model of the tropical Pacific, to paleoclimatic variations in solar insolation (due mainly to precession of the equinoxes). While the annual mean forcing from these orbital perturbations is close to zero in the tropics (or indeed averaged globally), the annual mean response of the system to the imposed forcing is not zero. This can be understood from the study of El Niño singular vectors of the coupled

ocean-atmosphere system discussed above, where it was noted that the dominant singular value of the tangent propagator M of a tropical Pacific coupled ocean-atmosphere model is larger for 6 month trajectory portions starting in (boreal) spring than from trajectory portions starting in (boreal) autumn. Hence a sinusoidal forcing $\varepsilon \sin \omega t$ with $2\pi/\omega = 1$ year, can excite an annual average response. *Cane and Clement* (1999) argue that to understand the fact that many observed paleoclimate changes are simultaneous in both hemispheres, the tropics must provide an important intermediary between global climate and orbital forcing. If these ideas prove correct, then this will provide a fundamentally new paradigm for understanding paleoclimate variability.

The same general ideas may apply when considering the atmosphere's response to anthropogenic perturbations, discussed in detail elsewhere (*Palmer*, 1993b, 1999). For example, it is commonly imagined that the climate response to anthropogenic forcing should be distinct from patterns of natural climate variability. On the other hand, consider Eq. (45). Largely independent of the details of the imposed forcing, the system's response to an imposed forcing f is associated with an increase in the pdf associated with one of the Lorenz regimes, and a decrease in the pdf associated with the other regime. (With $f = 0$ in Eq. (45), the pdfs associated with the two regimes are identical.) On the other hand, the phase-space locations of the regime centroids are largely unaffected by the imposed forcing. This can be understood by noting that the regime centroids are phase-space regions of relative stability; the system is sensitive to the imposed forcing in a region of the attractor between the two centroids (see Fig. 7 (b)).

If this picture were applicable to the climate system, it would imply that anthropogenically-forced changes in climate would project primarily onto the principal patterns of natural variability, even though such natural variability may occur predominantly on timescales much shorter than that of the imposed forcing. *Corti et al.* (1999) have argued that this is indeed the case based on an analysis of changes to the frequency of occurence of atmospheric circulation regimes defined from observed hemispheric circulation data for from 1949 to 1994.

Hence in predicting climate change, particularly on a regional basis, it will be very important to predict changes in the frequency of occurrence of familiar circulation patterns. For example, the UK can expect milder winters if the frequency of occurence of zonal circulation regimes which bring mild maritime air over the European continent (e.g. associated with the positive phase of the North Atlantic Oscillation; *Hurrell*, 1995), increases. Conversely winters can be expected to be more severe if the frequency of so-called blocked regimes with persistent anticyclonic flow, increases. Results suggest that these regime frequencies are sensitive to changes in the diabatic heating fields in the tropics (*Trenberth et al.*, 1998). From this point of view, expecting observed regional climate change to be directly predictable from perturbations to radiative forcing is a gross oversimplification of the problem.

Indeed, these results indicate that predictions of anthropogenic climate change, particularly on a regional basis, require models which can simulate accurately natural circulation regimes and their associated variability, even though the dominant timescale of such variability may be much shorter than the climate-change signal itself. Moreover, since variability in these circulations regimes is essentially chaotic, the impact of anthropogenic forcing on the frequency of occurrence of a particular regime over a particular period of time, cannot be deduced deterministically; such predictions are necessarily probabilistic. The generation and verification of probability forecasts from global climate prediction models are discussed in the sections below.

6.4 Initialising ensemble forecasts

A third application of the singular vector technique is in providing a manageable set of initial conditions which sample the unstable subspace of the initial pdf, to provide the starting conditions for an ensemble of integrations to determine the forecast pdf into the nonlinear range. This application will be discussed in detail in the next section.

7. FORECASTING UNCERTAINTY BY ENSEMBLE PREDICTION

In the discussion above, the problem of forecasting uncertainty in weather and climate prediction has been formulated in terms of a Liouville, or, including model uncertainty, a form of Fokker-Planck equation. In practice these equations are solved using ensemble techniques i.e. from multiple integrations of the governing equations from perturbed initial conditions, using multiple models and/or stochastic parametrizations to represent model uncertainty. Insofar as the perturbed initial conditions constitute a random sampling of the initial pdf, this could be described as a Monte-Carlo approach. However, the methodologies discussed below for estimating the initial perturbations do not necessarily constitute such a random sampling, for reasons to be discussed. As such, the more general term "ensemble prediction" is used.

In the discussion below, it is useful to describe separately the problem of forecasting predictability on timescales of days, seasons, and beyond.

7.1 Global weather prediction: from 1-10 days

On timescales of 1 day or so, the extratropical weather forecast problem could be described conceptually as the prediction of the development of individual baroclinic disturbances, and their associated weather (fronts, cloud, rainfall etc). The notion that extratropical weather systems can be described qualitatively in terms of baroclinic eigenmodes has been mentioned in Section 5 above.

On the other hand, atmospheric circulations exhibit a degree of organisation on timescales of 10 days or so. From a popular perspective, the atmosphere exhibits “spells of weather”, characterised by a run of similar weather systems, or an extended period marked by the absence of weather systems. More precisely there is evidence of “circulation regimes” characterised by persistence on timescales much longer than an individual weather system, but with transitions between regimes characterised by the faster timescale e.g. of the dominant baroclinic instability (*Cheng and Wallace*, 1993; *Kimoto and Ghil*, 1993; *Mo and Ghil*, 1988; *Molteni et al.*, 1990; *Vautard* 1990). The prediction of persistence or changes in circulation regime could be said to be characteristic of the 10-day medium-range forecast problem.

The existence of these two timescales is illustrated in the *Lorenz* (1963) model, whose attractor has two clear regimes of activity (see Fig. 2). For commonly-used values of the parameters in the model, a typical regime residence timescale is longer than the timescale of oscillation about a regime centroid, or of the timescale of transition between regimes. In this context, the Lorenz model is a paradigmatic 'toy model' of atmospheric circulation regime behaviour. Intermediate models based on atmospheric quasi-geostrophic equations also show regime behaviour. An example (discussed further below) is the nonlinear quasi-geostrophic model of *Reinhold and Pierrehumbert* (1982) with two zonal and two meridional wavenumbers producing a topographically-forced planetary scale interacting with a baroclinically- unstable cyclone scale. Phase-space trajectories vacillate aperiodically between two long-lived weather regimes; results from this model are discussed below.

There is evidence of similar behaviour in tropical atmospheric circulations. Consider, for example, the Asian monsoon. Individual monsoon ‘depressions’, associated with an instability of the monsoon jet stream, develop on timescales similar to (or faster than) extratropical instabilities, and thus characterise the short-range forecast problem. On the other hand, the so-called active and break phases of the monsoon (which characterise the intraseasonal fluctuations of the monsoon) appear to have distinct regime dynamics with typical residence timescales on the order of 10-20 days (*Webster et al.*, 1998).

Hence on the timescale of 10 days, it is important to be able to estimate uncertainty in predictions of persistence or change in regime type. In the last few years, ensemble forecasting on this 10-day timescale has become an established part of operational global weather prediction (*Palmer et al.*, 1993, *Toth and Kalnay*, 1993, *Houtekamer et al.*, 1996, *Molteni et al.*, 1996). Different strategies have been proposed for determining the ensemble of starting

conditions. Conceptually, the strategies can be delineated according to whether the initial perturbations merely sample observation error, or whether (given the dimension of phase space and the uncertainty in our knowledge of the initial pdf) the initial perturbations are constrained to lie in some dynamically unstable subspace. In the latter case, this dynamical subspace can be defined in different ways. Within the meteorological community, there has been a very lively debate on the merits and deficiencies of the different strategies.

A recent analysis of the problem has been performed by *Anderson* (1997) using the *Lorenz* (1963) model. First, points are sampled at regular intervals on the attractor. A pdf O of "observation error" is then assumed. A sampled point on the attractor is perturbed with a realisation of O; this is the "analysed state". A 2-member ensemble is then generated by perturbing about the analysed state with perturbations $\pm p$, where p is again drawn from O. The integration from the state on the attractor is taken as "truth". Based on standard measures of skill (see Section 8 below), this ensemble is more skilful than any other 2-member ensemble, e.g. based on singular-vector perturbations of the analysed state. Similar results have been found by *Hamill et al.* (1999) based on an intermediate quasi-geostrophic model. This study concludes that it is enough in generating an ensemble of initial conditions to randomly perturb the observations Y (see Section 3) consistent with the observation error covariance matrix O.

On the other hand, this type of experiment is rather idealised. On the basis of the discussion in section 3, a sampling of the initial pdf obtained from Eq. (13) by perturbing observations, is likely to be a gross underestimate of (for example) the second moment of the initial pdf. Unfortunately, as discussed, there are so many unquantified uncertainties in the actual details of the data assimilation procedure, it is not at all straightforward to quantify this underestimation.

The strategies that use dynamically-constrained perturbations, in some sense bypass the quantification of these uncertainties, and focus on perturbations that are necessarily growing, and hence are likely to contribute to significant forecast error. Two types of dynamically-constrained perturbation have been proposed: the first based on the "bred vectors" (*Toth and Kalnay*, 1993) and the singular vectors discussed in section 5 (*Mureau et al.*, 1993).

A study of differences between ensembles initialised using bred vectors and singular vectors has recently by *Trevisan et al.* (2000) using the intermediate nonlinear quasi-geostrophic model of *Reinhold and Pierrehumbert* (1982) described above. A preliminary study of the predictability properties of the model was made by performing relatively large Monte Carlo simulations to each of 2,500 initial states (on the model attractor). Much smaller ensemble integrations were then constructed using the two types of dynamically constrained perturbation. The study focussed on cases of regime transition. It was found that in general, the bred-vector ensemble provided an average error distribution more similar to the Monte Carlo distribution, whilst the singular-vector ensemble provided a more reliable estimate of an upper bound on error growth. For example, a prediction of a low probability of transition of weather regime was found to be much more reliable using the singular vector perturbations than using the bredvector perturbations in these dynamically-constrained ensembles.

This is of some relevance for user confidence in probability forecasts. As discussed in more detail in Section 9, consider a user who could potentially suffer a catastrophic loss L if an event E occurs (in this case a circulation regime change bringing severe weather of some type). The user can take protective action at cost C, which might itself be substantial, but can be presumed to be less than L. Based on the ensemble forecast, one can estimate a probability p that E will occur. If p is sufficiently low, then the user should be able to assume that protective action is not necessary. On these occasions of very low p, the incurrence of a loss L would be seen as a failure of the forecast system, and user confidence in the system would be seriously compromised. Of course, an a posteriori constant offset could be added to the forecast probability, so that very small probabilities would never be issued. But in this case, the ensemble would have no realistic value to the user as a decision tool, since the issued probabilities would always be sufficiently high that protective action would always be taken.

