A Nonlinear Dynamical Perspective on Model Error: a Proposal for Nonlocal Stochastic-Dynamic Parametrisation in Weather and Climate Prediction Models

T.N.Palmer
ECMWF

February 3, 2000

Abstract

Conventional parametrisation schemes in weather and climate prediction models describe the effects of sub-grid scale processes by deterministic bulk formulae which depend on local resolved-scale variables and a number of adjustable parameters. Despite the unquestionable success of such models for weather and climate prediction, it is impossible to justify the use of such formulae from first principles. Using low-order dynamical-systems models, and elementary results from dynamical-systems and turbulence theory, it is shown that even if unresolved scales only describe a small fraction of the total variance of the system, neglecting their variability can, in some circumstances, lead to gross errors in the climatology of the dominant scales. It is suggested that some of the remaining errors in weather and climate prediction models may have their origin in the neglect of sub-grid scale variability, and that such variability should be parametrised by nonlocal dynamically-based stochastic parametrisation schemes. Results from existing schemes are described, and meteorologically- based mechanisms which might account for the impact of random parametrisation error on planetary-scale motions are discussed. Proposals for the development of stochastic-dynamic parametrisation schemes are outlined, based on potential vorticity diagnosis, singular vector analysis and a simple stochastic cellular automaton model.

For want of a nail the shoe was lost
For want of a shoe the horse was lost
For want of a horse the rider was lost
For want of a rider the battle was lost
For want of a battle the kingdom was lost
And all for the want of a horseshoe nail!
(Anon.)

1 Introduction

What has been the single-most important development in the atmospheric sciences over the last 50 years? There cannot be any disagreement that high on the list, if not at the very top, is the numerical model of the global climate system. Global atmospheric models have transformed the daily weather forecast, and, coupled to global ocean models, are now used routinely to make seasonal predictions, and climate change projections. As a result, 7-day forecasts are as skilful as 2-day forecasts 30-years ago (eg Bengtsson; 1999); the onset of El Nino and its impact on global weather patterns have been successfully predicted 6 months in advance (eg WMO, 1999); and quantitative projections of anthropogenic climate change provide the principal scientific basis for major international protocols on reducing the burning of fossil fuels (eg IPCC, 1996). The success of global weather and climate models derives not only from the mind-boggling development of computers over the last 50 years, but also to the ingenuity of scientists in devising accurate and efficient computational representations of the equations that govern climate.

And yet, today, these models are far from perfect representations of reality. In the short and medium-range, model error is not a negligible source of forecast error (Harrison et al, 1998), and the effects of model error must somehow be included in ensemble prediction systems to prevent forecast ensembles becoming systematically under-dispersive in the late medium range. Whilst mean systematic error is quite small in the medium range, it is interesting to note that the pattern of mean systematic error has hardly changed over the last couple of decades (C.Brankovic, personal communication). On longer timescales, systematic model error is a dominant source of forecast inaccuracy. For example, based on the European Union project PROVOST studying verifiable short-range climate variability using some of the major global climate models in Europe, seasonal-mean systematic error in simulating mid-tropospheric circulation patterns with observed sea suface temperature (SST) is comparable in magnitude with observed interannual variability (Brankovic and Palmer, 2000; Palmer et al, 2000). An example is illustrated in Fig 1, showing such systematic errors to be associated with an erroneous strengthening of the zonal flow and a weakening of planetary-wave activity. Similarly, with interactive oceans, the magnitude of the systematic error in tropical East Pacific SST is comparable with the magnitude of typical El Niño SST anomalies (Stockdale et al, 1998).

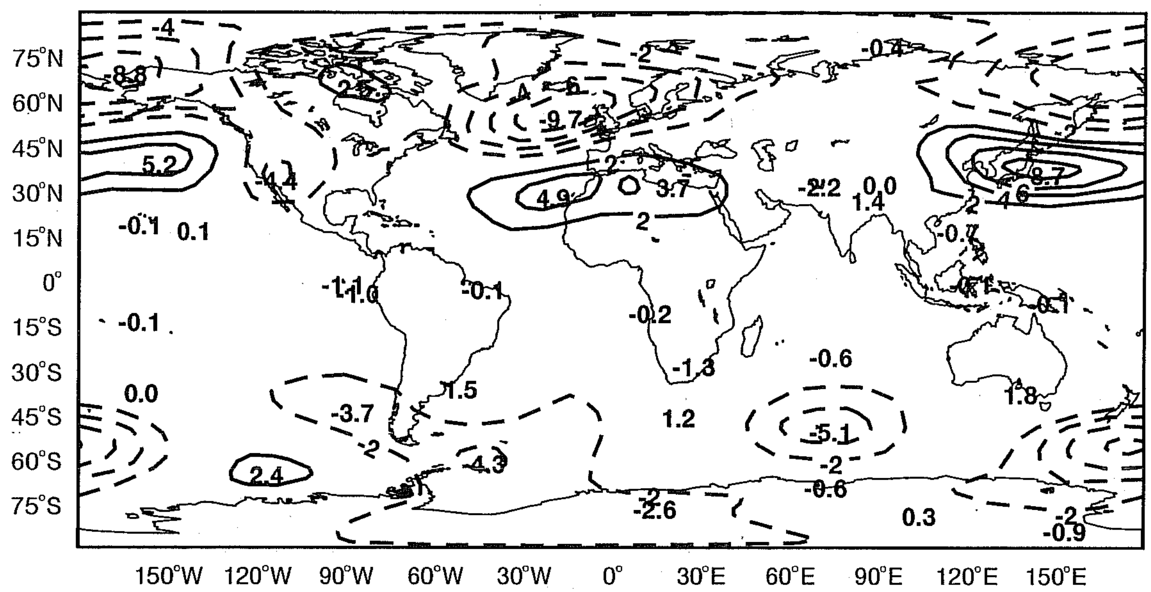

Figure 1: Day 31-120 mean 500hPa height systematic error during northern winter from a 14-winter set of seasonal integrations of the ECMWF atmosphere model with observed prescribed sea surface temperature, made as part of the PROVOST project data set (see Brankovic and Palmer, 2000; Palmer et al, 1999). Contour interval 2 dam.

The standard approach to modelling is based on the use of the explicit equations of motion truncated at some prescribed scale, and on the representation of scales below this by a number of deterministic bulk formulae which depend on the resolved flow and some adjustable parameters. Why are model errors still an important issue in weather and climate prediction? Perhaps the right parametrisations have yet to be formulated. Perhaps the right combination of existing parametrisations hasn't yet been found. Perhaps, with current parameter settings, substantial reduction of systematic error is just around the corner when global models can be run with increased resolution. However, there is a another possibility; perhaps the very methodology used to approximate the equations of motion for climate and weather prediction models is itself a source of large-scale systematic error. This possibility is not commonly discussed in the climate modelling community (though see, for example, Schertzer and Lovejoy, 1993; Mapes, 1996; Lander and Hoskins, 1997).

The practical difficulty in modelling climate lies in the fact that the governing equations describe a nonlinear coupling of scales of motion that potentially range from tens of thousands of kilometres to the viscous dissipation scale. The role of nonlinearity in generating scale-invariant geometries is vividly illustrated through the Mandelbrot set (eg Gulick, 1992). The nursery rhyme above captures the flavour of scale invariance as seen in the Mandelbrot set: if losing a kingdom is a tragedy, then losing a nail can also be. Any approximate scale invariance exists in the real climate is decisively broken in a conventional climate model. Does this matter? Are conventional parametrisations good enough, or are we missing important nails in the formulation of our weather and climate models?

The purpose of this paper is to make a case that explicit representation of sub-grid variability should be considered as part of the parametrisation process. More generally, it is proposed that the effects of unresolved scales should be represented by relatively-simple stochastic dynamical systems coupled to the resolved system over a range of scales, rather than by deterministic bulk formulae slave to the resolved dynamics at precisely the truncation scale. In section 2 we discuss the concept of parametrisation in the context of chaotic low-order dynamical systems models whose properties are known exactly. An example is shown (based on the Lorenz, 1963, system) of how the neglect of even an energetically-unimportant component of a dynamical system's state vector (a nail!) can lead to a major systematic error in the dominant component of the state vector (a kingdom!), and of how a simple stochastic representation of the energetically-weak component substantially improves the representation of the dominant component. An approach to the likely existence of an accurate parametrisation is discussed from a dynamical system's perspective, using geometric embedding theorems. With this approach, and using elementary scaling ideas from turbulence theory, it is suggested in section 3 that stochastic parametrisation may be needed in climate and weather prediction models, even taking foreseeable increases in model resolution into account. In section 4, simple existing stochastic parametrisations for weather and climate models are described, and their impact on weather and climate simulations discussed. Meteorologically-based mechanisms describing potential upscale cascades of error from the sub-grid scale to planetary scales are discussed in section 5. Two sets of mechanisms are discussed; the historical tendency for models to over-populate the more stable high zonal index circulation regimes, and the possible link between organised convection in the warm pool, the Madden-Julian oscillation (MJO), El Niño, and global-mean temperature. In section 6, a potential vorticity (PV) perspective on the stochastic representation of unresolved mesoscale organisation associated with convective and orographic systems is put forward; based on this two possible dynamically-based approaches to stochastic parametrisation are outlined in section 7; the first is based on an extension of the singular vector methodology used in medium-range ensemble prediction, the second is based on a simple cellular- automaton model.

