Predictability in dynamical systems relevant to climate and weather

Richard Kleeman
Center for Atmosphere Ocean Science
Courant Institute of Mathematical Sciences

Relevant Publications

R. Kleeman and A. M. Moore. A new method for determining the reliability of dynamical ENSO predictions. *Mon. Weath. Rev.*, **127**:694-705, 1999.

R. Kleeman. Measuring dynamical prediction utility using relative entropy. *J. Atmos Sci*, **59:**2057-2072, 2002.

M.K. Tippett, R. Kleeman, and Y. Tang. Measuring the potential utility of seasonal climate predictions. *Geophys. Res. Lett.*, **31**:L22201, 2004. doi 10.1029/2004GL020673.

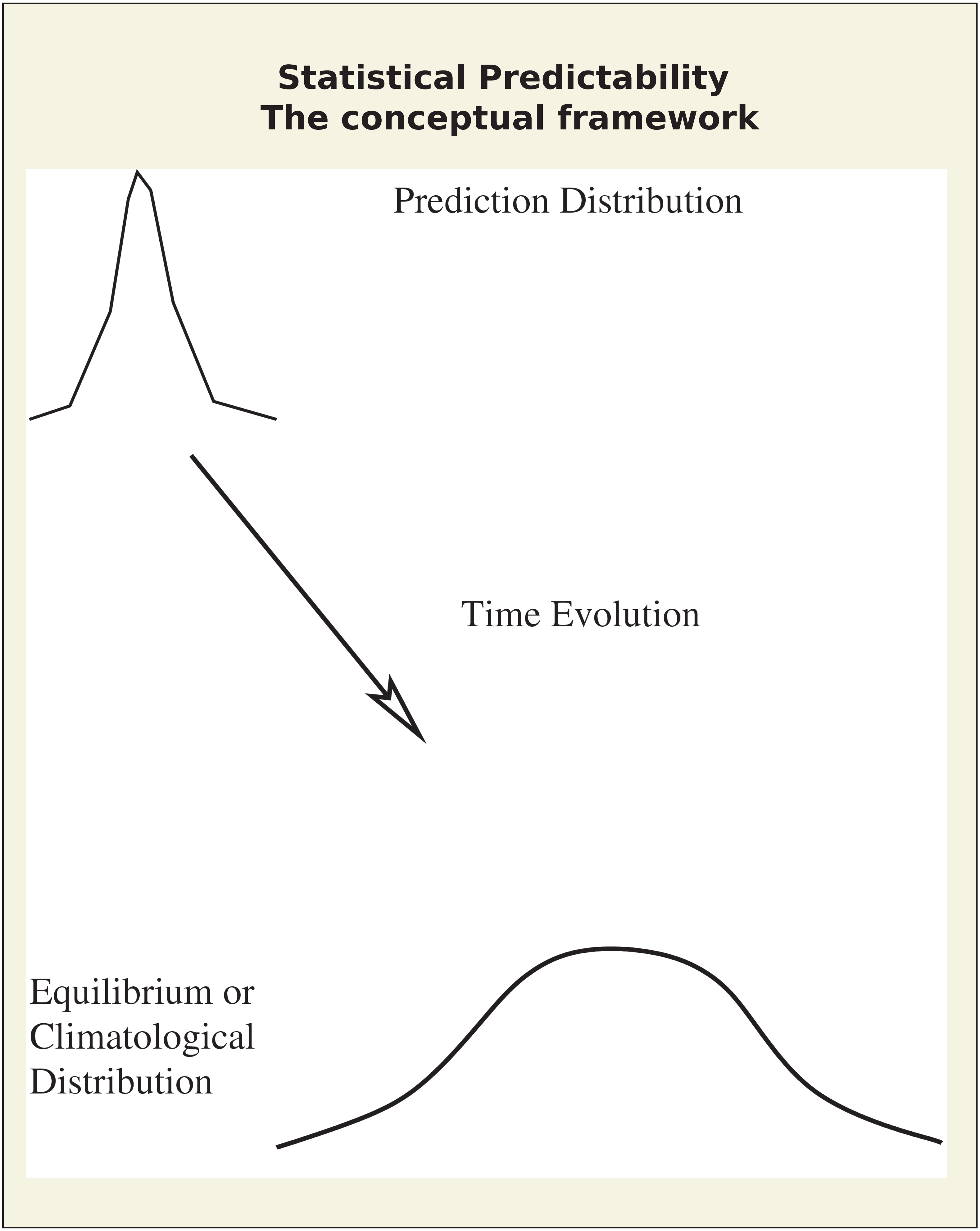
R. Kleeman and A. J. Majda. Predictability in a model of geostrophic turbulence. *J. Atmos Sci*, **62**:2864-2879, 2005.

R. Abramov, A.J. Majda, and R. Kleeman. Information theory and predictability for low frequency variability. *J. Atmos Sci*, **62**:65-87, 2005.

Y. Tang, R. Kleeman, and A.M. Moore. On the reliability of ENSO dynamical predictions. *J. Atmos Sci*, **62**:1770-1791, 2005.

R. Kleeman. Statistical predictability in the atmosphere and other dynamical systems. *Physica D,* 2005. Special issue on data assimilation. Revised.

R. Kleeman. Limits, variability and general behaviour of statistical predictability of the mid-latitude. *J. Atmos Sci*, 2006. In preparation.


Statistical Predictability
The conceptual framework
Prediction Distribution
Time Evolution
Equilibrium or
Climatological
Distribution

Bayesian Approach to Predictability

Without initial condition data or a dynamical projection of such information, the most reasonable assumption regarding target random variables is that they have the equilibrium distribution. This is thus the prior distribution in Bayesian terminology.

If the prediction data are then made available i.e. the initial condition distribution is deduced and projected using a dynamical system then this shifts the prior distribution to the posterior prediction distribution. The utility or information content of this shift is measured by the so-called relative entropy.

The relative entropy is also used in standard stochastic theory (it is called a Lyuponov functional there) to measure the convergence of the prediction to the equilibrium distribution. If the conditional distributions satisfy a causality condition i.e. if the future probability is uniquely determined by initial conditions, the relative entropy montonically declines (or is conserved).

$$D(p \parallel q) \equiv \sum_{x \in \mathcal{H}} p(x) \ln \left(\frac{p(x)}{q(x)} \right)$$

$$D(f \| g) \equiv \int_{all\ R^n} f(\overrightarrow{z}) \ln \left(\frac{f(\overrightarrow{z})}{g(\overrightarrow{z})} \right) d\overrightarrow{z}$$

The discrete form approaches the continuous form as partitions of state space are reduced.

Relative Entropy Properties

1. Suppose we have two probability densities f and g then $D(f\|g) \geq 0$ with equality if and only if $f = g$ almost everywhere (i.e. at points where $f \neq 0$).

2. Suppose we define a general non-linear transformation of our state space: $F : R^n \rightarrow R^n$which is non-degenerate i.e. $det(J(F)) \neq 0$ where J is the Jacobian, then the relative entropy of the transformed probability densities is left invariant.

3. Suppose we have two probability densities from two realizations of the same random process $F = F(x, t, x', t')$ and $G = G(x, t, x', t')$ with $x, x' \in R^n$ and $t, t' \in R$ and let us assume that the following causality condition

holds for the associated conditional densities:

$$F(x, t \mid x', t') = G(x, t \mid x', t') \qquad where\ t \geq t' \tag{1}$$

then the associated marginal distributions $f(x, t)$, $f(x', t')$ and $g(x, t)$, $g(x', t')$ satisfy

$$D_t(f \parallel g) \leq D_{t'}(f \parallel g) \tag{2}$$

where by definition

$$f(x, t) \equiv \int_{R^{n+1}} F(x, t, x', t') dx' dt'$$

and similarly for G and g. Note if we consider a subspace of our state-space then equation (1) will in general NOT hold because the future probability will depend on variables at the initial time in the complement of the subspace.