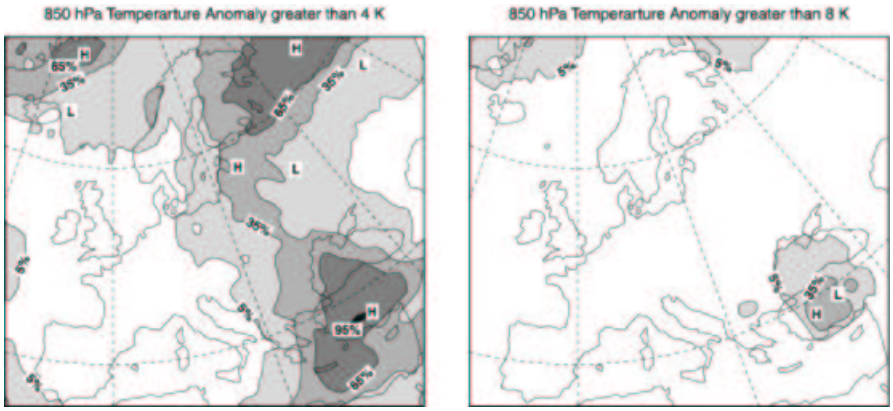

Figure 8: An example of an operational forecast product from the 50-member ECMWF ensemble prediction system. Probability that (a) 850 hPa temperature is at least 4°C above normal, (b) 8°C above normal, based on a 6-day forecast.

The implication of this analysis is that it is important, arguably paramount, to be sure that the user can be confident in a forecast where the ensemble spread is small. For this reason, it is important in an operational environment to ensure that the initial perturbations, are not conservative in the sense of not spanning phase space directions where the pdf is underestimated (due to unquantified errors in the data analysis procedures), and where forecast error growth could be large because of dynamical instability. For these reasons, the ECMWF ensemble system is based in part on initial perturbations using rapidly growing singular vectors (with approximations to the g_{A-1} metric). Examples of the beneficial impact of singular vector perturbations over other types of perturbation are given in *Mureau et al.* (1993) and *Gelaro et al.* (1998).

At the time of writing, the ECMWF Ensemble Prediction System (EPS) comprises 51 forecasts of the ECMWF forecast model (see *Buizza et al.*, 1998: *Buizza et al.*, 1999b). As above, the control forecast is run from the operational ECMWF analysis. The 50 perturbed initial conditions are made by adding and subtracting linear combinations of the dominant 25 singular vector perturbations to the operational analysis. In addition, the 50 perturbed forecasts are also run with the stochastic parametrisation defined in Eq. (25). The EPS is run every day and basic meteorological products are disseminated to all the national meteorological services of the ECMWF Member States. These products often take the form of probability forecasts for different events E, based on the fraction of ensemble members for which E is forecasts. For example, Fig. 8 shows day 6 probability forecasts over Europe for the events $E_{>8}$, $E_{>4}$ defined as: difference of lower tropospheric temperature from a long-term climatology is greater than 8°C, or 4°C, respectively.

Based on the EPS, many of the European national meteorological services are providing their customers detailed probability forecasts of possible weather parameters for site-specific locations. For example, based on the EPS, the United Kingdom Meteorological Office provides forecast pdfs of surface temperature and precipitation for various specific locations in the UK. Joint probability distributions are also estimated (e.g. the probability of precipitation with surface temperature near or below freezing).

As discussed above (cf. Fig. 2), a feature of nonlinear dynamical systems is the dependence of error growth on initial state. Fig. 9 illustrates this, based on two 10-day EPS integrations from starting dates exactly one year apart, the meteorological forecast variable being surface temperature over London. The unperturbed control forecast and verification are also shown. In the first example, the growth of the initial perturbations is relatively modest, and the associated forecast temperature pdf is relatively sharp. In the second example, the initial perturbations grow rapidly and the associated forecast temperature pdf is broad. Notice in the second example that the control integration is already very unskilful 3 days into the forecast. The large ensemble spread provides an a priori warning to the user

that decisions made from single deterministic forecasts during this period could be extremely misleading.

lthough forecasters have traditionally viewed weather prediction as deterministic, a culture change towards probabilistic forecasting is in progress. On the other hand, it is still necessary to demonstrate to users of weather forecasts that reliable probability forecasts provide greater value than imperfect deterministic forecasts. Such a demonstration will be given in Section 9 below.

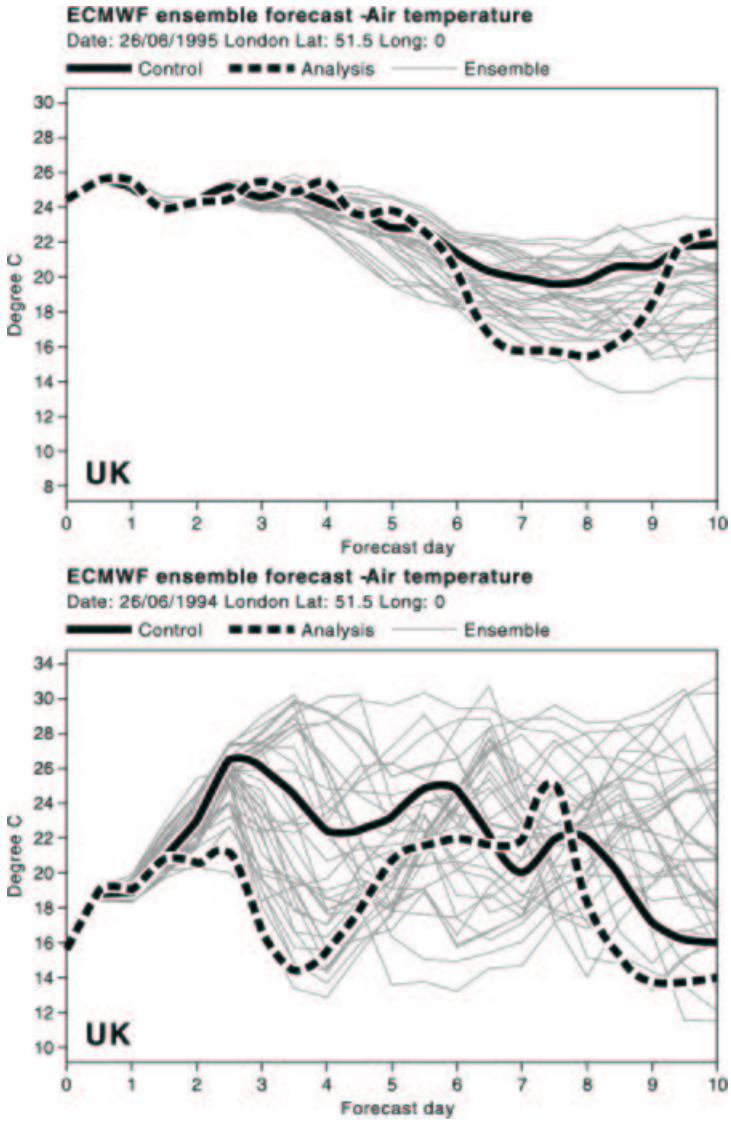

Figure 9: Time series of ensemble forecast integrations for surface temperature over London from starting conditions exactly 1 year apart. The unperturbed control forecast (heavy solid) and the verifying analysis (heavy dashed) are also shown. Top: a relatively predictable period. Bottom: an unpredictable period. (In these examples, the ensembles comprised 32 perturbed forecasts and the model did not include the stochastic representation as described in Eq. (25).)

7.2 Seasonal to interannual prediction

After the major 1997/98 El Niño, event (e.g. *Webster and Palmer*, 1998), the media and public alike have been fascinated by El Niño; indeed, almost every unusual weather event during that period was attributed to El Niño.

However, in addition to this explosion of general interest, there has, in the last few years, been a major development in our ability to predict El Niño and its global consequences. As a result, forecasts of the fluctuations of the climate system on seasonal to interannual timescales have matured to the stage of being produced routinely by operational forecast centres around the world (see http://www.iges.org/ellfb).

The term "El Niño" was coined by Peruvian fisherman to describe a seasonal warming of coastal waters, but was adopted by scientists to denote a fluctuation of the large-scale ocean atmosphere temperature and circulation field across the equatorial Pacific (*Philander* 1990). During normal years, equatorial Pacific sea surface temperatures (SSTs) along the equator are coldest in the east Pacific and warmest in the west Pacific. From a meteorological point of view, the release of energy from convective instability over these warmer waters drives the large-scale circulations in the tropical atmosphere including the surface trade winds which blow from east to west across the equatorial Pacific. From an oceanographic point of view, these westward (or "easterly") trade winds maintain the SST gradients across the. equatorial Pacific. In part, the trade winds "pile up" warm surface waters in the Western Pacific. In addition, due to the Coriolis effect, the trade winds cause an upwelling of colder subsurface water to the surface at the equator; this effect being strongest in the east Pacific (*Gill* 1982).

From this point of view, the ocean and atmosphere are in some sort of coupled equilibrium. However, this is not a completely stable equilibrium. Consider a weakening of the trade winds locally in the central equatorial Pacific. The effect of this perturbation would be to weaken the upwelling of cold water and hence to locally increase SST. The influence of this perturbation in the ocean can propagate eastwards and westwards, through oceanic Kelvin and Rossby modes respectively (*Gill* 1982). The faster-propagating Kelvin wave spreads the influence of this anomalous downwelling perturbation to the equatorial eastern Pacific, which leads to a weakening of the overall east-west gradient of SST. As a consequence, the convergence of moisture and consequent release of convective instability over the west Pacific will also weaken. The result is a further reduction in the strength of the trade winds. The amplification of this feedback can lead to an evolution of the ocean-atmosphere system to a state where both the trade winds and the east-west SST gradient are almost completely eliminated. This corresponds to a (major) EI Niño state.

However, this EI Niño state is not in a stable equilibrium either. The westward propagating oceanic Rossby mode, forced by the relaxed trade winds, reflects off the Indonesian archipelago and returns as a Kelvin mode restoring the ocean back to its normal state, or even to a overrelaxed "La Niña" state in which the east-west SST gradients are enhanced. It is these propagating oceanic Kelvin/Rossby modes which are believed to make the EI Niño event predictable several seasons ahead. The first coupled-model predictions of El Niño were performed using intermediate models of the Pacific ocean coupled to relatively simple atmospheric representations (*Zebiak and Cane*, 1987).