2 Parametrisation and low-dimensional dynamical systems

For the purposes of this paper, a model is a finite representation of a set of partial differential equations which govern the climate system, or the atmosphere in particular, in a form which can be integrated numerically. Let us write the unapproximated equations for climate schematically as

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

where $\tilde{X}$ is (effectively) infinite dimensional, and $\tilde{F}$ is some nonlinear functional. A climate or weather prediction model is conventionally constructed by performing some Galerkin decomposition on equation 1 to produce a set of N deterministic ordinary differential equations

$$\dot{X} = F[X] + P[X; \alpha(X)] \tag{2}$$

where $F[X]$ represents terms retained in the Galerkin decomposition, and $P[X;\alpha]$ represents some parametrised or bulk representation of the effects of the unresolved components $\tilde{X}_i$ of $\tilde{X}$ $(i > N)$. Conceptually, a parametrisation is based on the notion that there exists a statistical ensemble of unresolved sub-grid scale processes within a grid box x_j, in some secular equilibrium with the grid-box mean flow. For example, borrowing ideas from statistical mechanics, many familiar parametrisations involve the diffusive approximation, where α would be a diffusion coefficient, possibly dependent on the Richardson number of the resolved flow X. Two other examples, relevant to the discussion below, are mentioned. If the resolved-scale vertical temperature gradient associated with X is convectively unstable at x_j, then over some prescribed timescale (given by α) P would represent the effect of an ensemble of sub-grid convective plumes which operate to relax X back to stability at x_j (Betts and Miller, 1986) . For flow over unresolved topography, P could represent an ensemble of sub- grid orographic gravity waves, propagating vertically, and breaking at some height above the surface, leading to a drag on the resolved flow associated with X, at these heights (eg Palmer et al, 1986). In these and most other commonly-used parametrisations, the bulk representation of small-scale processes within x_j is assumed to be a local deterministic function of X at x_j ie

$$\dot{X}_j = F_j[X] + P[X_j; \alpha(X_j)] \tag{3}$$

where X_j and F_j represent projection into a subspace associated with x_j.

Despite weather and climate models being formulated in this way, it is generally agreed that the notion of unresolved scales in secular equilibrium with resolved scales is not rigorously justifiable; there is no known gap in spatio-temporal spectra of atmospheric circulations. However, it might be argued that by truncating the equations on a sufficiently small scale, the sub-grid motions will be so energetically weak compared with the large-scale circulations (in which we are primarily interested) that it should be possible to describe the effect of small scales, to reasonable accuracy, by conventional parametrisations. Is this a reasonable argument? Is it sufficient to say that just because a particular scale of motion only explains a small amount of variance, its effect can be represented in a truncated model where that scale is not explicitly represented, by a local bulk formula?

The following example shows that this idea is manifestly false. Consider the Lorenz (1963) model

$$\begin{aligned} \dot{X} &= -\sigma X + \sigma Y \\ \dot{Y} &= -XZ + rX - Y \\ \dot{Z} &= XY - bZ \end{aligned} \tag{4}$$

The familiar Lorenz attractor is illustrated in Fig 2a, using Lorenz's original choice of parameters ($\sigma = 10$, $b = 8/3$, $r = 28$,). Based on this parameter setting, the governing equations can be written in terms of the three empirical orthogonal functions (EOFs) of the Lorenz model $(\tilde{a}_1, \tilde{a}_2, \tilde{a}_3)$, so that equation 4 is transformed to (Selten, 1995)

$$\begin{aligned} \dot{\tilde{a}}_1 &= 2.3\tilde{a}_1 - 6.2\tilde{a}_3 - 0.49\tilde{a}_1\tilde{a}_2 - 0.57\tilde{a}_2\tilde{a}_3 \\ \dot{\tilde{a}}_2 &= -62 - 2.7\tilde{a}_2 + 0.49\tilde{a}_1^2 - 0.49\tilde{a}_3^2 + 0.14\tilde{a}_1\tilde{a}_3 \\ \dot{\tilde{a}}_3 &= -0.63\tilde{a}_1 - 13\tilde{a}_3 + 0.43\tilde{a}_1\tilde{a}_2 + 0.49\tilde{a}_2\tilde{a}_3 \end{aligned} \quad (5)$$

The 3^{rd} EOF only explains 4% of the total variance of the system. Hence we might consider that a reasonable approximation to the full model could be obtained by truncating the system to 2 EOFs, ie

$$\begin{aligned} \dot{a}_1 &= 2.3a_1 - 0.49a_1a_2 \\ \dot{a}_2 &= -62 - 2.7a_2 + 0.49a_1^2 \end{aligned} \quad (6)$$

Such a truncated model is a reasonable short-range forecast model, in the sense that, the initial tendencies $\dot{a}_1$ and $\dot{a}_2$ in equations 5 and 6 agree well with $\dot{\tilde{a}}_1$ and $\dot{\tilde{a}}_2$ (respectively) for points on the Lorenz attractor. However, (for reasons discussed immediately below) the climatology of equation 6 bears no relation to the climatology of the full model. In particular, instead of exhibiting chaotic variability, the truncated model evolves, from any initial state, to one of two fixed points (corresponding to the two regime centroids shown in Fig 2a). The climatology of this truncated model has gross systematic errors in both its mean state, and its internal variability.

We might naively consider, instead of neglecting the 3rd EOF, parametrising it in terms of the 1st 2 EOFs ie

$$\begin{aligned} \dot{a}_1 &= 2.3a_1 - 6.2a_3 - 0.49a_1a_2 - 0.57a_2a_3 \\ \dot{a}_2 &= -62 - 2.7a_2 + 0.49a_1^2 - 0.49a_3^2 + 0.14a_1a_3 \\ a_3 &= f(a_1, a_2) \end{aligned} \quad (7)$$

Whilst an astute choice of parametrisation f might well improve the skill of this model as a short-range forecast model (producing trajectories which shadow the full system for short periods of time), equation 7, like equation 6, is climatologically doomed to failure! The reason we can be sure about this is (eg Gulick, 1992)

The Poincaré-Bendixon Theorem
Consider a 2D autonomous system

$$\begin{aligned} \dot{x} &= f(x,y) \\ \dot{y} &= g(x,y) \end{aligned}$$

and a trajectory which starts at some point p and is confined to some bounded region of phase space. Then the trajectory must either: a) terminate at a fixed point b) return to p, or c) approach a limit cycle

Because of the Poincaré-Bendixon Theorem, 2-D autonomous systems of ODEs cannot be chaotic. In the Lorenz (1963) system, the mean state is intrinsically linked with the stability properties of the attractor. The Poincaré-Bendixon theorem provides a counterexample to the claim that good short-range forecasting models necessarily make good climate models. (A different argument for the failure of models with severe EOF truncation can be made using generalised stability theory; Farrell and Ioannou, 1996, 1999a,b).

Let us return to the truncated Lorenz (1963) system. Suppose the dynamical equation for the 3rd EOF is unknown. From the discussion above, it cannot be parametrised as a deterministic function of the dominant EOFs; we need to represent its variability somehow. Consider, therefore, the representation

$$\begin{aligned} \dot{a}_1 &= 2.3a_1 - 6.2a_3 - 0.49a_1a_2 - 0.57a_2a_3 \\ \dot{a}_2 &= -62 - 2.7a_2 + 0.49a_1^2 - 0.49a_3^2 + 0.14a_1a_3 \\ a_3 &= \beta \end{aligned} \quad (8)$$

where $\beta(t)$ is a stochastic variable randomly drawn from a gaussian probability density function (pdf) whose whose variance is equal to the explained variance associated with a_3. As such a stochastic model is not autonomous, it would not be constrained by the Poincaré-Bendixon theorem.

Fig 2b, c shows the impact of such a stochastic parametrisation (F. Selten, personal communication). Fig 2b shows results using the parametrised model (represented by equation 8 where the stochastic forcing is updated every 0.05 nondimensional Lorenz time units. The state vector oscillates irregularly about one of the regime centroids with unimodal pdf. Hence, both the mean and the variance in the space S_{EOF} of the dominant EOFs from this stochastic model are still seriously in error (though clearly the error in the variance is not as dire as with no stochastic parametrisation). On the other hand, Fig 2c shows the climate of the model when the stochastic variable is updated every 0.1 Lorenz time units. Now the model pdf is clearly bimodal similar to the exact Lorenz attractor (though the state vector tends to reside in a particular regime too long; see additional remarks in section 5). Hence, through a simple stochastic representation of the energetically-weak 3rd EOF, we have been able to fundamentally reduce

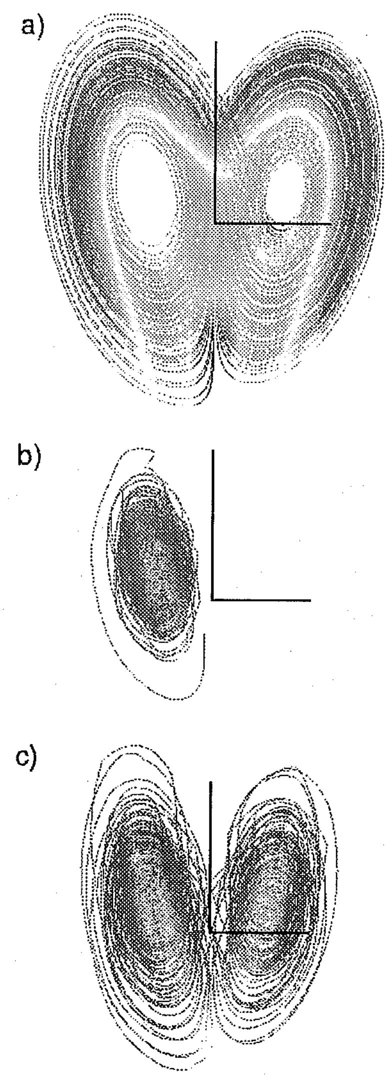

Figure 2: a) Lorenz attractor, b) Truncated Lorenz equations (cf equation 8) using stochastic forcing with random forcing updated every 0.05 Lorenz time units, c) truncated Lorenz equations using random forcing updated every 0.1 Lorenz time units. F.Selten, personal communication

the time-mean systematic error in S_{EOF}. This result depends fundamentally on the nonlinearity of the underlying system in S_{EOF}. (In fact, the use of multiplicative noise is not necessary to produce a bimodal pdf, the result illustrated in Fig 2c can be qualitatively replicated merely by adding noise terms to the right hand sides of equation 6 F. Selten, personal communication.)