Gaussian Formulae and Common Skill Measures

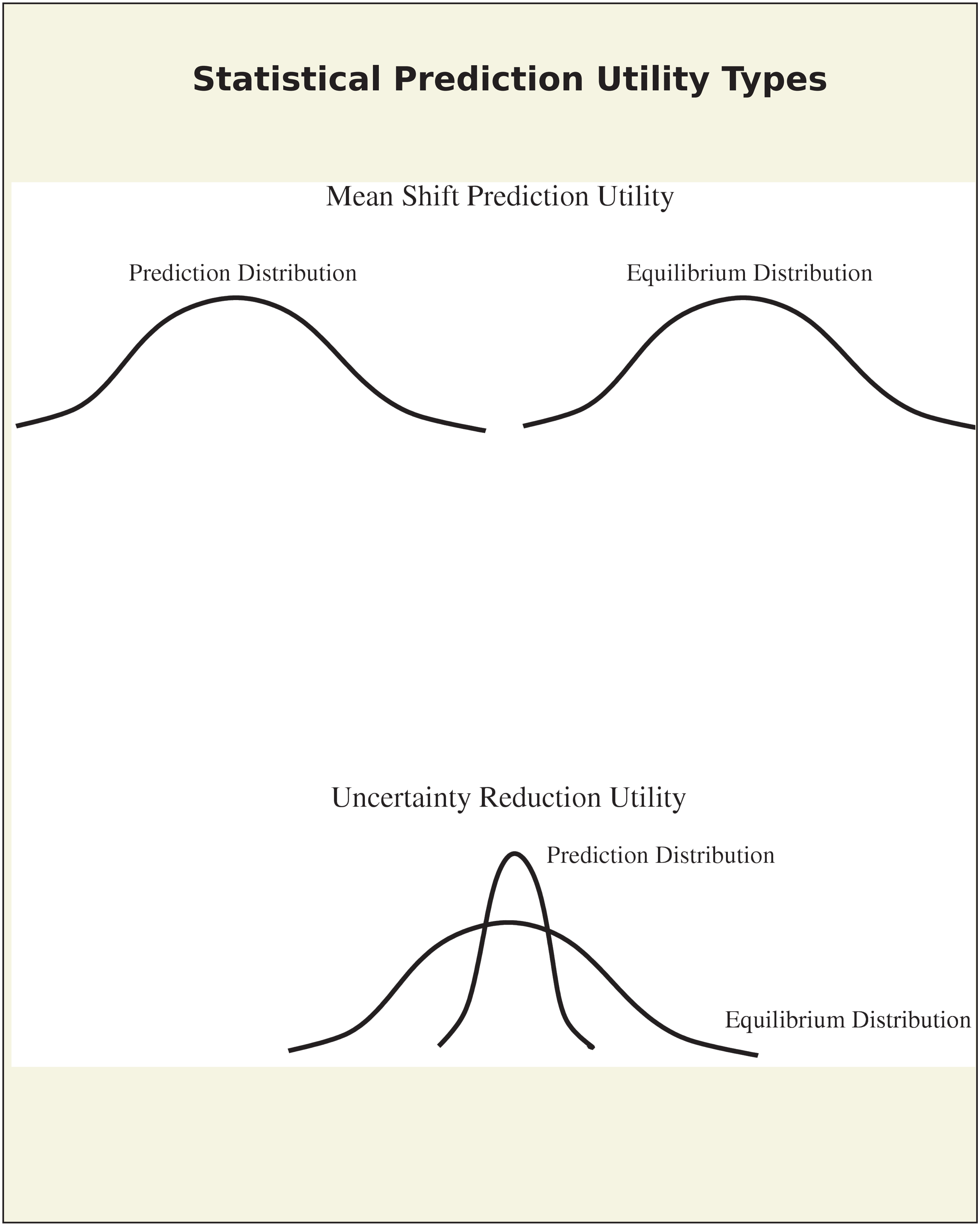
$$RE = Dispersion + Signal$$

$$Dispersion \;=\; \frac{1}{2}\left[\ln\left(\frac{\det(\sigma_q^2)}{\det(\sigma_p^2)}\right) + tr\left(\sigma_p^2(\sigma_q^2)^{-1}\right) - n\right]$$

$$Signal \;=\; \frac{1}{2}\left[(\overrightarrow{\mu_p})^t\,(\sigma_q^2)^{-1}\overrightarrow{\mu_p}\right]$$

$$RMS\;Error = \overline{tr(\sigma_p^2)}$$

$$Anomaly\;Correlation = \frac{\overline{(\overrightarrow{\mu_p})^t\,\overrightarrow{\mu_p}}}{\overline{(\overrightarrow{\mu_p})^t\,\overrightarrow{\mu_p}} + \overline{tr(\sigma_p^2)}}$$


Statistical Prediction Utility Types
Mean Shift Prediction Utility
Prediction Distribution
Equilibrium Distribution
Uncertainty Reduction Utility
Prediction Distribution
Equilibrium Distribution

Prediction Utility Variation

A question of prime importance to forecasters is the variation in predictability/utility from one forecast (or set of initial conditions) to the next. This may be studied by finding the causes of variation of relative entropy.

Since distributions are often (but not always) close to Gaussian the cause of variations may be strongly related to fluctuations in the Dispersion or fluctuations in the Signal.

Dispersion fluctuations. These are usually due to variations in the ensemble spread connected to changes in flow instability with initial condition.

Signal fluctuations. These are usually due to variations in the amplitude of (often persistent) anomalous modes. These may be due to random factors and have no particular dynamical cause.

Which particular fluctuation is most important is dynamical system dependent.

Specific Dynamical Systems Studied

Climate Systems

Stochastically Forced damped oscillator. Simple analog for ENSO.

Hybrid or Intermediate Coupled Models. Skilful ENSO prediction models.

Weather Systems

Lorenz 3 component system. Simple chaotic oscillator.

Baroclinic turbulence system.

Global primitive equation model with realistic meridional radiative forcing.

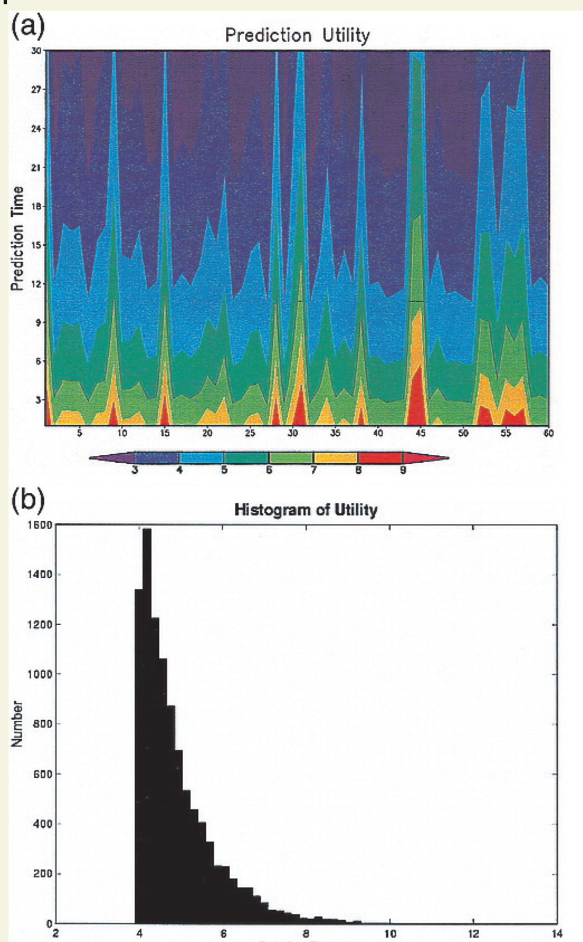
Stochastically forced oscillator

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}_t = \begin{pmatrix} 0 & 1 \\ \beta & \gamma \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} 0 \\ F \end{pmatrix},$$

F is a white noise forcing. The two components of u can be interpreted in an ENSO analog as the first two EOFs of tropical Pacific ocean heat content with the first being also highly correlated with the first EOF of SST. The matrix coefficients are given by

$$\gamma = -\frac{2}{\tau} \qquad \beta = -\left(\frac{4\pi^2}{T^2} + \frac{1}{\tau^2}\right).$$

where T is the period of the unforced oscillator and tau it's damping time. We assume firstly that these are fixed. If we assume a randomly chosen set of deterministic initial conditions drawn from a long run of the model the following relative entropy sample set (50) of evolutions and histogram for a particular time results:



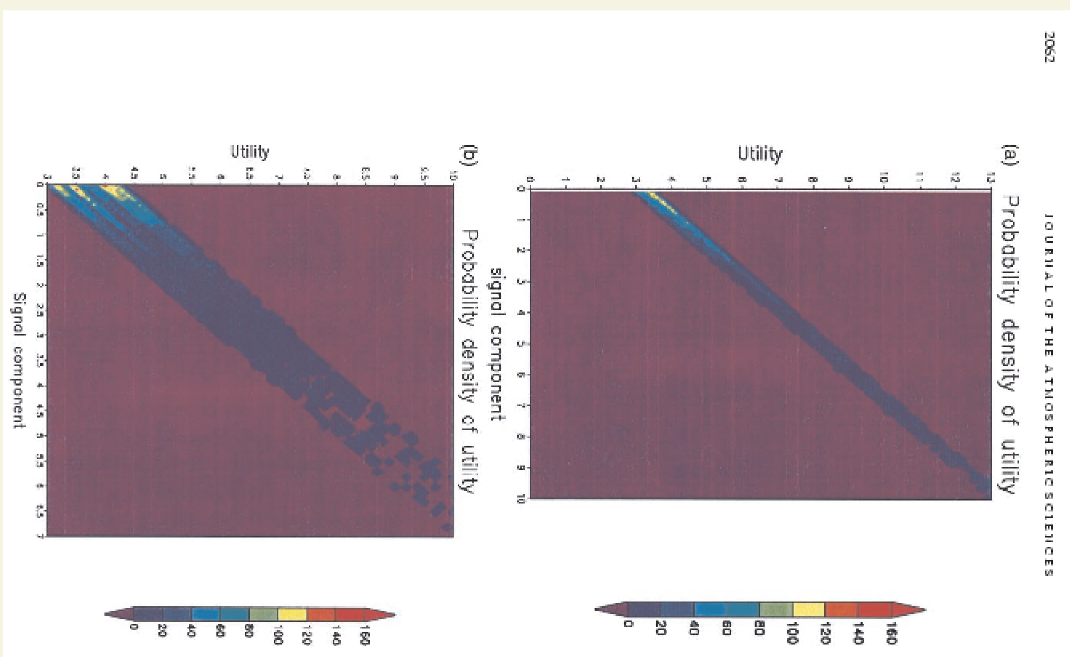
In such a model the distributions are all Gaussian and also the covariance matrix is independent of initial condition meaning that variations in utility/RE are completely determined by the Signal piece of the Gaussian relative entropy. Fluctuations in this are determined by random fluctuation in the amplitude of the vector u.

In general for a more realistic model of ENSO one might expect tau to vary according to the seasonal cyle or ENSO cycle. We can modify the model to allow for this:

$$\frac{1}{\tau} = \left(1 + 2\sin\frac{2\pi t}{P}\right)\Big/ T.$$

where P is the assumed period of the instability variation. Note that for certain parts of the instability cycle very strong unforced growth can occur.

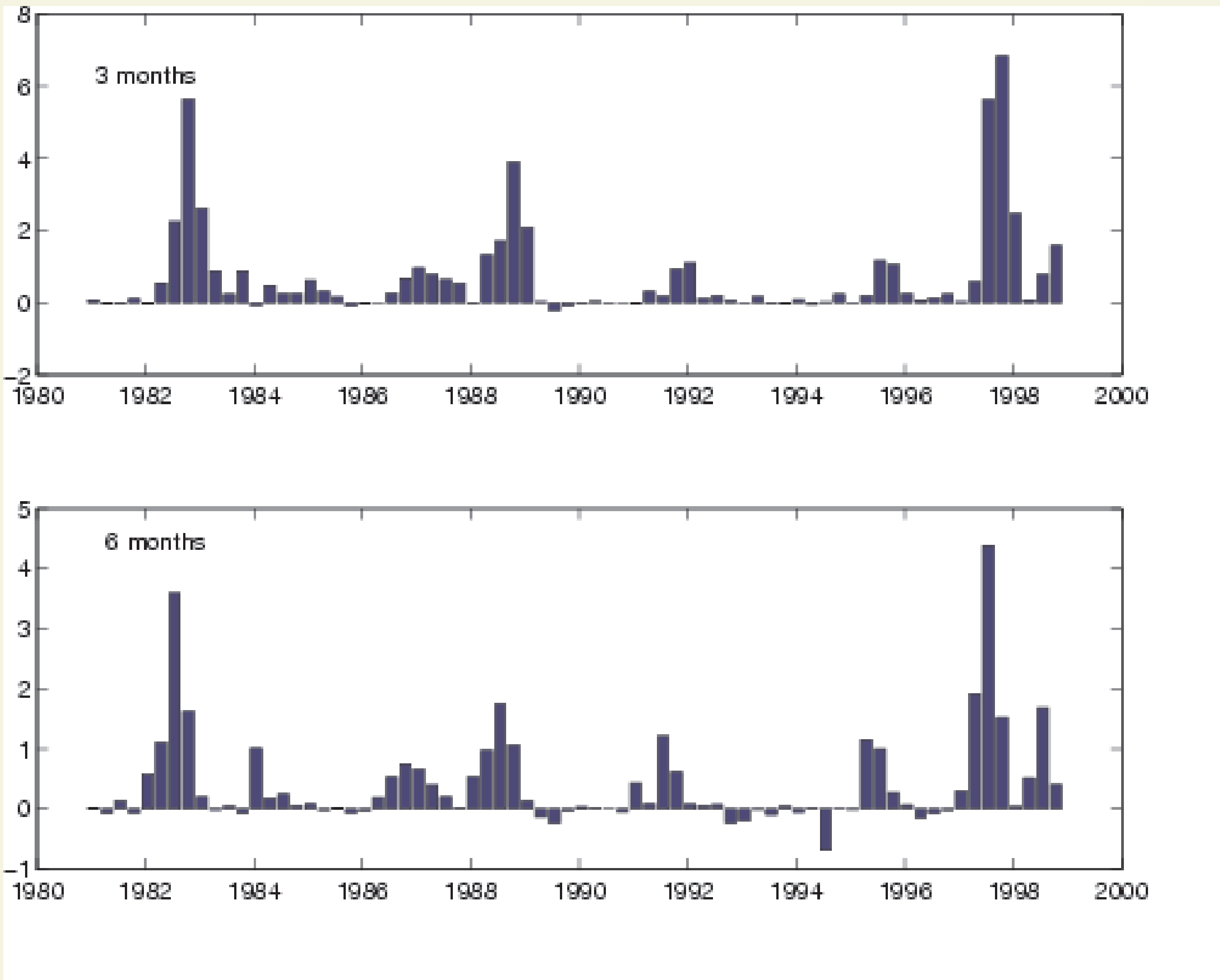
Despite this strong variation in instability, variations in utility are still dominated by the Signal piece of the Gaussian relative entropy. The right panel is for P=T=4 and the second is P=1 with T=4.



Hybrid and Intermediate Coupled Models of ENSO

Results to be shown are from a hybrid coupled model consisting of an OGCM coupled to a statistical atmospheric model. Results are qualitatively the same in three other coupled models, one hybrid and two intermediate. All models have roughly the same skill level which is close to many other coupled models used for routine ENSO prediction.

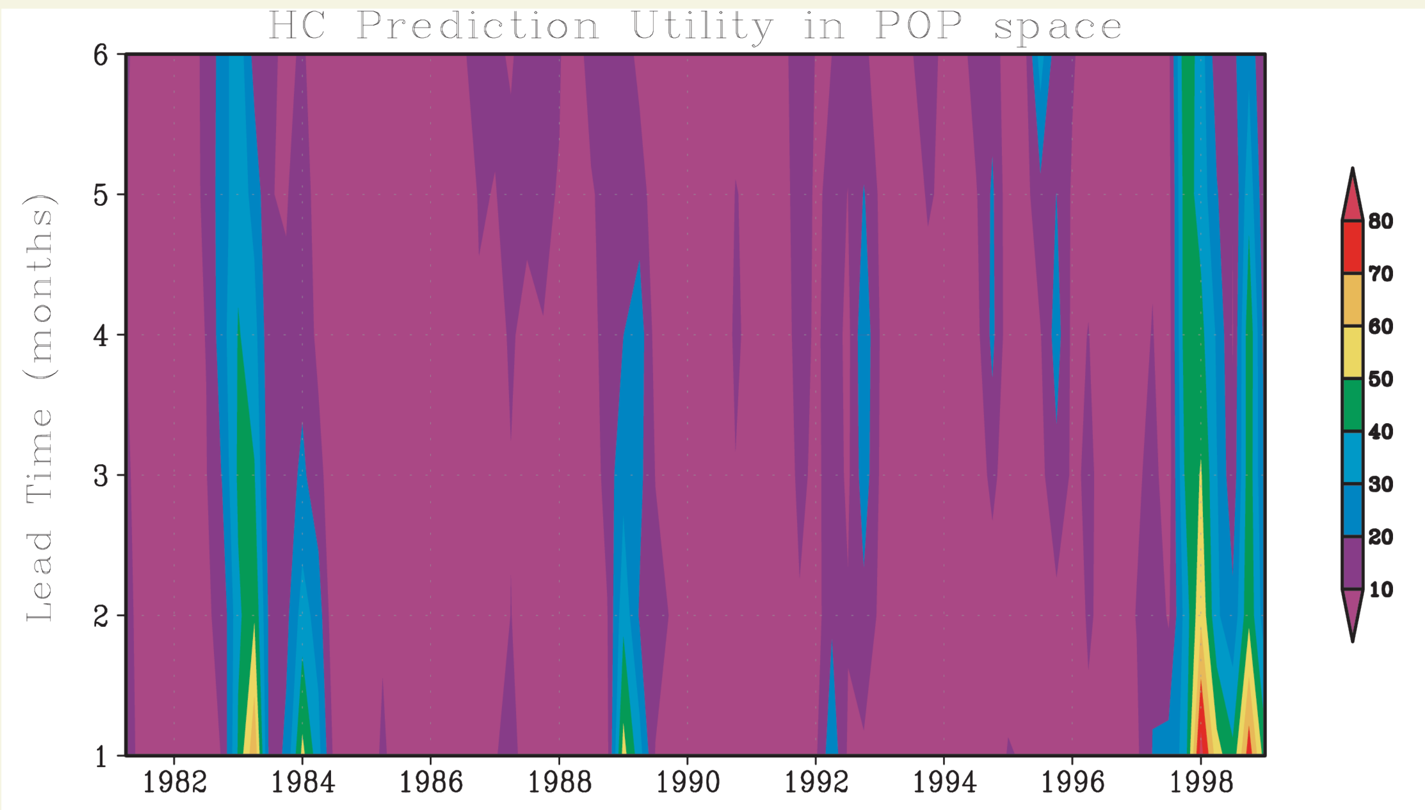
Contribution to anomaly correlation skill of particular forecasts