The basic dynamics and predictability of EI Niño can be qualitatively characterised by the ("toy-model") differential-delay equation (*Schopf and Suarez*, 1988; *Battisti and Hirst*, 1989)

$$\dot{T} = cT - bT(t-\tau) \tag{46}$$

where T denotes the SST anomaly in the eastern equatorial Pacific. The first term on the right-hand side of Eq. (46) represents the positive feedback between the strength of the trade winds and the strength of the SST gradient across the Pacific. The second term is a delayed negative feedback representing the adjustment processes in the ocean associated with wave dynamics. The delay τ is determined by the time it takes the Rossby wave to propagate westward across the Pacific basin, reflect off the Indonesian archipelago, and return as an eastward propagating Kelvin wave.

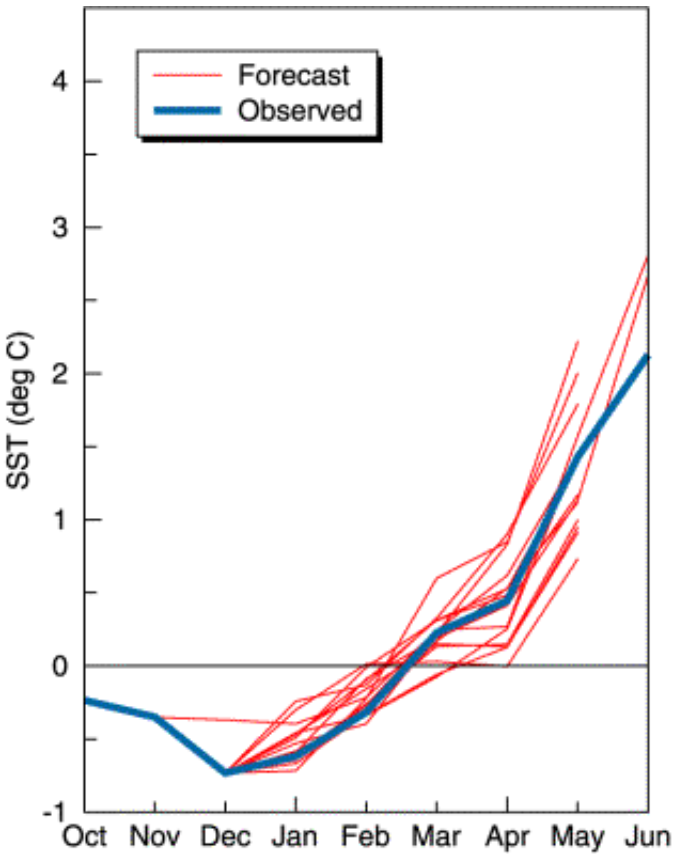

Figure 10: Time series of ensemble seasonal forecasts of monthly mean sea surface temperature anomalies predicted for the region of the tropical Pacific where the El Niño signal is strongest. Forecasts of the ECMWF coupled model are run over 6-month periods from consecutive days. The heavy line shows the observed sea surface temperatures. From the ECMWF seasonal prediction system (see http://www.ecmwf.int/seasonal).

El Niño influences atmospheric circulations around the globe. However, it is important to realise that these circulations are essentially chaotic. In this sense the impact of El Niño on global atmospheric circulations is not itself predictable. An analogy is provided by the *Lorenz* (1963) model, in which an additional prescribed forcing term is added (see Eq. (45)). With or without forcing, the timeseries of the Lorenz model state vector is chaotic. On the other hand, the probability of the state vector lying in one of the two Lorenz-attractor regimes is predictably altered by the imposed forcing. This analogy emphasises the essential point that reliable seasonal forecasts of the influence of El Niño on the global atmospheric circulation are necessarily probabilistic.

The successful prediction of El Niño events by intermediate complexity coupled ocean-atmosphere models, strongly motivated the development of operational forecast systems using comprehensive global coupled ocean-atmosphere models. The 1997/98 event was successfully forecast by a number of such models, including the ECMWF model (*Stockdale et al.*, 1998) which comprised the weather forecast model described above, coupled to a global ocean model developed at the Max Planck Institute for Meteorology (*Latif et al.* 1994).

Predicting seasonal climatic fluctuations is essentially an initial value problem; however, unlike the weather prediction problem, predictability arises from a memory of initial conditions primarily in the ocean. Hence, ocean data assimilation schemes, similar to that described in Section 3, are used to assimilate relevant ocean observations. (It should be noted that the land-surface can also impart seasonal-timescale predictability on the atmosphere; for example, the strength of the Indian summer monsoon is known to be influenced by snow amount accumulated over

the Asian continent during the previous winter; *Blanford*, 1884). In principle, the same techniques used for forecasting uncertainty in weather prediction can be used for seasonal prediction. For example, at ECMWF, an analysis of the ocean and atmosphere is performed every day, and a single 6-month integration of the global coupled ocean-atmosphere model is made from each analysis (see http://www.ecmwf.int/seasonal). Taken over a month, the set of such forecasts comprises an approximately 30-member "time-lagged" ensemble (*Hoffman and Kalnay*, 1983). Fig. 10 shows an example of such an ensemble of coupled model forecasts made just before the onset of the 1997/98 El Niño. The forecast variable is the mean SST anomaly in the central to eastern tropical Pacific. The ensemble shows a significant probability of the development of an El Niño event.

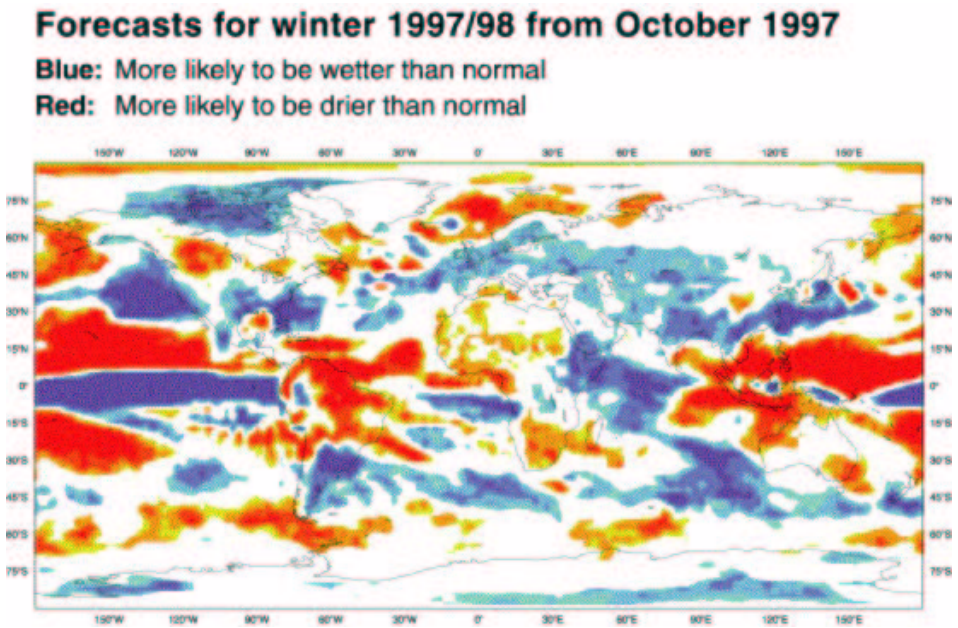

Figure 11: Probability that seasonal-mean precipitation is above or below normal based on a 30-member ensemble of 6-month integrations of the ECMWF seasonal forecast system for the boreal winter 1997/98 starting in October 1997. Dotted shading indicates a higher than normal probability of above average precipitation, the denser the dots, the higher the probability. Solid shading indicates a higher than normal probability of below average precipitation, the deeper the shading, the higher the probability. Regions shaded white indicate that the forecast probabilities are not significantly different to climatology. From the ECMWF seasonal prediction system (see http://www.ecmwf.int/seasonal for more details).

As mentioned, the effect of El Niño is global. Fig. 11 shows a probability forecast of precipitation, based on an ensemble of ECMWF seasonal integrations. The density of shading gives the probability that seasonal mean precipitation (for the northern winter 1997/98) will be above (dots) or below (solid) average. In the regions shown white, the influence of the anomalous boundary forcing is small (statistically insignificant) compared with chaotic variability arising from internal atmospheric fluctuations. For example, according to the forecast in Fig. 11 , there is a high probability of relatively dry conditions over Amazonia and Indonesia, and wet conditions over California across to Florida. Such conditions actually occurred and featured extensively in media coverage of weather anomalies during that winter.

At present, it is unclear what the ultimate limit of predictability of El Niño really is. Whether this limit is controlled purely by the internal nonlinear dynamics of El Niño (e.g. as represented by the intermediate Zebiak-Cane model; e.g. *Münnich et al.*, 1991), or by unpredictable forcing from atmospheric motions external to the coupled system (e.g. *Eckert and Latif*, 1997), remains to be established.

7.3 Decadal prediction and anthropogenic climate change

It is well known that anthropogenic increases of CO_2 have the propensity to change earth's climate. Even though climate is almost certain to warm on a global scale, regional changes in climate can only be predicted using the global coupled ocean-atmosphere models discussed above. But how confident can we be in such predictions? What are the principal uncertainties?

The problem of predicting climate change arising from anthropogenic perturbations has a different character to the weather or seasonal forecasts discussed above. In a general sense, anthropogenic climate-change prediction could be described as a problem of estimating how the climate attractor changes for a prescribed increase in atmospheric CO_2 content. A conceptually similar problem was discussed in Section 5, how the climate responds to slow changes in the orbit of the earth about the sun. As discussed in the introduction, the climate-change prediction problem, like the paleoclimate problem, is a prediction of the second kind.

In practice, this delineation is perhaps simplistic; climate-change scenario experiments are made by running climate models over the coming decades with slowly increasing CO_2. On these timescales there is in fact some evidence that natural fluctuations of the Physical climate system are partially predictable. For example, anomalies of ocean salinity and SST have been observed to propagate coherently for many years through high latitudes of the North Atlantic (*Dickson et al.*, 1988, *Belkin et al.*, 1998, *Hansen and Bezdek*, 1996; *Sutton and Allen*, 1997). It has been suggested that such anomalies may influence and/or be influenced by the global thermohaline circulation (*Dickson et al.*, 1996), a circulation which affects the mean climate over Northern Europe. Ensemble integrations of coupled ocean-atmosphere models suggest that the thermohaline circulation has some predictability on decadal timescales (*Griffies and Bryan*, 1997). There may be other ocean- atmosphere modes with decadal predictability (*Latif*, 1998). In addition, the climate system may have biochemical feedbacks which are partially predictable on decadal timescales (*Pielke*, 1999).