A more general way of looking at this problem is provided by a classic theorem in differential geometry (eg Dodson and Poston, 1979)

The Whitney Embedding Theorem
Consider an m-dimensional manifold M. Then M can be embedded in $\mathbf{R}^n$ *providing* $n > 2m$.

This theorem has been generalised by Takens (1981) so that M includes chaotic attractors (see also Sauer et al, 1991). Note that $n > 2m$ is a sufficient condition for embedding; a necessary condition is that $n \geq m$.

For the case of Lorenz (1963) the attractor dimension is greater than 2; hence the attractor cannot be embedded in the space spanned by the two dominant EOFs. However, there are dynamical models where these embedding ideas can point to the likelihood of an accurate parametrisation. Consider for example the hierarchical Lorenz (1996) system

$$\dot{\tilde{X}}_i = -\tilde{X}_{i-2}\tilde{X}_{i-1} + \tilde{X}_{i-1}\tilde{X}_{i+1} - \tilde{X}_i + F - \frac{c}{b}\sum_{j=1}^{N}\tilde{x}_{j,i} \quad (9)$$

$$\dot{\tilde{x}}_{j,i} = -cb\tilde{x}_{j+1,i}\tilde{x}_{j+2,i} + cb\tilde{x}_{j-1,i}\tilde{x}_{j+1,i} - c\tilde{x}_{j,i} + \frac{c}{b}\tilde{X}_i \quad (10)$$

where the $\tilde{X}_i$ are large-scale variables, and the $\tilde{x}_{j,i}$ are small-scale variables. For $N = 8$, the attractor of the Lorenz (1996) system has dimension ~ 4 (Orrell, 1999; Smith, 2000). The variables $X_1, X_2, ..., X_8$ span $\mathbf{R}^8$; hence, by the Whitney/Takens theorem, an embedding of the attractor into the space of large-scale variables, and hence an accurate parametrised model

$$\begin{aligned} \dot{X}_i &= -X_{i-2}X_{i-1} + X_{i-1}X_{i+1} - X_i + F + P_i \\ P_i &= P_i(X_1, X_2, ...X_N). \end{aligned} \quad (11)$$

is certainly possible. In fact, Orrell (1999) has studied the local linear parametrisation

$$P_i = \alpha_0 + \alpha_1 X_i \quad (12)$$

where the parameters α_0 and α_1 are determined by linear regression. Fig 3 shows a spectral bifurcation diagram for the exact system (Fig 3a) and for the linearly parametrised model (Fig3b). (The spectral bifurcation diagram gives a power spectrum for the X_1 variable, for different values of F.) For small values of F, both the exact system and the parametrised model show most of the power on a set of discrete frequencies; conversely for large F the power is distributed over a continuum of frequencies for both the exact system and the parametrised model. In general the model compares well with the system. However, there are values of F where the parametrised model fails. For example, near $F = 7$, the model's spectrum is dominated by power on discrete frequencies; in the exact system the power is distributed more uniformly with frequency. Orrell (personal communication) has found parametrisations P_i^{NL} which fit the exact system better than the parametrisation in equation 12. However, these P_i^{NL} are nonlocal in the sense that the parametrised tendency for X_1, say, is a function not only of X_1, but also of the other large-scale variables X_2... etc. Note that the Whitney/Takens embedding theorem does not require the parametrisations to be local ($P_i = P_i(X_i)$) in any way. This latter point is relevant when we consider stochastic parametrisation in weather and climate models.

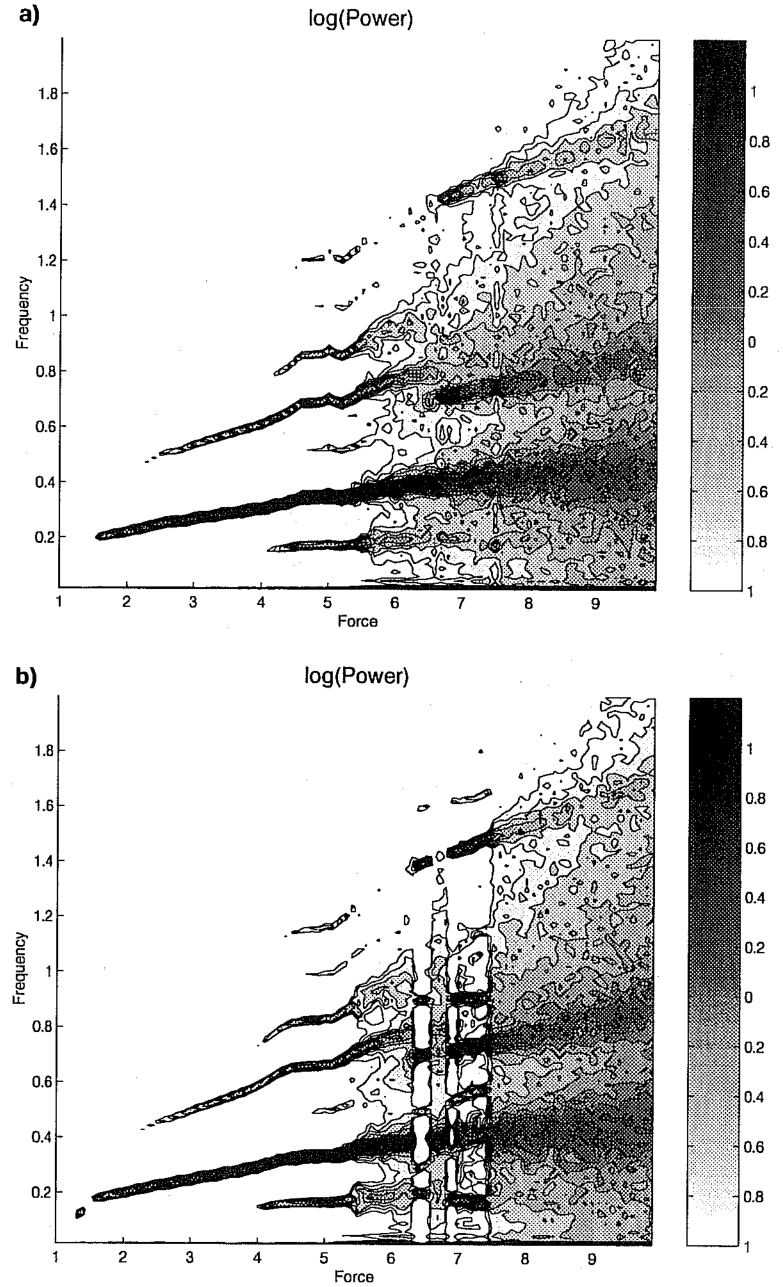

Figure 3: Power spectrum bifurcation diagram for the X_1 variable in Lorenz (1996) system. a) exact system (equation 9). b) parametrised system (equations 11 and 12). From Orrell, 1999.

3 Parametrisation and high-dimensional dynamical systems

In the previous section, the concept of parametrisation was discussed, not from the conventional notion of secular equilibrium, but from a dynamical-systems perspective. If the dimension of a system's attractor is small compared with the size N of a model truncation the system, then conventional parametrisation of unresolved scales may be possible.

Chaotic atmospheric behaviour can be simulated in low-order models (eg Ghil and Childress, 1987). The attractor dimension, $O(10)$, of such models is therefore much less that the number of resolved variables in a typical weather or climate prediction model. However, with an intermediate-resolution quasigeostrophic model with O(1000) degrees of freedom, the dimension of the simulated attractor appears to increase to O(100) (Palmer, 1997). One can therefore ask the question: how does the dimension of the simulated climate attractor increase as the number of resolved variables increases to values typical of weather and climate prediction models?

As discussed above, the dimension of the attractor cannot be assessed from the number of dominant EOFs of the flow. By the arguments in section 2 above, it may be that 100 EOFs describe 99% of the variance of the flow; however, both the variability and the mean state of the system within the space S_{EOF} of dominant EOFs may depend fundamentally on the remaining 1% of explained variance.

If we consider the whole range of atmospheric motions, then the atmosphere could be viewed as a high Reynolds-number turbulent fluid. A simple representation of such a high-dimension system is given by the hierarchical Gledzer, Ohkitani and Yamada (GOY) shell system (eg Bohr et al, 1999)

$$\dot{u}_n + \nu {k_n}^2 = ik_n({u^*}_{n+1}{u^*}_{n+2} - \frac{\delta}{2}{u^*}_{n-1}{u^*}_{n+1} - \frac{1-\delta}{4}{u^*}_{n-1}{u^*}_{n-2}) + f_n \qquad (13)$$

where $n = 1, 2...\hat{N}$ and $k_n = 2^n k_0$ and u_n is a complex variable. For $\delta < 1$ the GOY model has both energy and helicity-like invariance. For such a system, the Lyapunov dimension D of the attractor increases proportionally with the truncation shell $\hat{N}$ (Bohr et al, 1999). Such a system cannot therefore be described as 'low-order' in the limit of high Reynolds number ($\hat{N} \to \infty$). Hence the discussion above suggests it will be impossible to find an accurate parametrised model of the shell system with truncation scale $N \ll \hat{N}$. On the other hand, it could be imagined that the systematic effects of parametrisation error on the mean and variance of the large-scale ($n \sim 1$) flow might be small enough if $N >> 1$, even if $N \ll \hat{N}$. However, the scaling argument below suggests that this cannot be guaranteed.

A fundamental characteristic of error growth in climate and weather prediction is the upscale transformation of error, from small scale to large scale. This characterises the essential element of the 'butterfly effect' paradigm, as much as does amplitude growth.