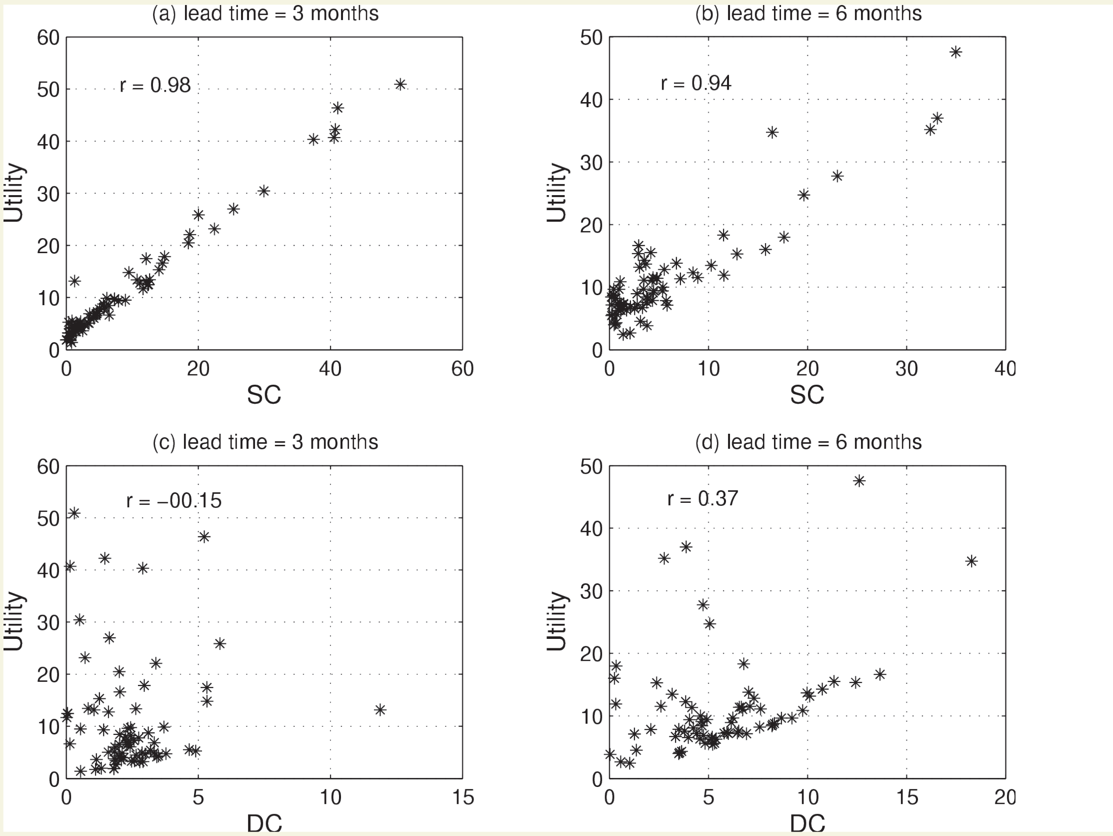
Methodology

Ensembles were created using stochastic forcing which was close to white on climate timescales and had a spatial pattern of the dominant stochastic optimal of the coupled model in climatological mode. A reduced state space was used for analysis consisting of the two components of the dominant POP of ocean heat content which corresponds approximately to the first two EOFs. The POPs and EOFs were obtained by a forced run of the OGCM and the latter explain around 80% of total heat content variance. Distributions on this two dimensional reduced state space were very close to Gaussian at all leads so we used the Gaussian RE formulae.