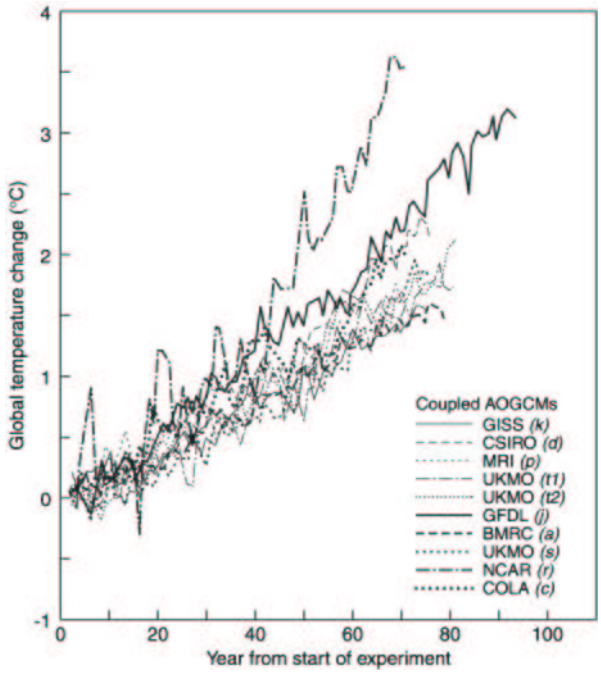

Figure 12: A multi-model ensemble of predictions of anthropogenic global temperature change, based on a prescribed increase in atmospheric CO_2 content. From *IPCC* (1996).

Let us suppose that such decadal predictability exists. Then, for example, it might be possible to estimate from en-

semble integrations initialised with current data, that there is a significant probability (e.g. significantly different than 0.5) that the thermohaline circulation will slow down in the next century as a result of natural climate variability. In this situation, the anthropogenic climate change problem becomes one of knowing how anthropogenic forcing would further influence this probability. However, given our limited knowledge of the current state of the climate system (especially the deep ocean), and the limited natural predictability that may in fact exist on decadal to centennial timescales, the problem of predicting the impact of anthropogenic forcing on the thermohaline circulation of the 21st century will in practice be a pure prediction of the second kind (e.g. as studied by *Manabe and Stouffer*, 1994).

A principal uncertainty in climate change prediction lies in the uncertainties in model formulation, e.g as discussed in Section 4. As discussed in Section 4, such uncertainty may necessarily imply some stochastic representation of the influence of sub-grid scale processes within the model formulation. At present, such stochastic representations have not been used in the context of climate simulation. Rather, the representation of model uncertainty in ensemble integrations has been based on the multi-model ensemble (cf. Section 4).

Multi-model ensemble prediction of anthropogenic climate change is a central feature of climate change assessment by the Intergovernmental Panel on Climate Change (IPCC). (IPCC was jointly established by the World Meteorological Organisation and the United Nations Environment Programme to assess available scientific information on climate change, to assess the environmental and socio-economic impacts of climate change, and to formulate response strategies.) For example, Fig. 12 shows a multi-model ensemble of climate change predictions, based on a 1%/year increase of atmospheric CO_2 concentration from 1990 (*IPCC*, 1996). The ensemble suggests there is no substantial probability of global warming being less that 0.25ºC, or more than 4ºC, 60 years from now.

On the other hand, the pdf spanned by the use of the different sets of models is ad hoc, and may not cover all of the uncertainties in model formulation. Moreover, in practice the models are not truly independent of one another; there is a tendency (possibly for pragmatic rather than scientific reasons) for groups to adopt similar sets of parametrisations. In addition, most of the models used in *IPCC* (1996) suffer from similar shortcomings; for example, none had sufficient equatorial resolution to simulate correctly the dynamics of El Niño. EI Niño does have an impact on the largest of atmospheric scales; for example, during El Niño years, global mean temperatures are enhanced. Hence if the frequency of EI Niño events is affected by increasing CO_2, this will impact on estimates of global warming. The effects of increasing CO_2 on EI Niño are now beginning to be studied in climate models (e.g. *Timmermann et al.*, 1999), though there is still uncertainty in the overall response.

Compared with weather and seasonal forecasts, where verifying data is reasonably plentiful, it is difficult to assess quantitatively the predictions of uncertainty in forecasts of climate change. Certainly, a model's ability to predict natural climate fluctuations on seasonal timescales is increasingly being taken as a prerequisite for a model to be useful for climate-change prediction. The *IPCC* (1990) report made this point explicitly: "Confidence in a model used for climate simulation will therefore be increased if the same model is successful when used in a forecasting mode". This idea is consistent with the paradigm for climate change discussed above (*Palmer*, 1993b, 1999) which emphasises the notion that anthropogenic forcing will change the pdfs of naturally-occurring regimes of climate variability. This paradigm puts a premium on a model's ability to simulate such natural variability.

Hence, albeit indirectly, the validation of forecasts of natural fluctuations in the climate system is relevant to the validation of climate-change scenarios. However, as we have emphasised, such validation techniques must take into account the probabilistic nature of such prediction. The next section deals explicitly with two commonly - used probabilistic validation techniques.

8. VERIFYING FORECASTS OF UNCERTAINTY

As discussed, the output from an ensemble forecast can be used to construct a probabilistic prediction. In this section, we discuss two basic measures of skill for assessing a probability forecast: the Brier Score and the Relative Operating Characteristic. Both of these measures are based on the skill of probabilistic forecasts of a binary event E, as discussed in Section 7 above. For example E could be: temperatures will fall below 0ºC in three days time; average rainfall for the next three months will be at least one standard deviation below normal; seasonal-mean rainfall will be below average and temperature above average, and so on.

8.1 The Brier score and its decomposition

Consider an event E which, for a particular ensemble forecast, occurs a fraction p of times within the ensemble. If E actually occurred then let $v = 1$, otherwise $v = 0$. Repeat this over a sample of N different ensemble forecasts, so that p_i is the probability of E in the i th ensemble forecast and $v_i = 1$ or $v_i = 0$, depending on whether E occurred or not in the i th verification ($i = 1, 2 \ldots N$).

The Brier score (*Wilks*, 1995) is defined by

$$b = \frac{1}{N}\sum_{i=1}^{N}(p_i - v_i)^2, \qquad 0 \le p_i \le 1, v_i \in \{0, 1\} \tag{47}$$

From its definition $0 \le b \le 1$, equalling zero only in the ideal limit of a perfect deterministic forecast. For a large enough sample, the Brier score can be written as

$$b = \int_0^1 [p-1]^2 o(p)\rho_{ens}(p)\mathrm{d}p + \int_0^1 p^2[1-o(p)]\rho_{ens}(p)\mathrm{d}p \tag{48}$$

where $\rho_{ens}(p)\mathrm{d}p$ is the relative frequency that E was forecast with probability between p and $p + \mathrm{d}p$, and $o(p)$ gives the proportion of such cases when E actually occurred. To see the relationship between (47) and (48) note that $\int_0^1 [p-1]^2 o(p)\rho_{ens}(p)\mathrm{d}p$ is the Brier score for ensembles where E actually occurred, and $\int_0^1 [p-0]^2 o(p)\rho_{ens}(p)\mathrm{d}p$ is the Brier score for ensembles where E did not occur.

Simple algebra on (48) gives *Murphy*'s (1973) decomposition

$$b = b_{rel} - b_{res} + b_{unc} \tag{49}$$

of the Brier score, where

$$b_{rel} = \int_0^1 [p - o(p)]^2 \rho_{ens}(p)\mathrm{d}p \tag{50}$$

is the reliability component,

$$b_{res} = \int_0^1 [\bar{o} - o(p)]^2 \rho_{ens}(p)\mathrm{d}p \tag{51}$$

is the resolution component

$$b_{unc} = \bar{o}[1-\bar{o}] \tag{52}$$

is the uncertainty component, and

$$\bar{o} = \int_0^1 o(p)\rho_{ens}(p)\mathrm{d}p \tag{53}$$

is the (sample) climatological frequency of E.

A reliability diagram (*Wilks*, 1995) is one in which $o(p)$ is plotted against p for some finite binning of width δp. In a perfectly reliable system $o(p) = p$ and the graph is a straight line oriented at 45° to the axes, and $b_{rel} = 0$. Reliability measures the mean square distance of the graph of $o(p)$ to the diagonal line.

Resolution measures the mean square distance of the graph of $o(p)$ to the sample climate horizontal line. A system with relatively high b_{res} is one where the dispersion of $o(p)$ about is as large as possible. Conversely, a forecast system has no resolution when, for all forecast probabilities, the event verifies a fraction $o(p) = \bar{o}$ times.

The term b_{unc} on the right-hand side of (49) ranges from 0 to 0.25. If E was either so common, or so rare, that it either always occurred or never occurred within the sample of years studied, then $b_{unc} = 0$; conversely if E occurred 50% of the time within the sample, then $b_{unc} = 0.25$. Uncertainty is a function of the climatological frequency of E, and is not dependent on the forecasting system itself. It can be shown that the resolution of a perfect deterministic system is equal to the uncertainty.

When assessing the skill of a forecast system, it is often desirable to compare it with the skill of a forecast where the climatological probability $\bar{o}$ is always predicted (so $\rho_{ens}(p) = \delta(p - o)$). The Brier score of such a climatological forecast is $b_{cli} = b_{unc}$ (using the sample climate), since, for such a climatological forecast $b_{rel} = b_{res} = 0$. In terms of this, the Brier skill score, B, of a given forecast system is defined by

$$B = 1 - \frac{b}{b_{cli}} \tag{54}$$

$B \leq 0$ for a forecast no better than climatology, and $B = 1$ for a perfect deterministic forecast.

Skill-score definitions can similarly be given for reliability and resolution, i.e.

$$\left[B_{rel} = 1 - b_{rel}/b_{cli}\right] \tag{55}$$

$$\left[B_{res} = b_{res}/b_{cli}\right] \tag{56}$$

For a perfect deterministic forecast system, $B_{rel} = B_{res} = 1$. Hence, from Eqs. (49) and (54)

$$B = B_{res} + B_{rel} - 1 \tag{57}$$

Fig. 13 shows two examples of reliability diagrams for the ECMWF EPS taken over all day-6 forecasts from December 1998 - February 1999 over Europe (cf. Fig. 8). The events are $E_{>4}$, $E_{>8}$:- lower tropospheric temperature being at least 4°C, 8°C greater than normal. The Brier score, Brier skill score, and Murphy decomposition are shown on the figure.