The evolution of a small-amplitude initial perturbation $\delta x(t_0)$ to an initial state X is determined by linearising equation 3 about some nonlinear ('basic state') solution. This can be written as

$$\delta x(t) = M(t, t_0)\delta x(t_0) \tag{14}$$

where M is the tangent propagator (or tangent-linear model) associated with equation 3. The initial perturbations with largest amplitude at time t are the dominant singular vectors of M (eg Farrell and Ioannou, 1996). For atmospheric flows, the dominant energy-norm singular vectors of M often describe an upscale transformation from sub- cyclone scales to cyclone and planetary scales (eg Molteni and Palmer, 1993; Buizza and Palmer, 1995). This non-modal behaviour arises because of the non-normality of M.

The 'butterfly effect' can describe not only small-scale error in the intial state, but also model error in representing sub-gridscale activity by bulk formulae. For example, consider equation 3 with an additional weak imposed forcing $f(t)$, representing random errors in the bulk formula P. The influence of f on the resolved scale over some finite interval $\Delta t = [t_0, t]$ can be written (cf equation 14) as

$$\delta x(t) = \int_{t_0}^{t} M(t, t')f(t')dt' \tag{15}$$

If f was constant over Δt, then forcing perturbations with largest impact on the the flow at time t would be given by the dominant singular vectors of

$$\mathcal{M}(t, t_0) \equiv \int_{t_0}^{t} M(t, t')dt'. \tag{16}$$

As with initial perturbations, the non-normality of this operator means that variability in f could efficiently force variability in the large-scale components X of the resolved flow.

To take this argument futher, consider the following well-known scaling argument for error propagation in homogeneous isotropic turbulence. Consider a model of such a system with truncation wavenumber k_N within the inertial subrange, and let f denote truncation error at k_N. Let $E(k)$ denote the energy kinetic energy per unit wavenumber of the atmosphere at wavenumber k. Following Lorenz (1969) and Lilly (1973), let us assume that the time it takes error at wavenumber $2k$ to infect wavenumber k (ie to propagate one 'octave') to be proportional to the 'eddy turn-over time' $\tau(k) = k^{-3/2}[E(k)]^{-1/2}$. Hence the time $\Omega(k_N)$ taken for uncertainty to propagate N_o octaves from wavenumber k_N to some large scale k_L of interest is given by

$$\Omega(k_N) \equiv \sum_{n=0}^{N_o-1} \tau(2^n k_L) \tag{17}$$

In the case of a two-dimensional turbulence in the enstrophy-cascading inertial subrange between some large-scale (eg baroclinic) forcing scale and dissipation scale, then $E(k) \sim k^{-3}$, τ

is independent of k, and $\Omega(k_N) \sim N_o$ which diverges as $k_N \to \infty$. By contrast if $E(k) \sim k^{-5/3}$ so that $\tau \sim k^{-2/3}$ then $\Omega(k_N)$ tends to a finite limit

$$\Omega(k_N) \sim \tau(k_L) \tag{18}$$

as $k_N \to \infty$.

From a physical point of view, this analysis suggests that for $k^{-5/3}$ flow the effect of neglecting unresolved-scale variance on the (mean and variance of the) large-scale flow cannot necessarily be made arbitrarily small by resolving more of the inertial subrange. (Mathematically, Ω^{-1} is related to the dominant singular value of $\mathcal{M}$, where the norm on the right singular vector space involves a projection onto wavenumber k_N, and the norm on the left singular vector space maximises growth onto wavenumber k_L. By contrast, $\tau^{-1}(k_N)$ is associated with the dominant Lyapunov timescale of the system. This demonstrates a fundamental difference between Lyapunov and singular vectors in a multi-scale system.)

Whilst the atmosphere has a k^{-3} spectrum on cyclone scales, there is evidence for a $k^{-5/3}$ spectrum on scales up to a few hundred kilometres (Nastrom and Gage, 1985; Gage and Nastrom, 1986; Cho et al, 1999). Lilly (1983) has suggested that this at least partly associated, not with 3-D motion, but with upscale energy transformations forced by organised mesoscale activity (including mesoscale convective complexes, MCCs). More recently, in addition to the inverse cascade process of 2-D turbulence, Lilly et al (1998) suggest a 'PVs-spreading' mechanism for upscale transformations, associated with a direct effect of mass outflow from MCCs (see section 6 below). The truncation scale of typical global climate and weather prediction models is generally within this range of observed $k^{-5/3}$ activity. The ECMWF model, even with truncation scales of tens of kilometres, shows little sign of the shallower $k^{-5/3}$ spectum (M.Hortal, personal communication), and there is evidence that this behaviour is typical amongst weather and climate prediction models (Koshyk et al, 1999a). This could be taken as evidence of an inability of global weather and climate prediction models to simulate mesoscale variability. According to the scaling argument above, such a shortcoming could have an impact on the variability (and mean value) associated with scales k_L within the k^{-3} range.

On the other hand, one particular climate model (the GFDL SKYHI model) has shown unambiguous evidence of $k^{-5/3}$ variability near its truncation scale (Koshyk et al, 1999b), though such a model cannot be expected to simulate mesoscale organisation (eg associated with MCCs). Recent diagnosis (Koshyk et al, 2000) suggests that the shallow mesoscale regime in this model is not associated with an inverse energy cascade, and much of the energy is contained in gravity-wave components near the truncation scale, consistent with VanZandt's (1982) analysis of the observations. As such, it would appear that there is still some ambiguity in the interpretation of the existence of the observed $k^{-5/3}$ spectra, and the mere existence of this shallow spectrum cannot be taken as unambiguous evidence of an inverse cascade. If the behaviour of the SKYHI model is realistic, then misrepresentation of mesoscale organisation in weather and climate models may not have a substantial effect on large-scale variability. On the other hand, if we assume that some form of inverse cascade associated with mesoscale organisation really does exist, then the results of Koshyk et al (2000) warn that the injection of unbalanced stochastic noise at the truncation scale of a climate model cannot be guaranteed to efficiently propagate energy upscale. In section 5 it is proposed to project the stochastic forcing onto those balanced PV structures which ensure upscale propagation.

4 Stochastic parametrisation in climate and weather prediction models

The use of stochastic noise to represent unpredictable small-scale variability is familiar in a number of geophysical models (eg Hasselmann, 1976; Farrell and Ioannou, 1993; DelSole and Farrell, 1995; Penland, 1996, Newman et al, 1996; Moore and Kleeman, 1996). Moreover, following Leith (1990), a stochastic representation of sub-grid scale stress variations has also been used in comprehensive models of three-dimensional turbulent flow, leading to improvement in the simulation of the resolved flow near rigid surfaces, and to a more accurate logarithmic flow profile in particular (Mason and Thompson, 1992).

A version of this stochastic parametrisations ('stochastic backscatter') has been applied to the UK Meteorological Office's global weather prediction model (Evans et al, 1998). The parametrisation was adapted to produce quasi non-divergent horizontal velocity increments with fixed amplitude, at all grid points, and at one model level in the lower troposphere. It was shown that an ensemble of integrations with different realisations of the stochastic parametrisation could produce significantly different cyclone- and planetary-scale variability by the late medium range. The effect of this scheme on model systematic error has not yet been quantified.

A somewhat different stochastic formulation was proposed by Buizza et al (1999), linking stochastic forcing to regions in the atmosphere where conventional sub-grid parametrisation is active, specifically:

$$\dot{X}_j = F_j[X] + \beta P[X_j; \alpha(X_j)] \tag{19}$$

where β is a stochastic variable drawn from a uniform distribution in $[0.5, 1.5]$. The random drawings were constant over a time range of 6 hours, and a spatial domain of $10^\circ \times 10^\circ$ latitude/longitude. As with the stochastic forcing in the truncated Lorenz (1963) model above, the choice of spatio-temporal autocorrelations strongly influence the performance of the scheme.

A dramatic effect of this stochastic representation (distinct to the impact of intial singular vector perturbations) was found in the simulation of isolated atmospheric vortices (Puri et al, 2000). An example is given in Fig 4 which shows sea-level pressure over part of Australia

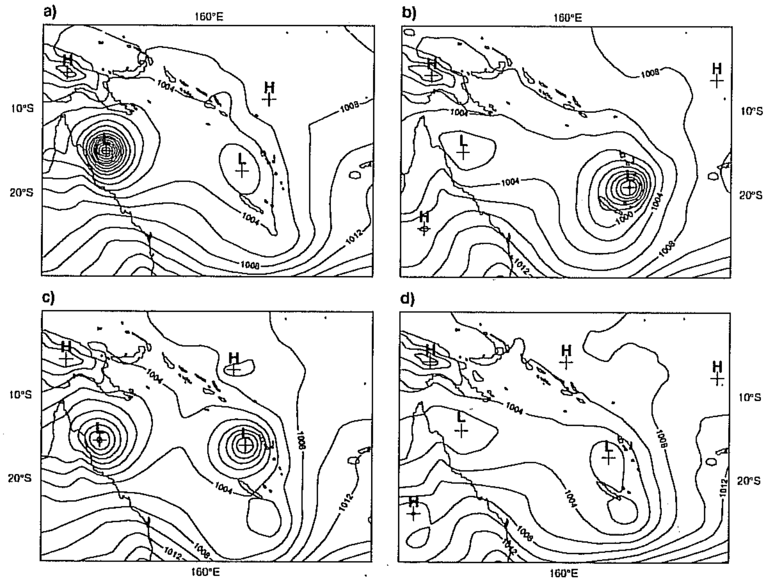

Figure 4: Four 2-day integrations of the ECMWF model from identical starting conditions but different realisations of the stochastic parametrisation scheme represented by equation 19 with parameter settings as given in Buizza et al (1999). The field shown is sea-level pressure over parts of Australia and the west Pacific. The depressions in the pressure field represent potential tropical cyclones. K.Puri, personal communication.

and the west Pacific from four 2-day integrations of the ECMWF model. The integrations have identical starting conditions, but different realisations of 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. This example clearly shows the difficulty in predicting tropical cyclone development, and its sensitivity to model parametrisation. More importantly, it shows how the variability of resolved circulation features, particularly in the tropics, is influenced by the stochastic representation. This fact is used to discuss further possible influences on large-scale systematic error (see section 5).