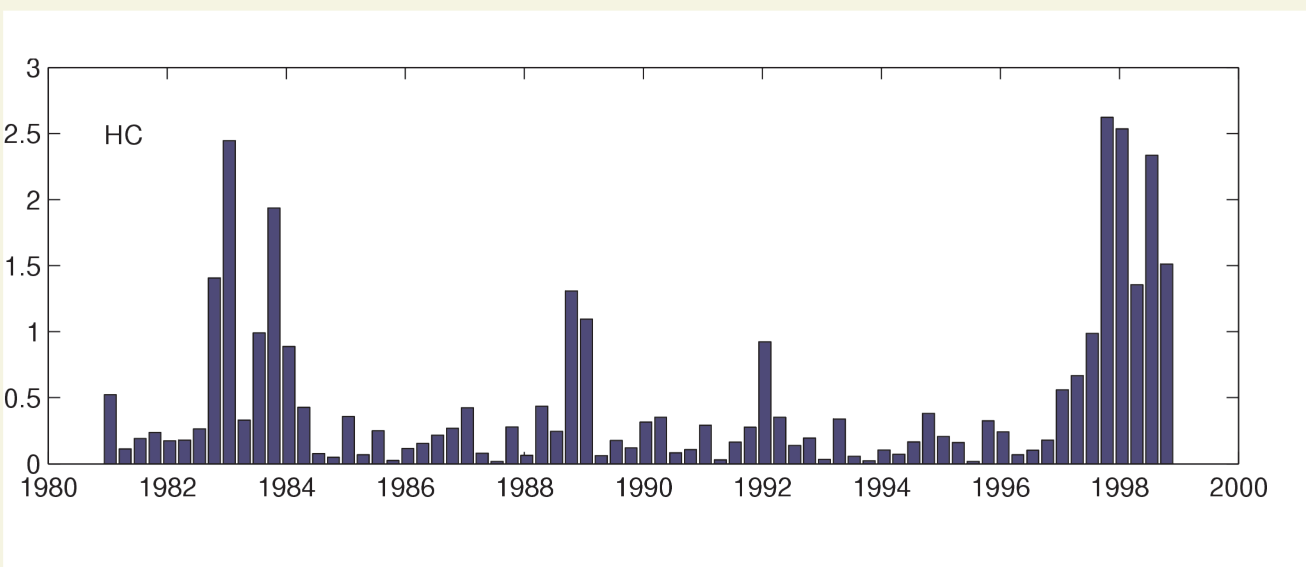
Relative Entropy



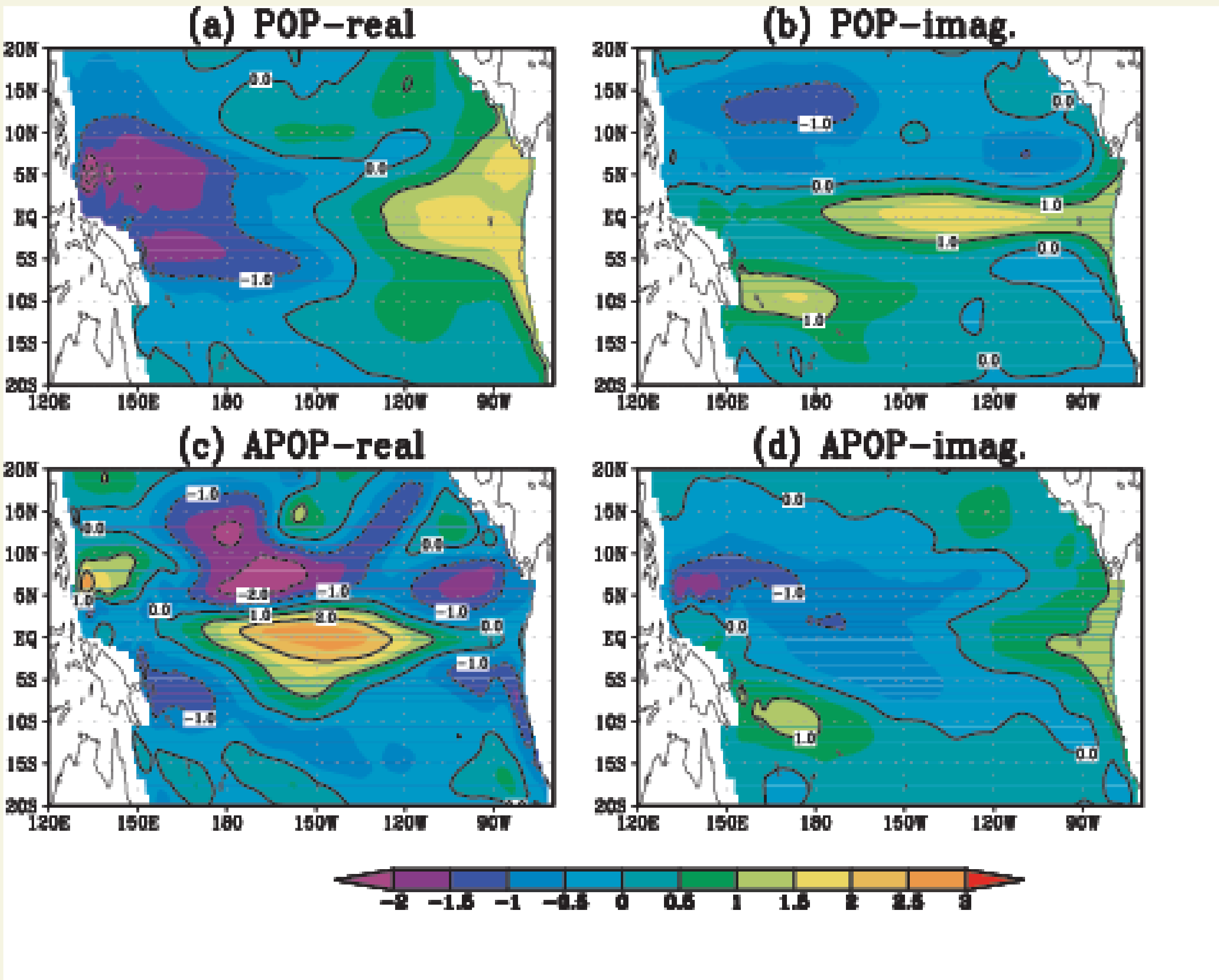
Signal versus Dispersion



POP amplitude in initial conditions

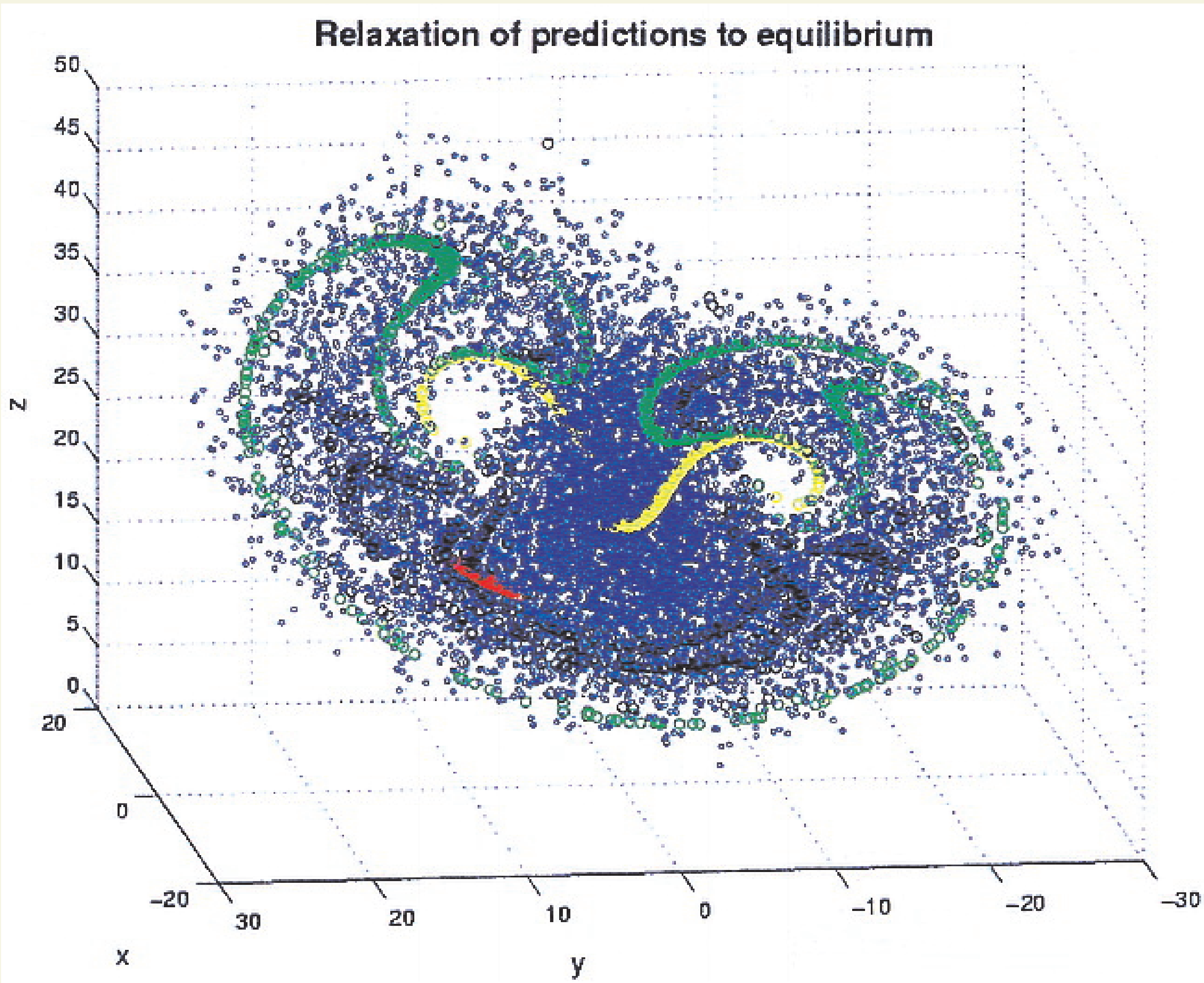


POP and APOP Heat Content Spatial Patterns (Complex and Real)



Classical Lorenz Three Component Model

$$\dot{x} = -\sigma(x - y) \qquad \dot{y} = \rho x - y - xz$$
$$\dot{z} = xy - \beta z.$$

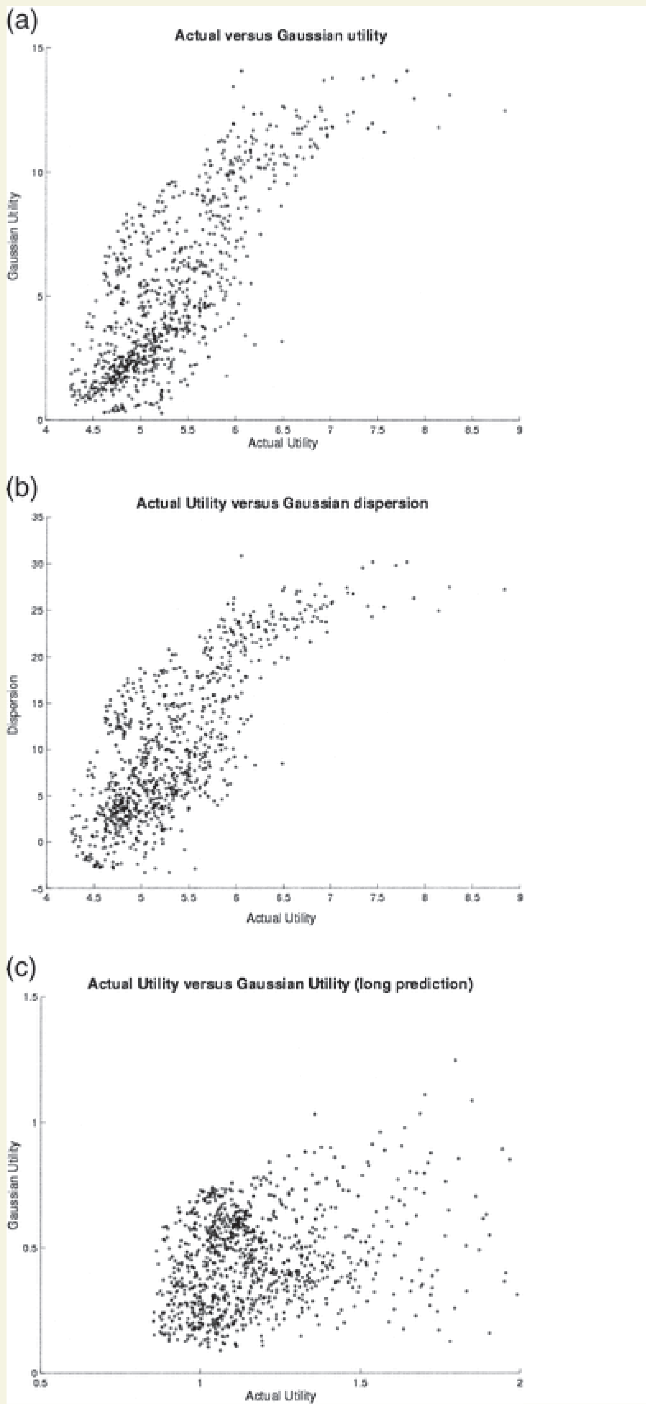


Note the Non-Gaussianicity for longer prediction ensembles

Methodology

Sample sizes of 10^5 were used and 1000 random initial conditions selected from the attractor. An initial condition Gaussian distribution was assumed with a standard deviation around two orders of magnitude less than the attractor size. Relative entropy was then calculated using a discrete formula. Monotonic decline in utility was always observed despite claims in the literature of a long lead return of skill in the Lorenz attractor. Such a result is likely a function of the skill measure deployed.