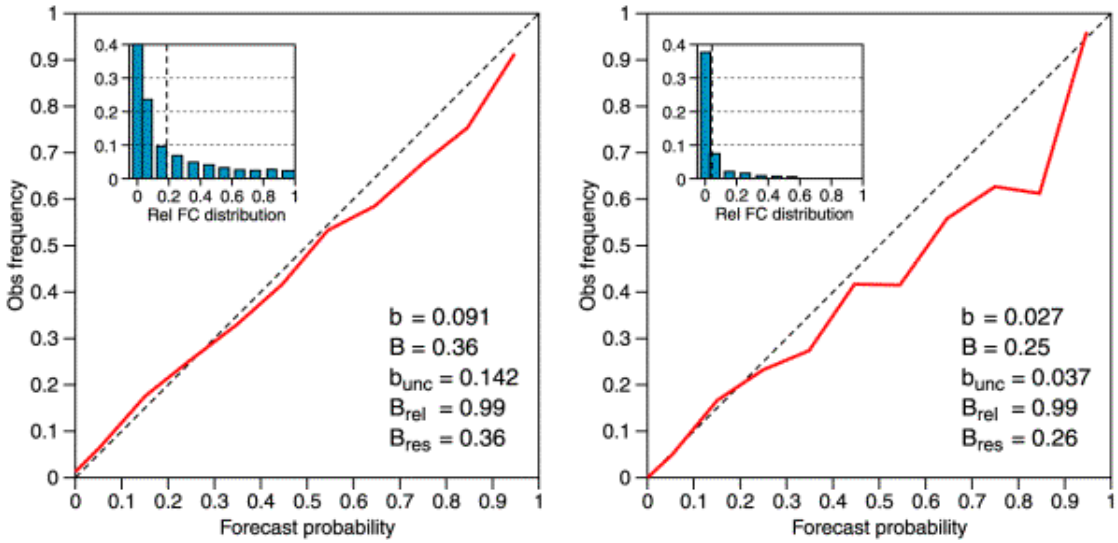

Figure 13. Reliability diagram and related Brier score skill score and Murphy decomposition for the events: (a) 850 hPa temperature is at least 4°C above normal and (b) at least 8°C above normal, based on 6-day forecasts over Europe from the 50-member ECMWF ensemble prediction system from December 1998 - February 1999. Also shown is the pdf $\rho_{ens}(p)$ for the event in question.

The reliability skill score B_{rel} is extremely high for both events. However, the reliability diagrams indicate some overconfidence in the forecasts. For example, on those occasions where $E_{>4}$ was forecast with a probability between 80% and 90% of occasions, the event only verified about 72% of the time. However, it should be remembered that the integrand in Eq. (50) is weighted by the pdf $\rho_{ens}(p)$, shown in each reliability diagram. In both cases, forecasts where $p > 0.4$ are relatively rare and hence contribute little to B_{rel}.

To see why probability forecasts of $E_{>4}$ have higher Brier skill scores than probability forecasts of $E_{>8}$, consider Eq. (57). From Fig. 13 , whilst B_{rel} is the same for both events, B_{res} is larger for $E_{>4}$ than for $E_{>8}$. This can be seen by comparing the histograms of $\rho_{ens}(p)$ in Fig. 13 which are more highly peaked for $E_{>8}$ than for $E_{>4}$; there is less dispersion of the probability forecasts of the more extreme event about its climatological frequency, than the equivalent probability forecasts of the more moderate event. This is hardly surprising; the more extreme event $E_{>8}$ is relatively rare (its climatological frequency is ~ 0.04) and most of the time is forecast with probabilities which almost always lie in the first probability category ($0 \le \delta p \le 0.1$). In order to increase the Brier score of this relatively extreme event, one would need to increase the ensemble size so that finer probability categories can be reliably defined. (For example, suppose an extreme event has a climatological probability of occurrence of p_{rare} Let us suppose that we want to be able to forecast probabilities of this event which can discriminate between probability categories with a band width comparable with this climatological frequency, then the ensemble size S_{ens} should be $>>(1/p_{rare})$.) With finer probability categories, the resolution component of the Brier score can be expected to increase. Providing reliability is not compromised, this will lead to higher overall skill scores.

However, this raises a fundamental dilemma in ensemble forecasting given current computer resources. It would be meaningless to increase ensemble size by degrading the model (e.g. in terms of "physical" resolution) making it cheaper to run, if by doing so it could no longer simulate extreme weather events. Optimising computer resources that on the one hand ensemble sizes are sufficiently large to give reliable probability forecasts of extreme but rare events, and that on the other hand the basic model has sufficient complexity to be able to simulate such events, is a very difficult balance to define.

The Brier score and its decomposition provide powerful objective tools for comparing the performance of different

probabilistic forecast systems. However, the Brier score itself does not address the issue of whether a useful level of skill has been achieved by the probabilistic forecast system. In order to prepare the ground for a diagnostic of probabilistic forecast performance which determines potential economic value, we first introduce skill score originally derived from signal detection theory.

8.2 Relative operating characteristic

The relative operating characteristic (ROC; *Swets*, 1973; *Mason* 1982; *Harvey et al.*, 1992) is based on the forecast assumption that E will occur, providing E is forecast by at least a fraction $p = p_t$ of ensemble members, where the threshold p_t is defined a priori by the user. As discussed below, optimal p_t can be determined by the parameters of a simple decision model.

Consider first a deterministic forecast system. Over a sufficiently large sample of independent forecasts, we can form the forecast- model contingency matrix giving the frequency that E occurred or did not occur, given it was forecast or not forecast, i.e.

		Occurs	
		No	Yes
Forecast	No	α	β
	Yes	γ	δ

Based on these values, the so-called "hit rat" (H) and "false-alarm rate" (F) for E are given by

$$\begin{aligned} H &= \delta/(\beta+\delta) \\ F &= \gamma/(\alpha+\gamma) \end{aligned} \tag{58}$$

Hit and false alarm rates for all ensemble forecast can be defined as follows. It is assumed that E will occur if $p \geq p_t$ (and will not occur if $p < p_t$). By varying p_t between 0 and 1 we can define $H = H(p_t)$, $F = F(p_t)$. In terms of the pdf $\rho_{ens}(p)$

$$\begin{aligned} H(p_t) &= \int o(p)\rho_{ens}(p)\mathrm{d}p/\bar{o} \\ F(p_t) &= \int \{1-o(p)\}\rho_{ens}(p)\mathrm{d}p/(1-\bar{o}) \end{aligned} \tag{59}$$

The ROC curve is a plot of $H(p_t)$ against $F(p_t)$ for $0 \leq p_t \leq 1$. A measure of skill is given by the area under the ROC curve (A). A perfect deterministic forecast has $A = 1$, whilst a no-skill forecast for which the hit and false alarm rates are equal, has $A = 0.5$.

Relative operating characteristic curve for seasonal timescale integrations (run over the years 1979-93, run with prescribed observed SST) for the event $E_{<0}$:- the seasonal-mean (December-February) 850 hPa temperature anomaly is below normal. Solid: based on a single model 9-member ensemble. bottom: based on a multi-model 36-member ensemble (see *Palmer et al.*, 2000) for more details. The area A under the two curves (a measure of skill) is shown.

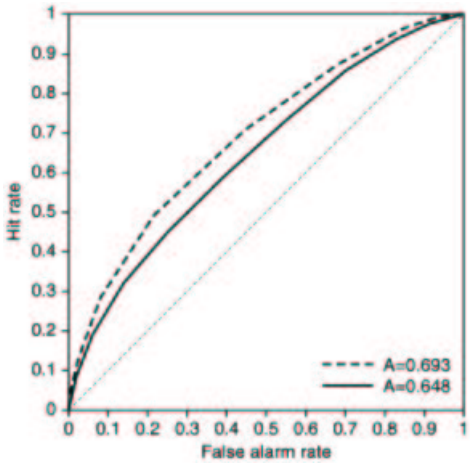

Figure 14. Relative operating characteristic curve for seasonal timescale integrations (run over the years 1979-93, run with prescribed observed SST) for the event $E_{<0}$:- the seasonal-mean (December-February) 850 hPa temperature anomaly is below normal. Solid: based on a single model 9-member ensemble. bottom: based on a multi-model 36-member ensemble (see *Palmer et al.*, 2000) for more details. The area *A* under the two curves (a measure of skill) is shown.

We illustrate in Fig. 14 the application of these measures of skill to a set of multi-model multi-initial condition ensemble integrations made over the seasonal timescale (*Palmer et al.*, 2000). The event being forecast is $E_{<0}$:- the seasonal-mean (December-February) 850 hPa temperature anomaly will be below normal. The global climate models used in the ensemble are the ECMWF model, the UK Meteorological Office Unified Model, and two versions of the French Arpège model; the integrations were made as part of the European Union "Prediction of Climate Variations on Seasonal to Interannual Timescales (PROVOST)". For each of these models, 9-member ensembles were run over the boreal winter season for the period 1979-1993 using observed specified SSTs. The values $H(p_t)$ and $F(p_t)$ have been estimated from probability bins of width 0.1. The ROC curve and corresponding *A* value is shown for the 9-member ECMWF model ensemble, and for the 36-member multi-model ensemble. It can be seen that in both cases, *A* is greater than the no-skill value of 0.5; however, the multi-model ensemble is more skilful than the ECMWF-model ensemble. Studies have shown that the higher skill of the multi-model ensemble arises mainly because of the larger ensemble size, but also because of a sampling of the pdf associated with model uncertainty.

9. The economic value of predicting uncertainty

Although *B* and *A* (defined in Section 8) provide objective measures of skill for ensemble forecasts, they do not determine measures of usefulness for seasonal forecasts. In an attempt to define this notion objectively, we consider here a simple decision model (*Murphy*, 1977; *Katz and Murphy*, 1997) whose inputs are probabilistic forecast information and whose output is potential economic value.

Consider a potential forecast user who can take some specific precautionary action depending on the likelihood that *E* will occur. Let us take some simple examples relevant to seasonal forecasting. If the coming winter is mild (*E*:- seasonal-mean temperature above normal), then overwintering crops may be destroyed by aphid growth. A farmer can take precautionary action by spraying. If the growing season is particularly dry (*E*:- seasonal-mean rainfall at least one standard deviation below normal), then crops may be destroyed by drought. A farmer can take precautionary action by planting drought-resistant crops. In both cases taking precautionary action incurs a cost *C* irre-

spective of whether or not E occurs (cost of spraying, or cost associated with reduced yield and possibly with more expensive seed). However, if E occurs and no precautionary action has been taken, then a loss L is incurred (crop failure).