Such a stochastic parametrisation also has a profound effect on the skill of probabilistic forecasts of rainfall, as given by the ECMWF ensemble prediction system (Palmer et al, 1993; Molteni et al, 1995; Buizza et al, 1999). Fig 5 shows a probabilistic measure of forecast skill (area under the relative operating characteristic curve) for the dichotomous event: 12hour accumulated precipitation is greater than 20mm, taken over individual grid points in the NH. It can be seen that there is a fairly substantial improvement in skill in both the summer and wintertime, when the stochastic physics parametrisation is included.

A full assessment of the impact of the stochastic physics parametrisation on the systematic error of the ECMWF coupled model (Stockdale et al, 1998) is currently in progress. Preliminary results showing the impact of stochastic physics on 6-month mean tropical SST is shown in Fig 6, based two ensembles of 30 6-month coupled-model integrations (F. Vitart, personal communication). The shaded region indicates areas where the impact on stochastic physics is

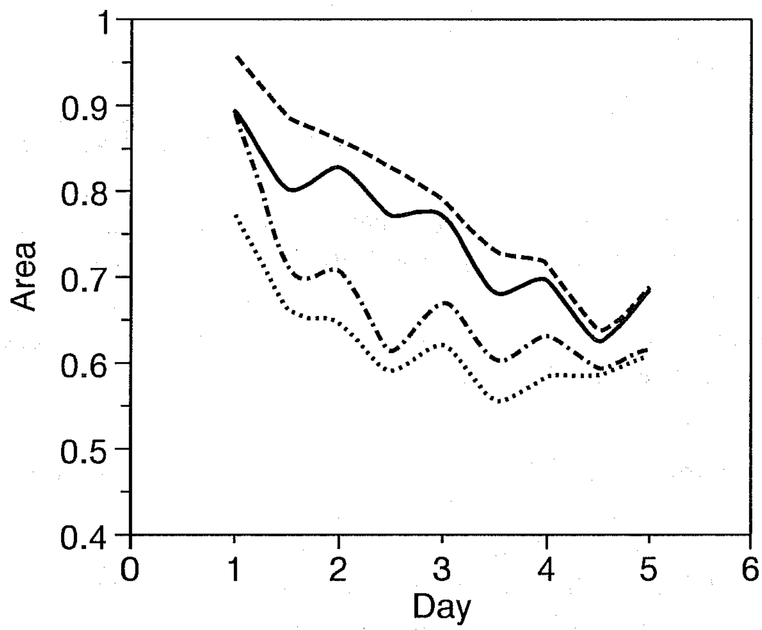

Figure 5: Area under the relative operating characteristic curve for the event: 12hr accumulated precipitation greater than 20mm. Based on 50-member ensemble integrations of the ECMWF ensemble prediction system. Without stochastic physics (solid) and with stochastic physics (dashed) for the period 16-22 December 1997. Without stochastic physics (dotted) and with stochastic physics (chain-dashed) for the period 16-22 December 1997. From Buizza et al, 1999

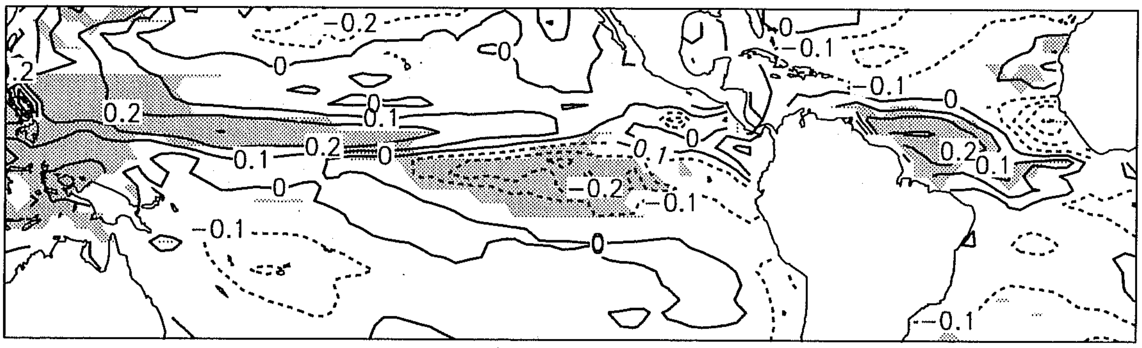

Figure 6: The impact of stochastic physics on the systematic error of SST in the ECMWF coupled model. Based on 30 pairs of 6-month integrations started one day apart in the spring of 1997. Contour interval 0.1K. Regions where the impact is statistically significant at the 95% level are shown shaded. F.Vitart, personal communication.

statistically significant at the 95% level. Across the western Pacific and western Atlantic, the impact is such as to reduce the systematic error in SST (a general cold bias) without stochastic physics.

It should be noted that whilst the scheme given in equation 19 has shown some positive impact on the skill of short and medium-range forecasts, and on systematic error, it is not energetically consistent with associated surface fluxes of heat and momentum. Indeed, if the surface fluxes were perturbed in such a way as to be consistent with stochastic perturbations to the parametrised tendencies, the impact on SST could be larger than that shown in Fig 6. The problem of energetic consistency could be addressed straightforwardly if, instead of the parametrisation tendency, the parametrisation input fields were stochastically perturbed, ie

$$\dot{X}_j = F_j[X] + P[X_j + \beta; \alpha(X_j)] \tag{20}$$

Experimentation with such a scheme is in preparation.

Indeed there is much scope for development of such schemes, and equations 19 and 20 are both simplistic. Suggestions for more dynamically-based stochastic parametrisations are outlined in section 7. However, before concluding this section, it should be noted that there is another commonly-used technique for representing model uncertainty in ensemble predictions: the multi-model ensemble . This is achieved by incorporating within the ensemble a number of quasi-independent weather or climate prediction models. Results suggest that probabilistic skill scores for the multi-model ensemble can exceed the skill of individual-model ensembles (Harrison et al, 1999; Palmer et al, 1999). A variant on the multi-model technique has been proposed by Houtekhamer et al (1996) where ensemble members are integrated within a common numerical framework, but with different parametrisations P or different values of the parameters α. This is similar to the multi-model technique insofar as the parametrisations and parameters are held fixed within a particular integration. These techniques are conceptually distinct from the type of stochastic physics scheme described schematically in equations 19 and 20. In multi-model ensembles, the effective model perturbations account for the fact that, even in circumstance where secular equilibrium might be a reasonable assumption, the expectation value of the pdf of sub-grid processes is not itself well known. By contrast, the schemes described in equations 19 and 20 are attempts to account for the fact that in circumstances of mesoscale organisation, the pdf of sub-grid processes, even if it was well known, would not be sharp around the mean.

5 Possible scenarios for an upscale cascade of model error

In studying the truncated Lorenz (1963) models in section 2, it was shown that the neglect of variability of energetically-weak components of a dynamical system can have a systematic impact on the mean state in the space S_{EOF} of dominant EOFs. Is there any evidence that this argument applies to climate dyamics? It was noted in section 4 that despite a positive impact on model performance, the schemes developed so far are simplistic and somewhat energetically inconsistent. Moreover, as discussed in section 6 below, they arguably do not explicitly perturb the most dynamically-relevant variables. As such there is a need for further development of such schemes before extensive integrations and detailed diagnostic analysis are performed (see section 7 below). Nevertheless, it is perhaps worth speculating on possible meteorological mechanisms in the real climate system whereby stochastic physics can influence the largest scales.

Consider first the extratropics. As discussed in many papers (see eg references in Palmer, 1993; Corti et al, 1999), there is evidence that in the northern extratropics, the pdf of S_{EOF} has non-gaussian, and possibly multi-modal properties. Let us assume that a climate model can correctly simulate the circulation regimes as diagnosed by Corti et al (1999) from operational analyses, but with inadequate small-scale variability to trigger regime transitions. In such a model, a plausible scenario (first suggested by Molteni and Tibaldi, 1990) is that the more stable circulation regimes will become overly-populated. Fig 7 (from Molteni and Tibaldi, 1990) illustrates this effect. Fig 7a shows a hypothetical double potential well; Fig 7b shows two bimodal pdfs computed using a Fokker-Planck equation for white- noise stochastic perturbations evolving in the double potential well. The dashed-line results are based on a stochastic variance which has been reduced by a factor of 2 compared with the solid-line results. With reduced stochastic variance, the mean state shifts towards the more stable and populated regime. According to the analysis in Corti et al (1999), the most populated regime in the real atmosphere in the recent past ('cluster A'; see Fig 8) corresponds to a relatively zonal flow with weak planetary waves (and is equivalent to the 'cold ocean/warm land pattern; COWL' pattern of Wallace et al, 1996. In the Pacific/ North American sector, this circulation regime projects onto the positive phase of the Pacific/North American pattern (Wallace and Gutzler, 1981) which is known to be a relatively stable pattern (Palmer, 1988). In this nonlinear perspective, therefore,

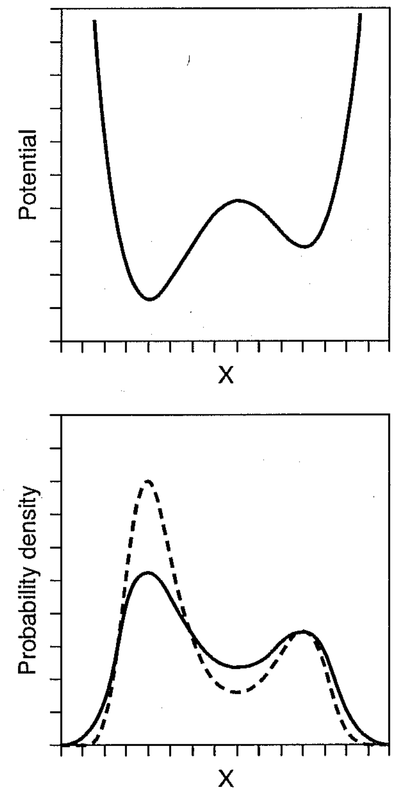

Figure 7: a) Potential-well function. b) PDFs obtained using a) and given white noise stochastic forcing, based on a solution of the Fokker-Planck equation. The white noïse for the dashed line solution is half that for the solid line solution. Relative to the solid-line solution, the mean state for the dashed-line solution is biased towards the more populated regime. From Molteni and Tibaldi, 1990

the systematic error of a climate model without stochastic physics would tend to correlate with the cluster A anomaly field. Comparing Fig 1 with Fig 8 in the northern hemisphere, one can clearly see a correlation between the ECMWF model systematic error and the most densely populated cluster from the recent past.