Relationship to the Gaussian Entropy



Quasigeostrophic Baroclinic Turbulence

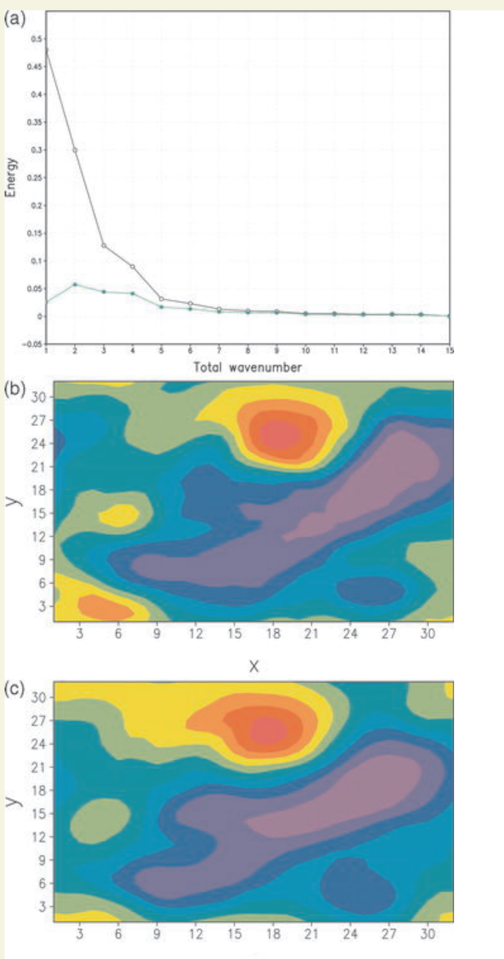
Quasigeostrophic two level model on a beta plane with a uniform mean shear to simulate the jetstream. Domain is doubly periodic in the zonal and meridional direction and the first 16 Fourier components are retained in both directions which is sufficient to resolve the Rossby radius for the parameter settings deployed. Model equations:

$$\frac{\partial q_i}{\partial t} + J[\psi_i + \overline{\psi}_i, q_i + \overline{q}_i + R(i)h_b] = -\kappa R(i)\nabla^2\psi_i + F_{\mathrm{hyp}},$$

$$q_1 + \overline{q}_1 = \nabla^2\psi_1 + \beta y + S[(\psi_2 - \psi_1) + \mathrm{US}y],$$
$$q_2 + \overline{q}_2 = \nabla^2\psi_2 + \beta y + S[(\psi_1 - \psi_2) + \mathrm{US}y],$$

$$\overline{\psi}_i = -U_i y,$$

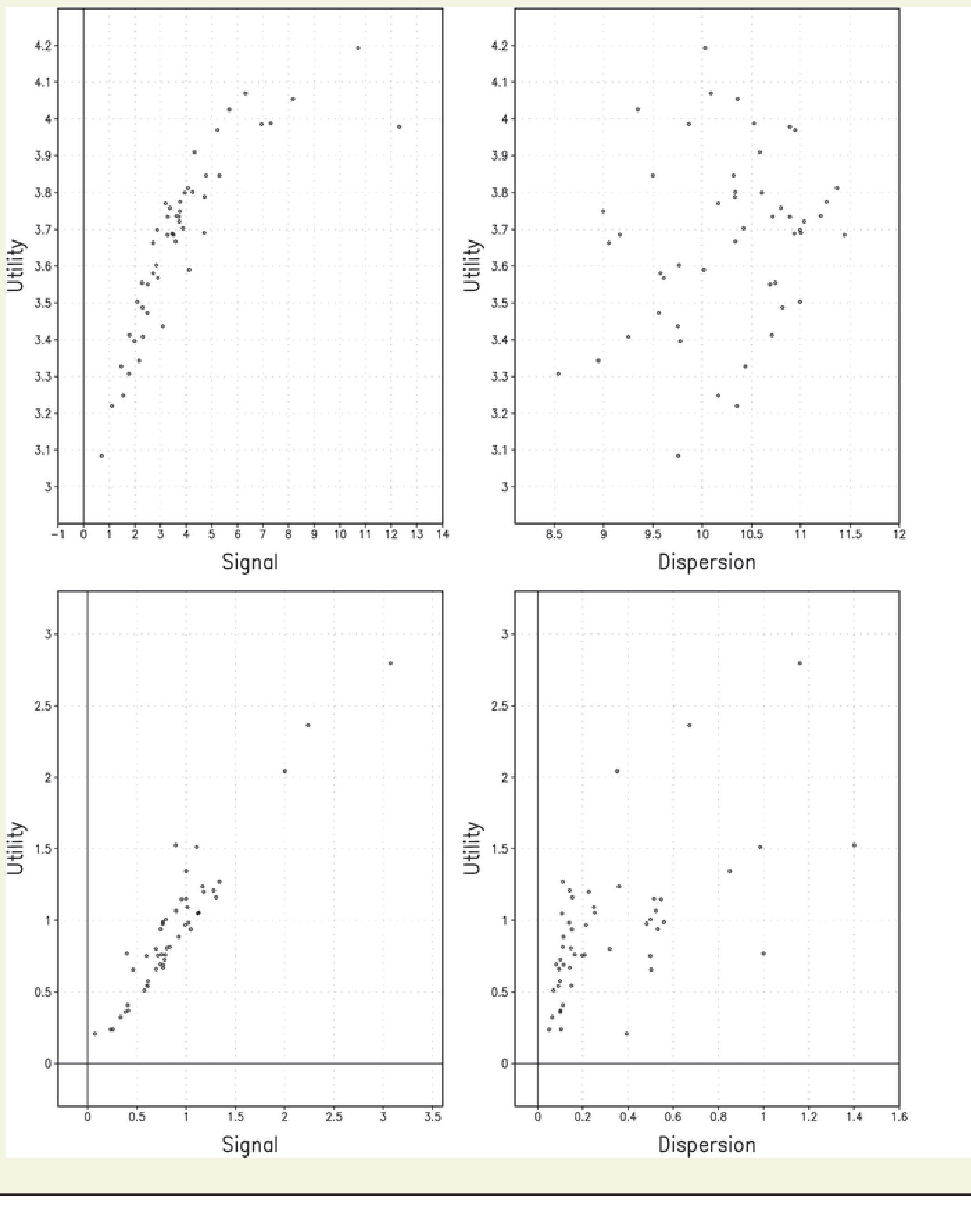
Equilibrium spectral energy (green baroclinic; black barotropic) plus a typical snapshot:



Entropy and Predictability Methodology

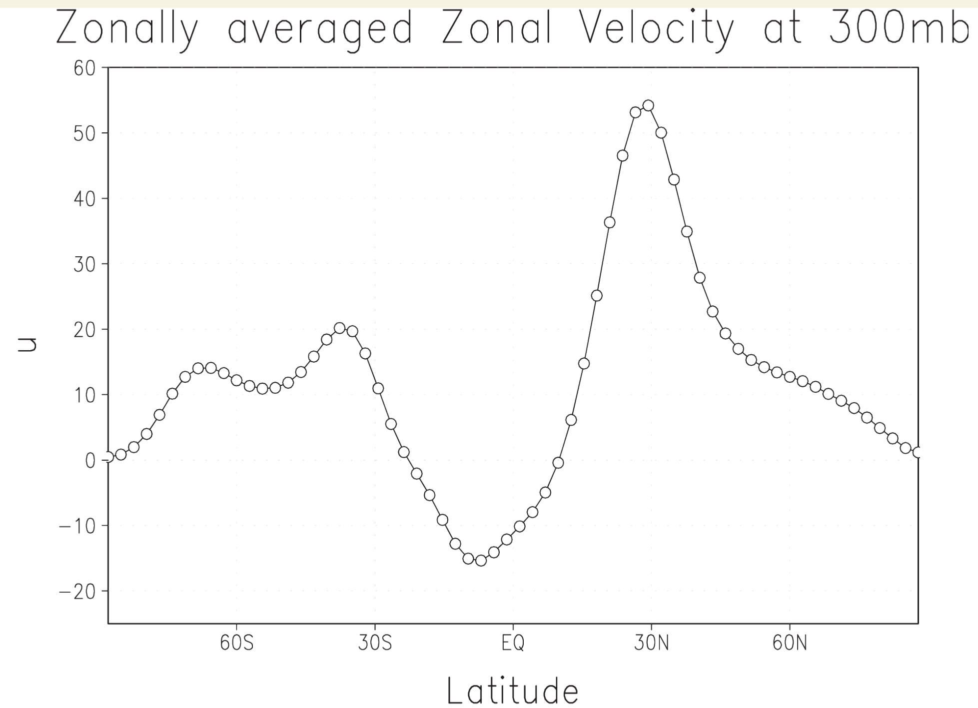
A reduced state was chosen consisting of the first two complex barotropic Fourier components. These modes explain around 60% of total variance in an equilibrium run. 50 random initial conditions from an equilibrium run were selected and Gaussian initial condition distributions with standard deviation two orders of magnitude smaller than climatology used. 1000 member prediction ensembles were used and each reduced state space dimension was partitioned into quartiles implying and average bin count of around 4. Ensembles appeared often quite close to Gaussian.