This simple "cost-loss" analysis is also applicable to much shorter range forecast problems (Richardson, 1999). For example, if the weather event was the occurrence of freezing conditions leading to ice on roads, and the precautionary action was to salt the roads, then C would correspond to the cost of salting, and L would be associated with the increase in road accidents, traffic delays etc.

In general, the expense associated with each combination of ac action /inaction and occurrence/non-occurrence of E is given in the decision-model contingency matrix

		Occurs	
		No	Yes
Take action	No	0	L
	Yes	C	C

It is presumed that the decision maker wishes wishes to maximise profits, or at least minimise overall expenses.

If only the climatological frequency $\bar{o}$ of E is known, there are two basic options: either always or never take precautionary action. Always taking action incurs a cost C on each occasion, whilst never taking action incurs a loss L only on the proportion $\bar{o}$ of occasions when E occurs, giving an expense $\bar{o}L$.

If the seasonal forecast data used in Section 8 above were used by a hypothetical decision maker, would his/her expenses would be reduced beyond what could be achieved using $\bar{o}$ alone? Consider first a deterministic forecast system with characteristics described by the forecast-model contingency matrix in Section 8.2. The user's expected mean expense M per unit loss is

$$M = \frac{\beta L + (\gamma + \beta)C}{L} \tag{60}$$

This can be written in terms of the hit-rate H and the false-alarm F using (58), so that

$$M = F\frac{C}{L}(1-\bar{o}) - H\bar{o}\left(1-\frac{C}{L}\right) + \bar{o} \tag{61}$$

For a perfect deterministic forecast $H = 1$ and $F = 0$, hence

$$M_{\text{per}} = \bar{o}\frac{C}{L} \tag{62}$$

To calculate the mean expense per unit loss knowing only $\bar{o}$, suppose first the decision maker always protects, then $M = C/L$. Conversely, if the decision maker never protects then $M = \bar{o}$. Hence if the decision maker knows only $\bar{o}$, M can be minimised by either always or never taking precautionary action, depending on whether $C/L < \bar{o}$, or $C/L > \bar{o}$, respectively. The mean expense per unit loss associated with a knowledge of climatology only is therefore

$$M_{\mathrm{cli}} = \min\left(\frac{C}{L}, \bar{o}\right) \tag{63}$$

The value V of forecast information is defined as a measure of the reduction in M over M_{cli}, normalised by the maximum possible reduction associated with. a perfect deterministic forecast, i.e.

$$V = \frac{M_{\mathrm{cli}} - M}{M_{\mathrm{cli}} - M_{\mathrm{per}}} \tag{64}$$

For a forecast system which is no better than climate, $V = 0$; for a perfect deterministic forecast system $V = 1$.

As discussed in Section 8, an ensemble forecast gives hit and false-alarm rates $H = H(p_t)$ and $F = F(p_t)$, as a functions of probability thresholds p_t (see (59)). Hence V is defined for each p_t, i.e. $V = V(p_t)$. Using (61), (62) and (63)

$$V(p_t) = \frac{\min\left(\frac{C}{L}, \bar{o}\right) - F(p_t)\frac{C}{L}(1-\bar{o}) + H(p_t)\bar{o}\left(1-\frac{C}{L}\right) - \bar{o}}{\min\left(\frac{C}{L}, \bar{o}\right) - \bar{o}\frac{C}{L}} \tag{65}$$

For given C/L and event E, the optimal value is

$$V_{\mathrm{opt}} = \max_{p_t}[V(p_t)] \tag{66}$$

Figs. 15 -16 show examples of optimal value as a function of user cost/loss ratio. Fig. 15 is for the ECMWF day-6 ensemble weather prediction system and the event $E_{<4}$ (as in Fig. 13). The solid curve is the optimal value for the ensemble system, the dashed curve shows value for a single deterministic forecast (the unperturbed "control" integration in the ensemble). Peak value tends to occur for $C/L \square \bar{o}$; for such users, it makes little difference to the climatological expense M_{cli} whether they always protect, or never protect. The figure also illustrates the fact that the ensemble forecast has greater "value" than a single deterministic forecast. For some cost/loss ratios (e.g. $C/L > 0.6$), the deterministic forecast has no value, whereas the ensemble forecast does have value. The reason for this can be understood in terms of the fact that for a probabilistic forecast, different users (with different C/L) would take precautionary action for different forecast probability thresholds. A user who would suffer a catastrophic loss ($C/L \ll 1$) if E occurred, would take precautionary action even when a small probability of E was forecast. A user for whom precautionary action was expensive in comparison with any loss ($C/L \sim 1$) would take precautionary action only when a relatively large probability of E was forecast. The result demonstrates the value of a reliable probability forecast.

Fig. 16 shows optimal value from the ECMWF seasonal integrations described in section 7. Two different events are shown based on seasonal-mean rainfall forecasts over all regions in the tropics. They are: rainfall less than normal, and rainfall less than one standard deviation below normal (the latter being a possible objective definition of drought). It can be seen that the peak value for the two events occurs at different C/L, arising because the two events have different values of $\bar{o}$. The figure also shows that the value of the forecasts of the more extreme "drought" event is more valuable than the forecasts of rainfall simply below normal. It would appear that the reason for this is that such events tend to be forecast more during El Niño years, when seasonal predictability is high, than during other years, when predictability is lower.

An absolute "dollar" estimate of value can be attached to these curves, providing the user knows the "dollar" value

of their cost and loss. An attempt to do this for seasonal forecasts in southern Africa suggests, for some applications, a potential value of O(10^9 $) (M. Harrison and N. Graham, personal communication).

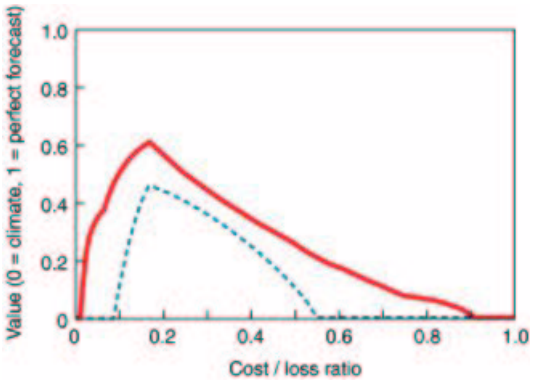

Figure 15: Potential economic value of the ECMWF ensemble prediction system as a function of user cost/loss ratio of day 6 weather forecasts over Europe (for the period December 1998 - February 1999) for the event $E_{<4}$: 850 hPa temperature at least 4°C below normal. Solid: value of the ECMWF ensemble prediction system. Dashed: value of a single deterministic forecast (the unperturbed "control" forecast of the EPS system). From *Richardson* (1999).

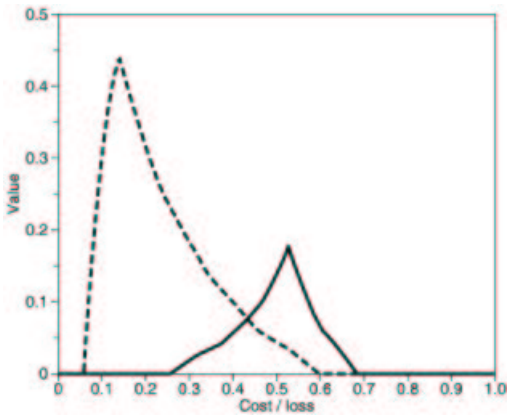

Figure 16: Potential economic value of seasonal ensembles as a function of user cost/loss ratio in the tropics, for the events $E_{<0}$: seasonal-mean rainfall below normal (solid), and $E_{<-\alpha}$: seasonal-mean rainfall at least one standard deviation below normal (dashed), based on 9-member ensembles of integrations with the ECMWF model over the years 1979-93, run with prescribed observed SST.

10. Concluding remarks

Our climate is a complex nonlinear dynamical system, with spatial variability on scales ranging from individual clouds to global circulations in the atmosphere and oceans, and temporal variability ranging from hours to millenia. Climate scientists interact with society through the latter's demands for accurate and detailed environmental forecasts: of weather, of El Niño and its impact on global rainfall patterns, and of man's effect on climate. The complexity of our climate system implies that quantitative predictions can only be made with comprehensive numerical models which encode the relevant laws of dynamics, thermodynamics and chemistry for a multi-constituent multiphase fluid. Typically such models comprise some millions of scalar equations, describing the interaction of circu-

lations on scales ranging from tens of kilometres to tens of thousands of kilometres; from the ocean depth to the upper stratosphere. These equations can only be solved on the world's largest supercomputers.

However, a fundamental question that needs to be addressed, both by producers and users of such forecasts, is the extent to which weather and climate are predictable; after all, much of chaos theory developed from an attempt to demonstrate the limited predictability of atmospheric variations. In the past, the topic of predictability has been a somewhat theoretical and idealised one, somewhat removed from the practicalities of prediction. A famous climatologist remarked some years ago: "Predictability is to prediction as romance is to sex!". However, the remark is perhaps not so apposite today; the science of predictability of weather and climate is now an integral part of the practical prediction problem - the two cannot be separated. The predictability problem can be formulated e.g. through a Liouville equation; however, in practice, estimates of predictability are made from multiple (ensemble) forecasts of comprehensive weather and climate prediction models. The individual members of the ensemble differ by small perturbations to quantities that are not well known. For example, the predictability of weather is largely determined by uncertainty in a forecast's starting conditions, whilst the predictability of climate variations is also influenced by uncertainty in representing computationally the equations that govern climate (for example, how to represent the effects of convective instabilities in a model that cannot resolve individual cloud systems).

Chaos theory implies that all such environmental forecasts must be expressed probabilistically; the laws of physics dictate that society cannot expect arbitrarily accurate weather and climate forecasts. These probability forecasts quantify uncertainty in weather and climate prediction. The duty of the climate scientist is to strive to estimate reliable probabilities; not to disseminate predictions to society with a precision that cannot be justified scientifically. Examples were shown that, in practice, the economic value of a reliable probability forecast (produced from an ensemble prediction system) exceeds the value of a single deterministic forecast with uncertain accuracy.