It is also possible that a systematic misrepresentation of mesoscale variability can have a systematic effect on larger-scale climatic variability in the tropics. Wang and Schlesinger (1999) have shown that a climate model's ability to simulate the MJO is sensitive to certain thresholds in the convection scheme, particularly the value of theshold relative humidity, below which the convection scheme will not trigger. Similar experiments confirming this sensitivity have been performed at ECMWF using convective available potential energy (CAPE) to define the threshold (L. Ferranti, D. Gregory, C. Jakob, personal communication). Specifically, with small threshold CAPE, the ECMWF has a rather poor simulation of the MJO. With convection triggering only when CAPE exceeds 600J/kg, the simulation of the MJO is much improved.

Diagnosis of these integrations (not illustrated) shows that large-scale rain dominates over con-

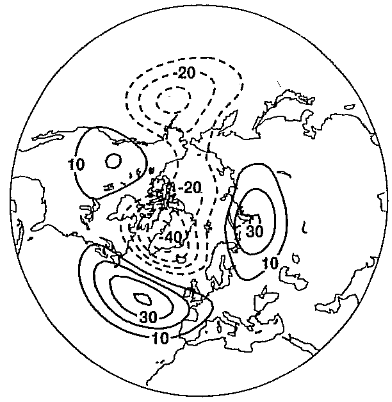

Figure 8: Geographical patterns of the most populated hemispheric circulation regime in the period 1971-1994, based on a regime analysis of Corti et al, 1999. Based on 500hPa height. Contour interval 10m. Compare with the systematic error field in Fig 1.

vective rain in the tropics with the 600J/kg CAPE threshold. Hence, the explicit dynamics is playing a significant role in releasing the convective instability and generating kinetic energy when the convection parametrisation scheme is suppressed. Indeed diagnosis of these results shows that excessive divergent kinetic energy is aliased onto the model grid, creating an overly strong Hadley circulation, for example. Conversely when the threshold for triggering the convective parametrisation is low, convective rain dominates over large-scale rain, and the parametrisation adjusts the convectively-unstable temperature profiles back to neutrality without any explicit production of kinetic energy. One could speculate that if, instead of relying on the CAPE threshold to generate convectively-forced balanced kinetic energy, the model was stochastically-forced (in particular with stochastic PV-dipole forcing, see section 6 below) especially in the warm-pool region where MCCs are common, a more satisfactory simulation of both the time-mean flow and MJO variability would ensue.

One can readily speculate about further possible upscale effects. For example, as suggested by Moore and Kleeman (1999), the (initial-time) singular vectors of the El Niño/ Southern Oscillation event (ENSO) have a strong projection onto the MJO, suggesting that models with a poor representation of the MJO may also have an excessively weak and/or regular El Niño climatotogy. Again, because El Niño itself is a nonlinear phenomenon (Münnich et al 1991), an excessively weak or overly-periodic ENSO may give rise to a systematic error in tropical Pacific SST. Finally, note that extratropical hemispheric-mean surface temperature is dependent on tropical Pacific SSTs (eg Palmer, 1996). The dynamical coupling of this speculative but plausible upscale error cascade is reminiscent of the poem at the beginning of the paper; the nail represents stochastic parametrisation, the shoe is the MCC, the horse is the MJO, the rider is ENSO, the battle is global-mean temperature, the kingdom our credibility!

6 Potential-vorticity and stochastic parametrisation

The specific representations, as given in equations 19 and 20 are of a rather restricted form, where only the tendencies associated with conventional parametrisations are perturbed. Is this a physically-justified restriction?

Consider for example the MCC. A characteristic of such systems is the balanced mesoscale circulation field which exist on scales much larger than the component cumulonimbi. These circulations can be conveniently described in terms of PV: cold upper level anticyclonic perturbations associated with a region of near-zero PV, and warm mid-level convergent cyclonic perturbations associated with a significant positive PV anomaly (Shutts and Gray, 1994). The existence of such PV anomalies is in part a consequence of the generation of balanced kinetic energy associated with horizontal variations in convective heating. On the basis of high resolution modelling in which 3-D convection is explicitly simulated, Shutts (1997) finds that if the mass convected in an MCC is equal to M_c, then the balanced energy generated in the MCC is proportional to $M_c^{5/3}$. Hence, for a given mass M_c being convected in a grid box of say 100km square, the ratio of balanced energy produced by a single MCC, to the balanced energy produced by an ensemble of say 100 convecting elements , is of the order of 10^2. On the basis that conventional parametrisation schemes are designed to describe the latter situation (where the convective kinetic energy is implicitly assumed to be dissipated within a grid box), such schemes may misrepresent situations of mesoscale organisation, missing an potentially important PV-source of upscale-cascading kinetic energy.

Gray (1999) has attempted to quantify the impact of such MCC PV forcing in a series of forecast experiments with the UK Meteorological Office weather prediction model. A control forecast is run using the operational weather prediction model, and a second forecast is run using an initial analysis which has been modified to include PV anomalies associated with MCCs diagnosed from satellite imagery. The PV anomalies were determined by idealised conceptual models and observational studies. Results were either significantly positive in terms of reducing forecast error, or at worst neutral.

Sensitivity studies reported by Gray (1999) suggest that the mid-level positive PV anomaly has a larger impact on forecast evolution than the upper-level negative PV anomaly. This is consistent with the singular vector analysis of Molteni and Palmer (1993) and Buizza and Palmer (1995) who show that the large-scale extratropical flow is sensitive to perturbations near the baroclinic steering level. Such perturbations can propagate vertically and lead to rapid energy growth in the upper tropospheric. Although the dominant singular vectors can have strong baroclinic tilt, Badger and Hoskins (1999) have shown that a monopole low-level PV perturbation which itself has no such tilt, can nevertheless have sufficient projection onto these rapidly-growing strucures to give impressive growth characteristics.

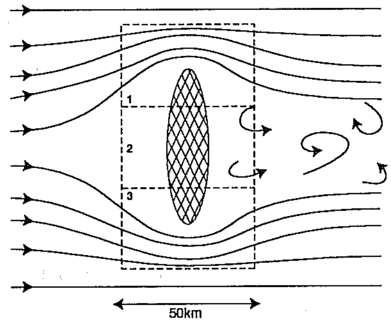

Figure 9: A schematic illustration of flow around topography with mesoscale structure but coherence across model grid boxes (shown dashed). Gravity-wave orographic parametrisation would diagnose sub-grid orographic variance in boxes 1, 2 and 3, but erroneously apply a drag to the grid box mean flow in boxes 1 and 3. A nonlocal and possibly stochastic parametrisation based on a horizontal PV dipole forcing might generate a more realistic response. From Palmer, 1997.

Non-local PV forcing may also be relevant to the problem of partially-resolved orography in global weather and climate models (consider Greenland, for example). Fig 9 shows, schematically, two dimensional flow around a poorly-resolved obstacle. With inviscid flow, a dipole pair of vortex sheets is created at the fluid boundary on either side of the obstacle. With viscosity, the vorticity perturbations can advect and diffuse outwards in the wake behind the cylinder (specific examples in the case of Greenland are shown in Doyle and Shapiro, 1999). If the obstacle is only partially resolved, as in Fig 9, then this PV dipole will be represented, both by the explicit dynamics, and by the sub-grid orographic parametrisations. For example, orographic gravity-wave drag (in which sub-grid scale orographic variance exists in all three grid boxes) will apply a negative tendency to the grid-box mean flow in all three grid boxes (eg Palmer et al, 1986). However, in reality the flow in grid boxes 1 and 3 is enhanced (compared with an unperturbed flow) as it moves past the obstacle ('tip jets', as discussed in Doyle and Shapiro, 1999). Similar to the convective case, the effect of such mesoscale topographic organisation might be best represented by some nonlocal PV dipole forcing, but oriented in the horizontal rather than the vertical. In the case of orography, it is not obvious that such PV dipoles should be represented stochastically; unlike convection, the orography itself is precisely known. Nevertheless, it is well known that the response to mesoscale orography can depend extremely sensitively on the direction of the large-scale upstream flow. In these circumstances, the strength of the imposed PV dipole might indeed be only representable as a pdf.

The downstream spreading of PV anomalies associated with localised orographic forcing and ambient downstream wind shear, is a process which can generate an upscale transformation of PV and associated wind field. Similar downstream and upscale spreading effects associated with convectively-forced PV have been suggested by Lilly et al (1998) as contributing to the observed $k^{-5/3}$ spectrum as discussed in section 3.

7 Dynamical approaches to stochastic parametrisation

From a dynamical systems perspective, one can imagine the scales above and below the truncation scale as represented by two coupled dynamical nonlinear systems, $S_{\leq N}$ and $S_{>N}$. The system $S_{\leq N}$ is given by the finite-N Galerkin representation of equation 1. Based on the discussion above, it is proposed to represent $S_{>N}$ as a simple dynamically-based stochastic system coupled to $S_{\leq N}$ over a range of scales, rather than as a 'lifeless' bulk formula depending on $S_{\leq N}$ only at the truncation scale. The weakness of coupling the parametrised processes to the resolved dynamics at precisely the truncation scale has already been exposed by Lander and Hoskins, 1997. Based on the PV perspective outlined above, we consider two complementary approaches which make more explicit use of the underlying dynamics in the formulation of $S_{>N}$. The first follows the approach used at ECMWF to initialise medium-range ensembles, the second utilises a cellular automaton approach to parametrisation. Consistent with the comments made in section 2 regarding the Whitney/Takens theorem, these stochastic-dynamic parametrisations are necessarily nonlocal.