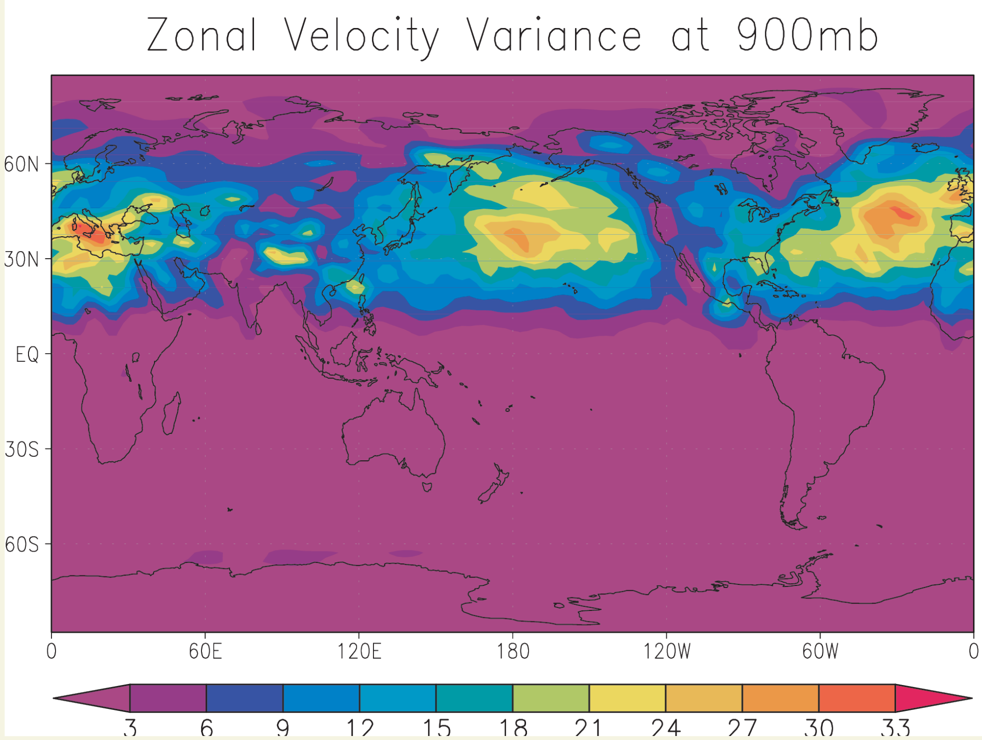
Relation of Entropy to Signal and Dispersion



Realistic Primitive Equation Atmospheric Model

T42 and 5 vertical levels. Realistic orography and Northern Winter thermal forcing. Simplified physics with Newtonian cooling to a zonally uniform "radiative/convective" profile replacing radiation and convection. Good simulation of the mid-latitude storm tracks regions in both hemispheres. Realistic jetstream behaviour.





Methodology

Reduced state space taken from the EOF spectrum of primitive equation prognostic variables (rescaled so that each has equal global mean variance). In general convergence of total explained variance is much slower than in the ENSO case and around 20 EOFs are required to explain approximately 50% of variance while 60 EOFs explain around 85% of variance. Results shown here are for a 40 dimensional reduced space.

High dimensional spaces imply that with practical ensembles only marginal distributions are able to be estimated. To deal with this we introduce a heirarchy of so-called m'th order marginal relative entropies which measure the mean information content of m dimensional subspaces:

$$D^m(p \parallel q) \equiv$$

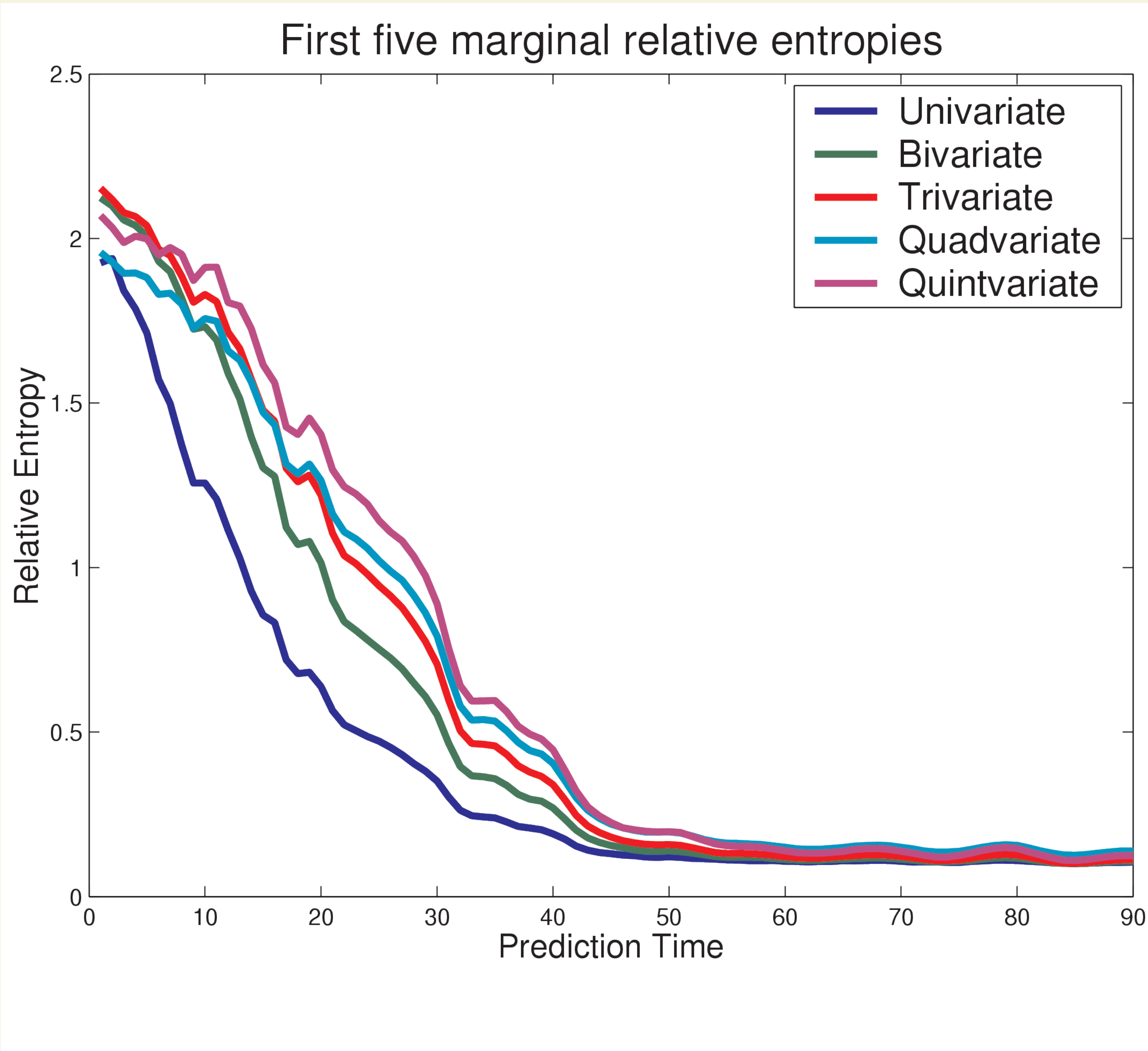
$$\frac{1}{C^n_m} \sum_{j_1,,j_2,\ldots,j_m} D\left(p(X_{j_1}, X_{j_2}, \ldots, X_{j_m}) \parallel q(X_{j_1}, X_{j_2}, \ldots, X_{j_m})\right)$$

$$D^1(p \parallel q) \leq D^2(p \parallel q) \leq \ldots \leq D^n(p \parallel q) = D(p \parallel q)$$

In general finer partitioning per dimension is possible for lower order marginal distribution estimation for a fixed ensemble size. We chose the largest number consistent with an adequate bin count and concentrated on the first five marginal entropies. Note the inequality heirarchy above only holds for constant partitioning across marginal entropies.