However, producing reliable probability forecasts from ensembles of climate model integrations puts enormous demands on computer resources. Computer power is needed to be able to resolve the details of the climate system. It has been argued that, because of the nonlinear nature of our climate system, a systematic mistreatment of small-scale phenomena (such as organised convective instabilities in the tropics) could lead to a systematic mistreatment of large-scale climate phenomena (such as EI Niño). However, there are reasons for studying small-scale phenomena in their own right. From a weather prediction perspective, it is necessary to be able to simulate such detail if models are to be able to forecast damaging weather events such as hurricanes and related extratropical systems (such as the one that struck Southern England in October 1987). From a climate point of view it may also be necessary to be able to simulate such detail in order to be able to assess the impact of anthropogenic forcing on the probability distribution of these damaging weather events. Will hurricanes or October storms become more likely or, on average, more intense in an enhanced CO_2 atmosphere?

However, this poses a fundamental dilemma given current computing resources. To be able to simulate extreme events requires models with considerable complexity and resolution. On the other hand, estimating reliably changes to the probability distributions of extreme and hence relatively rare events, requires large ensembles. One thing is certain; the more the need to provide reliable forecasts of uncertainty in our predictions of weather and climate, the more the demand for computer power exceeds availability, notwithstanding the unrelenting advance in computer technology. Indeed the need for quantitative predictions of uncertainty in weather and climate science is a relevant consideration in the design of future generations of supercomputers; ensemble prediction is a perfect application for parallel computing! There can be little doubt that the benefits to society of reliable regional probability forecasts of extreme weather events, seasonal variability and anthropogenic climate change justify such technological developments.

ACKNOWLEDGEMENTS

I would like to thank colleagues at ECMWF, Drs D. Anderson, J. Barkmeijer, C. Brankovic, R. Buizza, M. Ehrendorfer, F. Lalaurette T. Stockdale, and D. Richardson for innumerable discussions and for help in producing many of the figures in this paper. I would also like to thank Drs K. Puri and T. Iversen in particular for help in producing Figs. 4 and 7 , respectively.

REFERENCES

Anderson, J. L., 1997: Impact of dynamical constraints on the selection of initial conditions for ensemble predictions: low-order perfect model results. *Mon. Wea. Rev.*, 125, 2969-2983.

Andersson, E. and M. Fisher, 1999: Background errors for observed quantities and their propagation in time. ECMWF workshop on Diagnosing Data Assimilation Systems. ECMWF Shinfield Park, Reading, Berkshire RG2 9AX, UK.

Barkmeijer, J., M. van Gijzen and F. Bouttier, 1998: Singular vectors and estimates of the analysis error covariance metric. *Q. J. R. Meteorol. Soc*., 124, 1695-1713.

Battisti, D. S. and A. C. Hirst , 1989. Interannual variability in a tropical atmosphere- ocean model. Influence of the basic state, ocean geometry and non-linearity. *J. Atmos. Sci*., 46, 1687-1712.

Belkin, I. M., S. Levitus, J. Antonov, and S.-A. Malmberg, 1998: "Great Salinity Anomalies" in the North Atlantic. Progress in Oceanography, 41, 1-68.

Bengtsson, L., 1999: From short-range barotropic modelling to extended-range global weather prediction: a 40-year perspective. *Tellus*, 51A-B, 13-32.

Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part, II: single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets. *Q. J. R. Meteorol. Soc.*, 112, 693-709.

Blanford, H. F., 1884: On the connexion of Himalayan. snowfall kind seasons of drought in India. *Proc. Roy. Soc. London*, 37, 3-22.

Buizza, R., 1997: Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, 125, 99-119.

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

Buizza, R., T. Petroliagis, T. N. Palmer, J. Barkmeijer, M. Hamrud, A. Hollingsworth, A. Simmons and N. Wedi, 1998: Impact of model resolution and ensemble size on the performance of an ensemble prediction system. *Q. J. R. Meteorol. Soc*., 124, 1935-1960.

Buizza, R., M. J. Miller and T. N. Palmer, 1999a: Stochastic simulation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. So*c., 125, 2887-2908.

Buizza, R., A. Hollingsworth, A. Lalaurette and A. Ghelli, 1999b: Probabilistic predictions of precipitation using the ECMWF ensemble prediction system. *Weather and Forecasting*, 14, 168-189.

Cane, M. and A. Clement 1999: A role for the tropical Pacific coupled ocean-atmosphere system on Milankovic and Millenial timescales. Part II: Global Impacts. American Geophysical Union Monograph: Mechanisms of Millenial-Scale Global Climate Change.

Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteorol*., 4, 135-163.

Chen, Y. Q., D. S. Battisti, T. N. Palmer, J. Barsugli and E. S. Sarachik, 1997: A study of the predictability of tropical Pacific SST in a coupled atmosphere/ocean model using singular vector analysis: the role of the annual cycle and the ENSO cycle. *Mon. Wea. Rev.*, 125, 831845.

Cheng, X. and J. M. Wallace, 1993: Cluster analysis of the Northern Hemisphere wintertime 500hPa height field: spatial patterns. *J. Atmos. Sci*., 50, 2674-2696.

Cho, J. Y. N., Y. Zhu, R. E. Newell, B. E. Anderson, J. D. Barrick, G. L. Gregory, G. W. Sachse, M. A. Carroll and G. M. Albercook, 1999: Horizontal wavenumber spectra of winds, temperature and trave gases during the Pacific Exploratory Missions: 1. Climatology. *J. Geophys. Res*., 104, 5697-5716.

Clement, A. and M. Cane, 1999: A role for the tropical Pacific coupled ocean- atmosphere system on Milankovic and Millenial timescales. Part I: A modeling study of tropical Pacific variability. American Geophysical Union Monograph: Mechanisms of Millenial-Scale Global Climate Change.

Corti, S., F. Molteni and T. N. Palmer, 1999: Signature of recent climate change in frequencies of natural atmospheric circulation regimes. *Nature*, 398, 799-802.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Meteorol. Soc.*, 120, 1367-1387.

Courtier, P., E. Andersson, W. Heckley, R. Pailleux, D. Vasilevic, M. Hamrud, A. Hollingsworth, F. Rabier and M. Fisher, 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Q. J. R. Meteorol. Soc.*, 124, 1783-1808.

Daley, R., 1991: Atmospheric Data Analysis. Cambridge University Press. Cambridge, UK. pp457.

Dickson, R. R., J. Meinke, S. A. Malmberg and A. A. Lee, 1988: The "Great Salinity Anomaly" in the northern North Atlantic, 1968-1982. *Progress in Oceanography*, 20, 103 151.

Dickson, R. R., J. Lazier, J. Meincke and P. Rhines, 1996: Long-term coordinated changes in the convective activity of the North Atlantic. *Progress in Oceanography*, 38(3), 241-295.

Eady, E.T., 1949. Long waves and cyclone waves. *Tellus*, 1, 33-52.

Eckert, C. and M. Latif, 1997: Predictability of a stochastically-forced hybrid coupled model of El Niño. *J. Clim*., 10, 1488-1504.

Ehrendorfer, M., 1994a: The Liouville equation and its potential usefulness for the prediction of forecast skill. Part I: theory. *Mon. Wea. Rev.*, 122, 703-713.

Ehrendorfer, M., 1994b: The Liouville equation and its potential usefulness for the prediction of forecast skill. Part II: Applications. *Mon. Wea. Rev.*, 122, 714-728.

Ehrendorfer, M., 1999: Kalman filtering and atmospheric predictability. In: Proceedings of the ECMWF workshop on diagnosis of data assimilation systems. November 1988. ECMWF, Shinfield Park, Reading RG2 9AX.

Emanuel, K. A., 1994: *Atmospheric Convection*. Oxford University Press, New York. 588pp.

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

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic ocean model. *J. Geophys. Res.*, 99 (C5), 10 143 - 10162.

Farrell, B. F. and P. J. Ioannou, 1996: Generalised stability theory. Part I: Autonomous operators. *J. Atmos. Sci.*, 53, 2025-2040.

Fisher, M., 1998: Development of a simplified Kalman filter. ECMWF Technical Memorandum No 260. ECMWF

Shinfield Park, Reading, RG2 9AX, UK.

Fisher, M. and P. Courtier, 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. ECMWF Technical Memorandum No 220. ECMWF, Shinfield Park, Reading RG2 9AX, UK.

Frisch, U., 1995: *Turbulence*. Cambridge University Press. Cambridge, UK.

Gage, K. S. and G. D. Nastrom, 1986: Theoretical interpretation of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft during GASP. *J. Atmos. Sci.*, 43, 729-739.

Gates, W. L., 1992: AMIP: The atmospheric model intercomparison project. *Bull. Am. Meteorol. Soc.*, 73, 1962-1970.

Gelaro, R., R. Buizza, T. N. Palmer and E. Klinker, 1998: Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. *J. Atmos. Sci.*, 55, 10 12- 1037.

Gill, A. E., 1982: *Atmosphere-Ocean Dynamics*. Academic Press Inc. London. 662pp.

Gleeson, T. A., 1966: A causal relation for probabilities in synoptic meteorology. *J. App. Met.*, 5, 365-368.

Griffies, S. M. and K. Bryan, 1997: A predictability study of simulated North Atlantic multi-decadal variability. *Climate Dynamics*, 13, 459-487.

Hamill, T. M., C. Snyder and R. E. Morse, 1999: A comparison of probabilistic forecasts from bred, singular vector and perturbed observation ensembles. *Mon. Wea. Rev.*, In press.

Hansen, D. V., and H. F. Bezdek, 1996: On the nature of decadal anomalies on North Atlantic sea surface temperature. *J. Geophys. Res.*, 101(CA), 8749-8758.

Hansen, J., 1998: Adaptive Observations in Spatially-extended, Nonlinear Dynamical Systems. D.Phil. thesis. Mathematical Institute, Oxford University. 207pp.

Harrison, M., T. N. Palmer, D. S. Richardson and R. Buizza, 1999: Analysis and model dependencies in medium-range ensembles: two transplant cast studies. *Q. J. R. Meteorol. Soc.*, 125, 2487-2515.

Harvey, L.O., K. R. Hammond, C. M. Lusk and E. F. Mross, 1992: The application of signal detection theory to weather forecasting behaviour. *Mon. Wea. Rev.*, 120, 863-883.

Hasselmann, K., 1976: Stochastic climate models. Part I. Theory. *Tellus*, 28, 473 - 485.

Hoffman, R. N. and E. Kalnay, 1983: Lagged average forecasting. *Tellus*, 35, 110-118.

Holland, G. J., T. McGeer and H. Youngren, 1992: Autonomous aerosondes for economical atmospheric soundings anywhere on the globe. *Bull. Am. Met. Soc.*, 73, 1987 - 1998.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Richie, and H. L. Mitchell, 1996. A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, 124, 1225-1242.

Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, 269, 676-679.