7.1 Stochastic forcing in singular-vector space

As discussed in section 2, a number of geophysical phenomena have been modelled using a background stochastic forcing to excite the singular vectors of a stable but non-normal linearised operator. This raises the possibility that the stochastic forcing could be directly computed from the relevant singular vectors. This idea is closely related to philosophy used in the ECMWF ensemble prediction system (Palmer 1993; Molteni et al, 1986) to generate initial perturbations.

It is worth recalling the philosophy used to justify the singular vector strategy for generating initial EPS perturbations (see also Palmer, 1999). In principle, an unbiased ensemble of initial states should ideally be created by randomly sampling the initial pdf $\rho_i(X, t = 0)$ which determines the probability that the true state lies in some neighbourhood of a point X in phase space. There are two related problems that complicate such a procedure. Firstly, ρ_i is not well known; there are many assumptions in data assimilation (eg in quality control, representativity of observations, the tangent approximation, the role of model error in the data assimilation process and so on) whose contribution to ρ_i is not well quantified. Secondly, for a contemporary weather prediction model, if $\rho_i(X, t = 0)$ is Gaussian, then $O(0^{14})$ numbers are needed to specify it, orders of magnitude more than the maximum available sample size. In practice, a poorly-known and inadequately sampled pdf can lead to an underdispersive ensemble and overconfident probability forecasts.

A similar problem arises for the problem of stochastic forcing. A pdf $\rho_m(\dot{X}, t)$ can be defined giving the probability that the actual grid-box tendency lies within some neighbourhood of the model tendency $\dot{X}$. We have argued, for example, that this pdf should be relatively broad in circumstances where mesoscale organisation is likely to occur. However, this hardly fixes the actual pdf in any precise sense. Hence, like the initial pdf ρ_i, the sub-grid pdf ρ_m is not well known.

It is in principle straightforward to define a strategy for determining a stochastic sampling of ρ_m in the space of dominant singular vectors of $\mathcal{M}$ (see equation 16) where Δt is the autocorrelation time associated with meso-scale sub-grid processes. A metric can be used to constrain the (right) singular vectors to regions where the parametrised tendencies are large (in the same way that the initial ensemble perturbations are weighted towards regions where observation errors are likely to be large), and to scales close to the truncation scale. For each Δt, the stochastic forcing perturbation can be taken as a random linear combination of these singular vectors. Since the forcing singular vectors have direct projection onto all the model variables, including vorticity and temperature, they are broadly consistent with the PV-thinking approach discusssed in section 6. (For example, as noted above, there is evidence that the lower tropospheric component of the PV dipole associated with extratropical MCCs may well have significant projection onto baroclinic energy-metric singular vectors.)

7.2 A cellular automaton model

An alternative approach to the singular vector method discussed above, would be to try to model ρ_m more explicitly. One possible type of dynamical system on which to base ρ_m is the cellular automaton (CA) from which coherent structures (like tesselating hexagons reminiscent of organised Rayleigh-Bénard convection) can be readily produced (see eg Adamatsky, 1996). The CA model, first applied by von Neumann to biological problems, is a dynamical system with a state vector which takes on a number of discrete (often just two) states. This CA state vector is defined on a discrete grid of points in space and time. There is a rule (either deterministic or stochastic) which determines the state at some space point, as a function of the state of the CA at surrounding points, and at the concurrent and earlier times.

It is possible that a stochastic CA model could form the basis of a simple representation of the MCC; indeed such an approach has already been suggested by Randall and Huffman (1980). Fig 10 is a snapshot of an example of an extremely simple CA representation of mesoscale organisation (Palmer 1997). The dashed-line grid is presumed to be equivalent to the grid of a climate or weather prediction model. The CA grid is much finer than the climate model grid, and the CA is presumed to have bivalent states, black representing a convecting (or 'on') state, and white representing a non-convecting state. The CA grid could possibly be initialised from high resolution satellite imagery, otherwise by some random seed. A rule is needed to determine whether the CA is 'on' or 'off' at the next CA timestep, much shorter than the GCM timestep. In Fig 10, the probability of the CA being 'on' depends both on the large-scale CAPE (interpolated to a particular CA grid point) and the number of adjacent CAs which were 'on' at the previous timestep. For given CAPE, isolated 'on' cells represent individual cumulonimbi and have relatively short lifetimes, whilst the aggregates of 'on' cells, representing MCCs, have

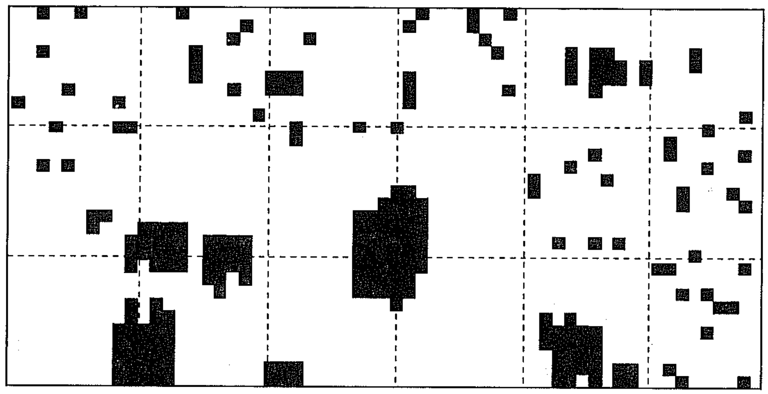

Figure 10: A snapshot of a cellular automaton model of organised convection. The dashed-line grid is presumed to be equivalent to the grid of a climate or weather prediction model. The probability of a cell remaining 'on' (shown black) would depend on the large-scale CAPE, and the number of adjacent 'on' cells. The CAs would feed PV back to the resolved-scale flow. The strength of a PV-dipole forcing would depend nonlinearly on the number of adjacent 'on' cells. From Palmer 1997

relatively long lifetimes. In addition to the evolution rules described above, the CA cells can be made to advect (across model grid-box boundaries) with the resolved-scale wind, and the ability to organise could be made dependent on the resolved- scale wind shear.

The feedback of the CA onto the large-scale flow could be based on the nonlinear formula of Shutts (1997) discussed in section 6. If we imagine that each 'on' state is convecting a mass M, then the 'blobs' with N connected elements represent an organised MCC convecting a mass NM. From the discussion in section 6, this determines a specific PV forcing onto the large-scale flow (proportional to $(NM)^{5/3}$). The total PV forcing field can be put through PV inversion software to determine the wind and temperature forcing at each GCM gridpoint (as was done by Gray, 1999).

Similar ideas could be applied to the orography problem. For example, let h be the mean height of the orography within a CA cell. The probability of a CA element being in an 'on' or blocked state would depend on Nh/U, where N and U are resolved-scale values of static stability and wind speed, interpolated to a CA grid point. A PV forcing dipole would be centred on the centroid of aggregates of 'on' CAs, representing mesoscale orography. The strength of the PV dipole would depend on the number of connected 'on' cells of an aggregated cellular automata, its width would depend on the dimensions of the aggregate 'on' blob, and the dipole would be oriented relative to the large-scale flow.

8 Discussion

On the odd occasions when meteorologists and economists get together, prediction techniques used in the two discliplines are often compared. At some stage, the meteorologists will raise the fact that whilst the governing equations of climate are known from underlying theoretical principles, the equations which describe the economy are only known from empirical analysis.

Whilst this may well be an excellent debating ploy, there is some dishonesty in making too much of this point. Whilst we know Newton's laws of motion extremely well, there is no unique prescription for representing the governing equations of climate computationally, since the process of parametrisation is not a rigourously (or even heuristically) justifiable procedure in regions of mesoscale organisation. Obviously uncertainties in parametrisation in no way invalidate meteorological prediction, after all global weather and climate models over the last 50 years have been one of the most important and fruitful products of our field of research. However, these uncertainties may impede future development, unless they are recognised explicitly.

On the other hand, it could be argued that once models have sufficient resolution, then the effect of such uncertainties on scales of interest will be minimal. However, arguments presented in this paper suggest that under-representation of subgrid variability could have an impact on large-scale systematic error, for any foreseeable resolution. (In any case, climate and weather forecasts must include estimates of uncertainty, and this requires significant utilisation of computer time for producing forecast ensembles. This requirement will limit the extent to which very high resolution climate integrations are possible.)

The real problem is that climate is a complex nonlinear system with many interacting scales. The generic procedure of truncating the equations to some 'hard' limit, and parametrising the unresolved scales as local deterministic bulk formulae depending on the resolved scales (at the truncation limit) may well be an underlying factor for why some model systematic errors (see Fig 1) have so stubbornly resisted upgrades in model resolution and parametrisation complexity. A case has been put for the generalisation of parametrisation schemes from local deterministic systems, slave to the large scale flow at precisely the truncation scale, to stochastic nonlocal nonlinear dynamical systems, weakly coupled to the large-scale flow over a range of scales. Such systems would not 'die' if the variability from the large-scale flow is held fixed; they have internal variability of their own. The weakness of coupling the parametrised processes to the resolved dynamics at precisely the truncation scale has already been exposed by Lander and Hoskins, 1997. The migration away from a purely deterministic approach to modelling, puts further emphasis on the importance of recognising that all meteorological prediction problems, from weather forecasting to climate-change projection, are essentially probabilistic.