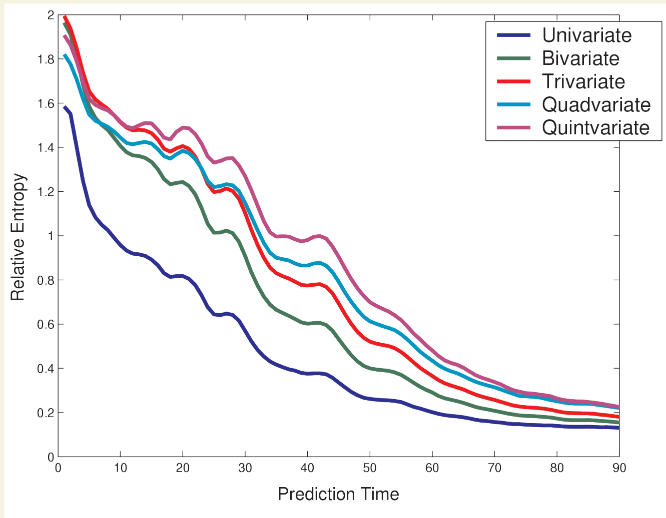
Generic behaviour

We chose one particular initial condition and derived the initial condition distribution as for the quasigeostrophic case. Results here are for a 9600 member ensemble.

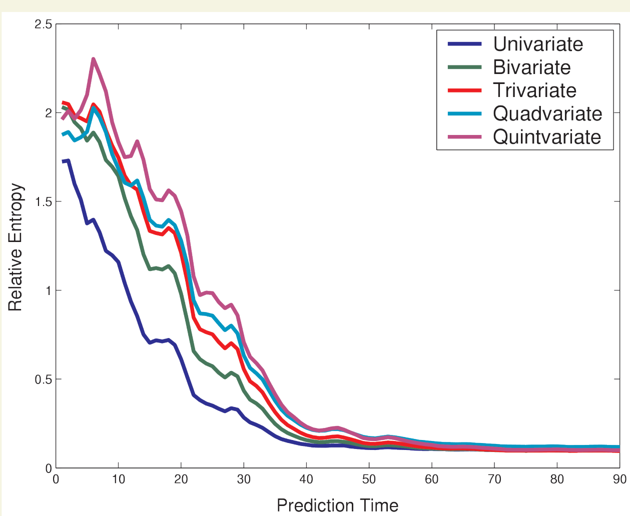


Regional Variation

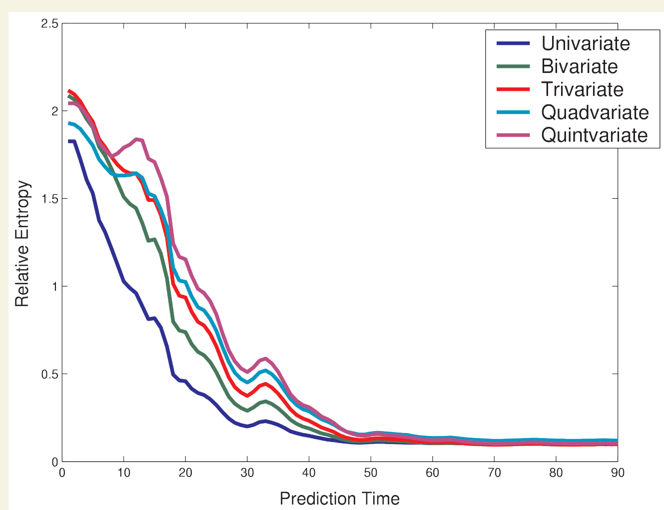
Southern Ocean



North Atlantic

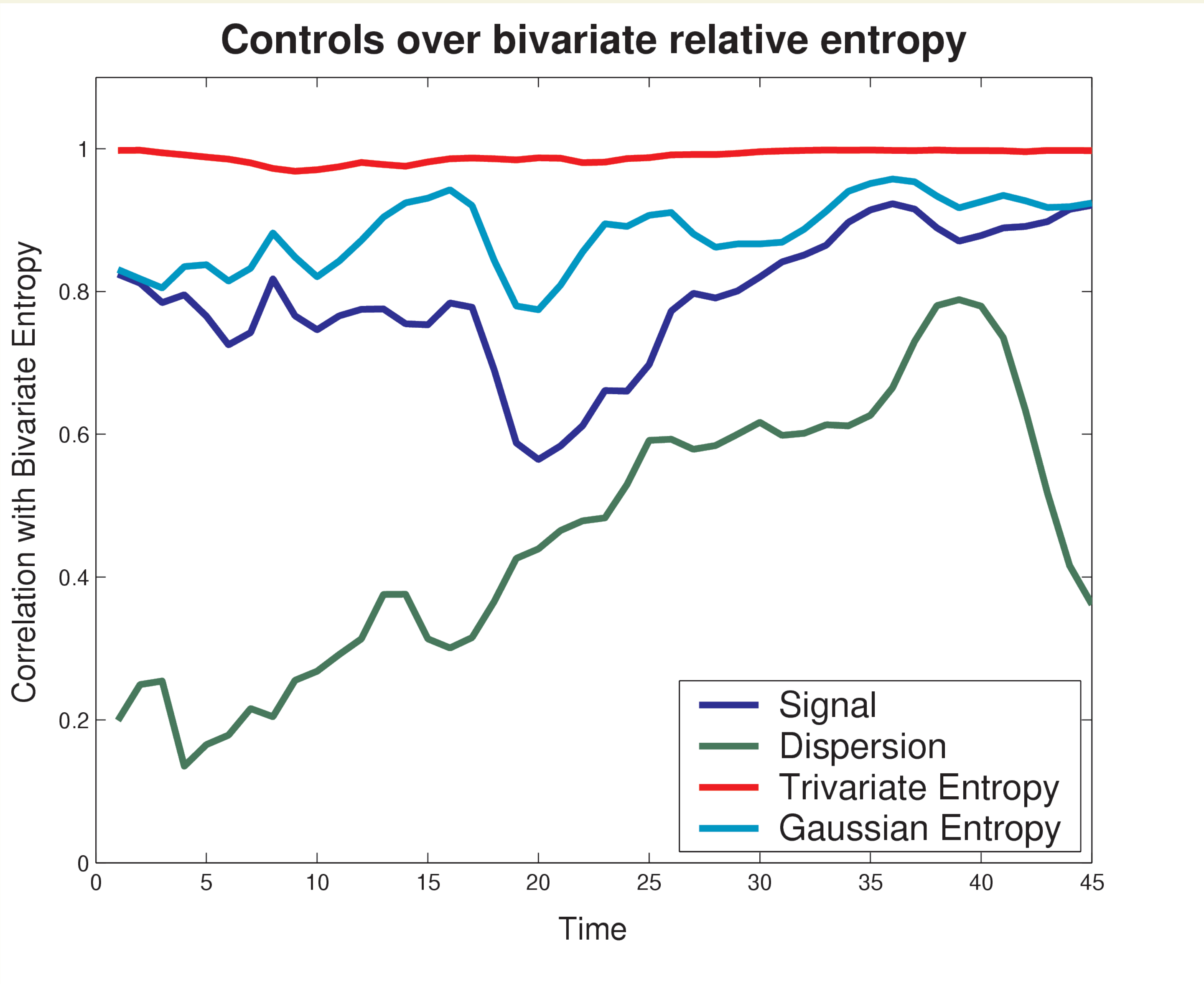


North Pacific



Variation of Utility with Initial Conditions. Importance of Signal and Dispersion

1000 member ensembles. Second and third marginal entropies and 50 initial conditions.



Conclusions

Prediction utility can be defined rigorously using the information theoretic functional of relative entropy

Two kinds of utility can be identified based on the near Gaussianicity of many prediction distributions. SIGNAL measures the shift in the mean of the prediction from the climatological distributions. DISPERSION measures the reduction in uncertainty of the prediction versus climatological distributions.

Variations in utility from one set of initial conditions to another are almost always strongly related to variations of Signal rather than to Dispersion. This conclusion seems to hold for realistic climate and weather prediction models. It does not appear to hold for the exceptional case of strongly chaotic systems such as the Lorenz 3 component model.

Preliminary evidence from a simplified weather model suggests that there is an approximate finite time cutoff for predictability of between 1 and 2 months. Beyond this cutoff initial condition information is not important to statistical prediction. Of course boundary condition information still can be.