IPCC, 1990: *Climate Change: the IPCC Scientific Assessment*, J.T. Houghton, J.G. Jenkins and J.J. Ephraums (eds). Cambridge University Press, Cambridge, UK, 365pp

IPCC, 1996: *Climate Change 1995: The science of climate change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK, 572pp.

Joly, A., and coauthors, 1997: The Fronts and Atlantic Storm-Track Experiment (FASTEX): Scientific objectives

and experimental design. *Bull. Am. Met. Soc.*, 78, 1917 - 1940.

Katz, R. W. and A. H. Murphy, 1997. Forecast value: prototype decision-making models. *In Economic value of weather and climate forecasts*, Katz, R. W., and Murphy, A. H., Eds. Cambridge University Press, 222 pp

Kimoto, M. and M. Ghil, 1993: Multiple flow regimes in the northern hemisphere winter. Part I: Methodology and hemispheric regimes. *J. Atmos. Sci.*, 50, 2625-2643.

Kraichnan, R. H., 1971: Inertial-range transfer in two-dimensional turbulence. *Phys. Fluids*, 10, 1417-1423.

Lacarra, J.-F. and O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus*, 17, 321-333.

Latif, M., T. N. Stockdale, J. WoIff, G. Burgers, E. Maier-Reimer, M. M. Junge, K. Arpe and L. Bengtsson, 1994: Climatology and variability in the ECHO coupled GCM. *Tellus*, 46A, 351-366.

Latif, M., 1998: Dynamics of interdecadal variability in coupled ocean-atmosphere models. *J. Clim.*, 11, 602-624.

Lilly, D. K., 1973: Lectures in sub-synoptic scales of motions and two-dimensional turbulence. pp353 - 418. In *Dynamical Meteorology.* Edited by P. Morel. Reidel Publishing Company. Boston, USA. 622pp.

Lilly, D. K., 1983: Stratified turbulence and the mesoscale variability of the atmosphere.. *J. Atmos. Sci.*, 40, 749-761.

Lilly, D. K., 1989: Two-dimensional turbulence generated by energy sources at two scales. *J. Atmos. Sci.*, 46, 2026-2030.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, 42, 433-471.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, 21, 289-307.

Lorenz, E. N., 1975: *Climate predictability. The physical basis of climate modelling*. WMO, GARP Publication Series, 16, 132-136. World Meteorological Organisation. Geneva.

Lorenz, E. N., 1996: Predictability - a problem partly solved. In “Predictability” 1995 ECMWF seminar proceedings. ECMWF Shinfield Park, Reading, Berkshire RG2 9AX, UK.

Lorenz, E. N., 1993: *The Essence of Chaos*. University of Washington Press. Seattle, USA, 227pp.

Madden, R. A. and P. R. Julian, 1994: Observations of the 40-50 day tropical oscillation - a review. *Mon. Wea. Rev.*, 122, 814-837.

Manabe, S., and R. J. Stouffer, 1994: Century-scale effects of increased atmospheric CO2 on the ocean-atmosphere system. *Nature*, 364, 215-218.

Marshall, J. and F. Molteni, 1993: Toward a dynamical understanding of planetary flow regimes. *J. Geophys. Res.*, 93D, 10927 - 10952.

Mason, I, 1982: A model for assessment of weather forecasts. *Aust. Meteorol. Mag.*, 30, 291-303.

Mason, P. J. and D. J. Thompson, 1992: Stochastic backscatter in large-eddy boundary layers. *J. Fluid. Mech.*, 242, 51-78. simulations of

Mo, K. C. and M. Ghil, 1988: Cluster analysis of multiple planetary-scale flow regimes. *J. Geophys. Res.*, 93D, 10927-10952.

Molteni, F. and T. N. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Q. J. Meteorol. Soc.*, 119, 269-298.

Molteni, R, S. Tibaldi and T. N. Palmer, 1990: Regimes in the wintertime circulation over northern extratropics. I: Observational evidence. *Q. J. R. Meteorol. Soc*., 116, 31-67.

Molteni, R, R. Buizza, and T. N. Palmer, 1996: The ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc*., 122, 73-119.

Moncrieff, M. W., 1992: Organised convective systems: archetypal dynamical models, mass and momentum flux theory and parametrization. *Q. J. R. Meteorol. Soc*., 118, 819-850.

Montani, A. A. J. Thorpe, R. Buizza and P. Unden, 1999; Forecast skill of the ECMWF model using targeted observations during FASTEX. *Q. J. R. Meteorol. Soc*., in press.

Moore, A.M. and R. Kleeman, 1996: The dynamics of error growth and predictability in a coupled model of ENSO. *Q. J. R. Meteorol. Soc*., 122, 1405 -1446.

Moss, F. and P. V. E. McClintock, Eds., 1989: *Noise in Nonlinear Dynamical Systems. Vol 1, Theory of Continuous Fokker-Planck Systems.* Cambridge University Press, 353pp.

Münnich, M., M. A. Cane and S. E. Zebiak, 1991: A study of self-excited oscillations of the tropical ocean-atmosphere system. Part II: Nonlinear cases. *J. Atmos.Sci.*, 48, 1238-1248.

Mureau, R., F. Molteni and T. N. Palmer, 1993: Ensemble prediction using dynamically-conditioned perturbations. *Q. J. R. Meteorol. Soc.*, 119, 299-323.

Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteorol*., 12, 595-600.

Murphy, A. H., 1977. The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Mon. Wea. Rev.*, 105, 803-816.

Nastrom, G.D. and K. S. Gage, 1985: A climatology of atmospheric wavenumber spectra observed by commercial aircraft. *J. Atmos. Sci*., 42, 950-960.

Nicolis, C. and G. Nicolis, 1998: Closing the hierarchy of moment equations in nonlinear dynamical systems. *Phys. Rev.* E, 58, 4391-4400.

Palmer, T. N., 1993a: Extended-range atmospheric prediction and the Lorenz model. *Bull. Am. Met. Soc.*, 74, 49-65.

Palmer, T. N., 1993b: A nonlinear dynamical perspective on climate change. *Weather*, 48, 314-326.

Palmer, T. N., 1999: A nonlinear dynamical perspective on climate prediction. *J. Clim.*, 12, 575-591.

Palmer, T. N., F. Molteni, R. Mureau, R. Buizza, P. Chapelet and J. Tribbia, 1993: Ensemble Prediction. 1992 ECMWF Seminar Proceedings.

Palmer, T. N., R. Gelaro, J. Barkmeijer and R. Buizza, 1998: Singular vectors, metrics and adaptive observations. *J. Atmos. Sci.*, 55, 633-653.

Palmer, T. N., C. Brankovic: and D. S. Richardson, 2000: A probability and decision model analysis of PROVOST seasonal multi-model ensemble integrations. *Q. J. R. Meteorol. Soc*., to appear

Philander, S. G. H., 1990: *El Niño, La Niña and the Southern Oscillation.* Academic Press, San Diego, 293pp.

Pielke, R. A., 1999: Climate prediction as an initial value problem. *Bull. Am. Met. Soc.*, 79, 2743-2745.

Reynolds, C.A. and T. N. Palmer, 1998: Decaying singular vectors and their impact on analysis and forecast correction. *J. Atmos. Sci.*, 55, 3005-3023.

Reinhold, B. B. and R. T. Pierrehumbert, 1982: Dynamics of weather regimes: quasi- stationary waves and block-

ing. *Mon. Wea. Rev.*, 110, 1105-1145.

Richardson, D. S., 1999: Skill and economic value of the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.,*. to appear.

Schopf, P. S. and Suarez, M. J., 1988. Vacillations in a coupled ocean-atmosphere model. *J. Atmos. Sci.*, 45, 549-566.

Smith, L.A., C. Ziehmann, K. Fraedrich, 1999: Uncertainty dynamics and predictability in chaotic systems. *Q. J. R. Meteorol. Soc.*, 125, 2855-2886.

Stockdale, T. N., D. L. T. Anderson, J. O. S. Alves and M. A. Balmaseda, 1998: Global seasonal rainfall forecasts using a coupled ocean-atmosphere model. *Nature*, 392, 370-373.

Strang, G., 1986: *Introduction to applied mathematics*. Wellesley-Cambridge Press. pp758.

Sutton, R. T. and M. R. Allen, 1997: Decadal predictability of North Atlantic sea surface temperature and climate. *Nature*, 388, 563-567.

Swets, J. A., 1973: The relative operating characteristic in psychology. *Science*, 182, 990-1000.

Tennekes, H., 1991: Karl Popper and the accountability of numerical forecasting. In New developments in Predictability. ECMWF Workshop Proceedings, ECMWF, Shinfield Park, Reading UK, 1991.

Timmermann, A., J. Oberhuber, A. Bacher, M. Esch, M. Latif and E. Roeckner, 1999: Increased El Niño frequency in a climate model forced by future greenhouse warming. *Nature*, 398, 694-697.

Thompson, P. D., 1957: Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus*, 9, 275-295.

Toth, Z. and E. Kalnay, 1993: Ensemble forecasting at NMC: the generation of perturbations. *Bull. Am. Met. Soc.*, 74, 2317-2330.

Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, 12, 3297-3319.

Trenberth, K. E., 1992: *Climate System Modelling*. Cambridge University Press. Cambridge, UK, 788pp

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau and C Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperature. *J. Geophys. Res.*, 103, 14291-14324.

Trevisan A. and R. Legnani, 1995: Transient error growth and local predictability: a study in the Lorenz system. *Tellus*, 47A, 103-117.

Trevisan, A., Pancotti, F. and F. Molteni, 2000: Ensemble prediction in a model with flow regimes. *Q. J. R. Meteorol. Soc.*, in press.

Vautard, R. 1990: Multiple weather regimes over the North Atlantic. Analysis of precursors and successors. *Mon. Wea. Rev.*, 109, 784-812.

Webster, P. J. and T. N. Palmer, 1998: The past and future of El Niño. *Nature*, 390, 562- 564.

Webster, P. J., V. O. Magana, T. N. Palmer, J. Shukla, R. A. Thomas, M. Yanai and T. Yasunari, 1998: Monsoons: Processes, predictability and the prospects for prediction. *J. Geophys. Res.*, 103, 14451-14510.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press. 467pp

Xue, Y., M. A. Cane and S. E. Zebiak, 1997: Predictability of a coupled model of ENSO using singular vector anal-

ysis. *Mon. Wea. Rev.*, 125, 2043-2056.

Zebiak, S. E. and M. A. Cane,1987: A model El Niño-Southern Oscillation. Mon. Wea. Rev., 115, 2262-2278.