Acknowledgements

I am very grateful to Kamal Puri (BMRC and ECMWF), Frank Selten (KNMI) and Frederick Vitart (ECMWF) for supplying me with unpublished material (Figs 4, 2 and 7 respectively) for this paper; to David Orrell and Lenny Smith (University of Oxford) for many helpful discussions on the applications of dynamical systems theory to the parametrisation problem; to Laura Ferranti, David Gregory and Christian Jakob (ECMWF) for helpful discussions on convective parametrisation and the Madden-Julian Oscillation; to Roberto Buizza (ECMWF), Martin Miller (ECMWF), Franco Molteni (ICTP) and Brian Hoskins (University of Reading) for discussions for devlopment for stochastic physics schemes, and to Paul Mason and Glenn Shutts (UK Meteorological Office) for discussions on stochastic backscatter and potential vorticity generation by mesoscale convection.

References

Adamatzky, A., 1994: Identification of cellular automata. Taylor and Francis, London, 337pp.

Badger, J. and B.J.Hoskins, 1999: Simple initial value problems and mechanisms for baroclinic growth. J. Atmos.Sci., to appear.

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.Meteorol.Soc., 112, 693-709.

Bohr, T., M.H.Jensen, G. Paladin and A. Vulpiani, 1998: Dynamical Systems Approach to Turbulence. Cambridge Nonlinear Science Series 8. Cambride University Press.350pp

Brankovic, C. and T.N.Palmer, 2000: Seasonal skill and predictability of ECMWF PROVOST ensembles. Q.J.R.Meteorol.Soc., to appear.

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

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

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. Horizontal wavenumber spectra of winds, temperature and trave gases during the Pacific Exploratory Missions: 1. Climatology. J.Geophys.Res., 104, 5697-5716.

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.

DelSole, T. M., and B. F. Farrell 1995:A stochastically excited linear system as a model for quasigeostrophic turbulence: analytic results for one- and two-layer fluids. Journal of the Atmospheric Sciences, 52, 2531-2547.

Dodson, C.T.J, and T. Poston, 1979: Tensor Geometry. Pitman, London. 598pp

Doyle, J. and M.A.Shapiro, 1999: Flow response to large-scale topography: the Greenland tip jet. Tellus, 51A, 728-748.

Evans, R.E., R.J.Graham, M.S.J.Harrison and G.J.Shutts, 1998: Preliminary Investigations Into The Effect Of Adding Stochastic Backscatter To The Unified Model. Forecasting Research Technical Report No. 241. Meteorological Office, London Road, Bracknell, Berkshire, RG12 2SZ, UK. 11pp.

Farrell, B.F., and P. J. Ioannou, 1993:Stochastic Dynamics of Baroclinic Waves. Journal of the Atmospheric Sciences ,50, 4044-4057.

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

Farrell, B.F., and P. J. Ioannou, 1999a: Perturbation Growth and Structure in Time Dependent Flows. J. Atmos. Sci., 56, 3622-3639.

Farrell, B.F., and P. J. Ioannou, 1999b: Approximating an Autonomous Stable Linear Dynamical System by a System of Lower Order. J. Atmos. Sci., to appear.

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.

Ghil, M. and S. Childress, 1987: Topics in geophysical fluid dynamics: atmospheric dynamics, dynamo theory and climate dynamics. Applied Mathematical Sciences, 60. Springer-Verlag 485pp

Gulick, D., 1992: Encounters with Chaos. McGraw-Hill, 308pp

Gray, M.E.B., 1999: Potential vorticity anomalies in mesoscale convective systems and their impact on forecast evolution. Met O (APR) Turbulence and Diffusion Note No. 261.United Kingom Meteorological Office. Bracknell. Berks. UK

Harrison, M., T.N.Palmer, D.S.Richardson and R. Buizza, 1998: Analysis and model dependencies in medium-range ensembles: two transplant cast studies. Q.J.R.Meteor.Soc., 125, 2487-2515.

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

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.

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.

Koshyk, J.N., B.A.Boville, K.Hamilton, E.Manzini and K.Shibata, 1999a: Kinetic energy spectrum of horizontal motions in middle-atmosphere models. J. Geophys.Res., 104, 27177- 27190.

Koshyk, J.N., K. Hamilton and J.D. Mahlman, 1999b: Simulation of the $k^{-5/3}$ mesoscale spectral regime in the GFDL SHYHI general circulation model. Geophys. Res. Lett., 26, 843-846.

Koshyk, J.N. and K. Hamilton, 2000: The horizontal kinetic energy spectrum and spectral budget simulated by a high-resolution troposphere-stratosphere GCM. J. Atmos.Sci., submitted.

Lander, J. and B.J.Hoskins, 1997: Believable scales and parametrisations in a spectral transform model. Mon.Wea.Rev., 125, 292-303.

Leith, C.E., 1990: Stochastic backscatter in a subgrid-scale model: plane shear mixing layer. Phys.Fluids A, 2, 297-299.

Lilly, D.K., 1973: Lectures in sub-synoptic scales of motions and two-dimensional turbulence. Dynamical Meteorology. ed. P.Morel. Boston: Reidel. p353-418.

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

Lilly, D.K., G. Bassett, K. Droegemeier and P. Bartello, 1998: Stratified turbulence in the atmospheric mesoscales. Theoret.Comput.Fluid Dynamics, 11, 139-153.

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., 1996. Predictability - a problem partly solved. In 'Predictability'. ECMWF, 1995 Seminar Proceedings, Shinfield Park, Reading, RG2 9AX, UK.

Mapes, B.E., 1997: Equilibrium vs. activation control of large-scale variations of tropical deep convection. In 'The physics and parameterization of moist atmospheric convection' p321-358. ed. R.K.Smith. Kluwer Academic Publishers. Netherlands.

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

Molteni, F. and Tibaldi, S., 1990: Regimes in the wintertime circulation over northern extratropics. II: Consequences for dynamical predictability. Q.J.R.Meteorol.Soc., 116, 1263-1288.

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

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

Moore, A.M. and R. Kleeman, 1999: Stochastic forcing of ENSO by the intraseasonal oscillation. J.Clim., 12, 1199-1220.

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.

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

Newman, M., P. D. Sardeshmukh and C. Penland, 1997: Stochastic forcing of the wintertime extratropical flow. J. Atmos. Sci, 54, 435-455.

Orrell, D., 1999: A shadow of a doubt: model error, uncertainty and shadowing in nonlinear dynamical systems. Thesis. University of Oxford, Mathematical Institute (OCIAM), St Giles, Oxford.

Palmer, T.N., 1988: Medium and extended-range predictability, and stability of the PNA mode. Q.J.R.Meteorol.Soc., 114, 691-713.

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

Palmer, T.N., 1996: Predictability of the atmosphere and oceans: from days to devades. NATO Advanced Study Institute Series, SubseriesI 'Global Environmental Change'. Vol 44 'Decadal Climate Variability: Dynamics and Predictability'. Eds D.Anderson and J. Willebrand. Springer.

Palmer T.N, 1997: On parametrizing scales that are only somewhat smaller than the smallest resolved scales, with application to convection and orography. 1996 ECMWF workshop on convection. ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

Palmer, T.N., 2000: Predicting uncertainty in forecasts of weather and climate. Rep.Prog.Phys., 62, 1-47.

Palmer, T.N., G.J.Shutts and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrisation. Q.J.R.Meteorol.Soc., 112, 1001-1031.

Palmer, T.N., F.Molteni, R.Mureau, R.Buizza, P.Chapelet and J.Tribbia, 1993: Ensemble prediction. 1992 ECMWF Seminar proceedings.
ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

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.

Penland, C., 1996: A stochastic model of IndoPacific sea surface temperature anomalies, Physica D, 98, 534-558.

Puri, K., J. Barkmeijer and T.N.Palmer, 2000: Ensemble prediction of tropical cyclones using targeted singular vectors. Q.J.R.Meteor.Soc., submitted.

Randall, D.A. and Huffman, G.J., 1980: A stochastic model of cumulus clumping. J.Atmos.Sci., 37, 2068-2078.

Sauer, T., J.A.Yorke and M. Casdagli, 1991: Embedology. J.Stat.Phys., 65, 579-616

Schertzer, D. and S. Lovejoy, 1993: Nonlinear variability in geophysics 3: scaling and multifractal processes. Lecture Notes. Institut D'Etudes Scientifiques de Cargése. 292pp

Selten, F.M., 1995: An efficient empirical description of large-scale atmospheric dynamics. Ph.D. Thesis. Vrije Universiteit te Amsterdam. 169pp.

Shutts, G.J. and Gray, M.E.B., 1994: A numerical modelling study of the geostrophic adjustment process following deep convection.
Quart.J.Roy.Meteorol.Soc., 120, 1145-1178.

Shutts, G.J., 1997: A balanced flow perspective on the effect of deep convection. 1996 ECMWF workshop on convection. ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

Smith, L.A., 2000: Disentangling uncertainty and error: on the predictability of nonlinear systems. Nonlinear dynamics and statistics, A.Mees(ed.). Birkhauser, 2000.

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.

Takens, F., 1981: Detecting Strange Attractors in Turbulence. In Dynamical Systems and Turbulence edited by D.A. Rand and L.-S. Young. Lecture Notes in Mathematics. Vol 898, 336-381. Springer-Verlag.

VanZandt, T.E., 1982: A universal spectrum of buoyancy waves in the atmosphere. Geophys. Res. Lett., 9, 575-578.

Wallace, J.M. and D.S. Gutzler, 1981: Teleconnections in the geoptential height field during the northern hemisphere winter. Mon.Wea.Rev., 109, 784-812.

Wallace, J.M., Zhang, Y. and Bajuk, L., Interpretation of interdecadal trends in northern hemisphere surface air temperature. J.Clim., 9, 249-259.

Wang, W. and M.E.Schlesinger, 1999: The dependence on convection parameterization of the tropical intraseasonal oscillation simulated by the UIUC 11-layer atmospheric GCM., J.Clim., 12, 1423-1457.

WMO, 1999: The 1997-1998 El Niño event: a scientific and technical perspective. WMO-No.905. World Meteorological Organization. Avenue de la Paix. PO Box 2300. CH-1211 Geneva-2. 96pp.