# Mid-latitude statistical predictability of the atmosphere. Variability and the potential impact of tropical convection

Richard Kleeman and Minora Kadota
Center of Atmosphere Ocean Science
Courant Institute of Mathematical Sciences

Relevant References

R. Kleeman. Statistical predictability in the atmosphere and other dynamical systems. Physica D, 2006. Special issue on data assimilation. In press.

R. Kleeman. Limits, variability and general behaviour of statistical predictability of the mid-latitude atmosphere. J. Atmos Sci, 2006. Submitted and revised.

A relaxation perspective on statistical predictability
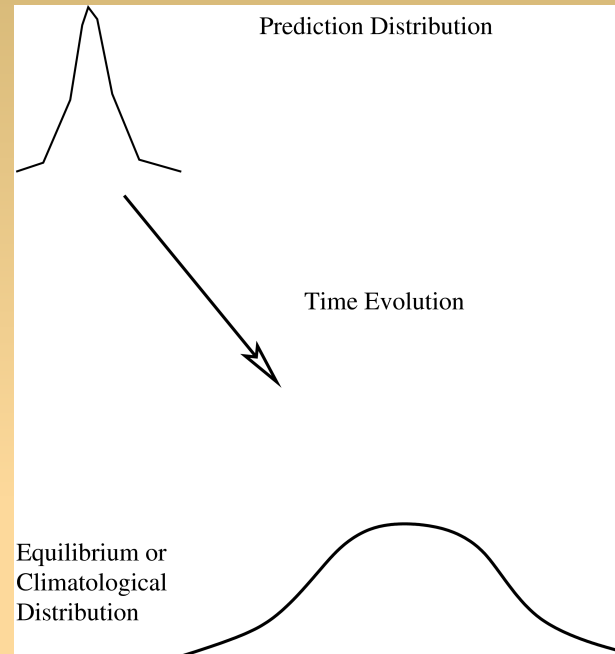
The relaxation of a prediction ensemble back toward the climatological ensemble is the fundamental characteristic of statistical prediction from uncertain initial conditions. Once convergence has occurred then initial condition information is useless to the prediction. This relaxation can be effectively measured using the relative entropy of the two ensembles as indeed is common for general classes of stochastic equations such as the Chapman-Kolmogorov or Fokker-Planck equations. This measure also has the practical interpretation of the potential utility of the statistical prediction.

Gaussian Ensembles and Common Skill Measures

$$RE = Dispersion + Signal$$

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

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

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

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

Many ensembles of practical interest are close to Gaussian perhaps because they are from dynamical systems with many degrees of freedom and the central limit theorem of statistics is possibly appropriate. It is therefore a worthwhile analysis tool to consider the Gaussian relative entropy. It splits into two pieces with intuitively clear interpretations. There is a also a relation to commonly used perfect skill measures. Note that such measures are ensemble means (the overbar) while entropy is defined on one particular ensemble.

Simplified Primitive Equation Atmospheric Model

Zonally averaged Zonal Velocity at 300mb
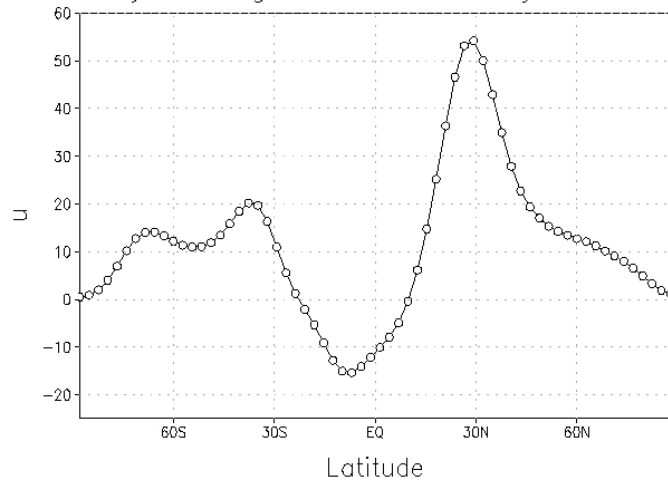
Zonal Velocity Variance at 900mb
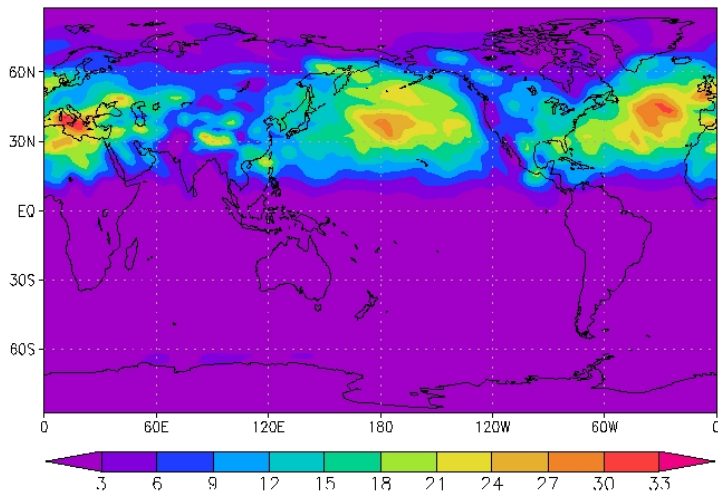
T42 and 5 vertical levels. Realistic orography and northern winter thermal forcing. Simplified physics with Newtonian cooling to a zonally uniform "radiative/convective" profile replacing radiation and convection. The vertical relaxation profile varies meridionally from close to a moist adiabat in the deep tropics to a more unstable profile at higher latitudes. This reflects the approximately the observed zonally averaged temperature structure. The model has a good simulation of the mid-latitude storm tracks regions in both hemispheres and also has fairly realistic jet stream behaviour. When forced diabatically from the West Pacific it produces a realistic PNA wave train over the Pacific and upper North America.

Ensembles are generated by sampling from an isotropic Gaussian distribution with a 1000km horizontal decorrelation and no correlation vertically or between prognostic variables. Variances are chosen to be two orders of magnitude smaller than the global average climatological/equilibrium values for each prognostic variable. This is obviously a highly simplified representation of model initialization.

Ensemble Methodology

A reduced state approach is adopted using rescaled prognostic variable EOFs. The four prognostic variables (at all vertical levels) are rescaled to give equal mean global variances before EOFs are calculated. Globally the EOF spectrum at T42 converges relatively slowly and approximately 60 EOFs are required to capture 90% of total variance. For the regional case fewer are required. We selected the first 60 EOFs as our reduced space. This EOF space was different for each regional calculation shown below.

In such a high dimensional space practical ensembles effectively sample only low dimensional subspaces of the full multivariate pdf due to the so-called curse of dimensionality. We addressed this issue in two ways. Firstly we used marginal relative entropies which are the relative entropies averaged over all possible subspaces of given dimensionality. Secondly since distributions were usually close to Gaussian we calculated the relative entropy under the assumption that "converged" pdfs were indeed Gaussian. More technical details can be found in the papers cited earlier.

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

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

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

The marginal relative entropies, when defined with respect to the same partitioning of state space, satisfy a natural hierarchy because they represent the information content of covariations of increasing degree. When the dimensionality of the marginal subspace grows to the dimensionality of the full space then the marginal entropy becomes the full entropy.

The converged Gaussian relative entropy is generally much larger than the marginal entropies for practical partitioning of state space because it is assumes considerable multivariate structure not adequately sampled by the marginal entropies. Nevertheless after rescaling it behaves in practical situations very similarly to marginal entropies.
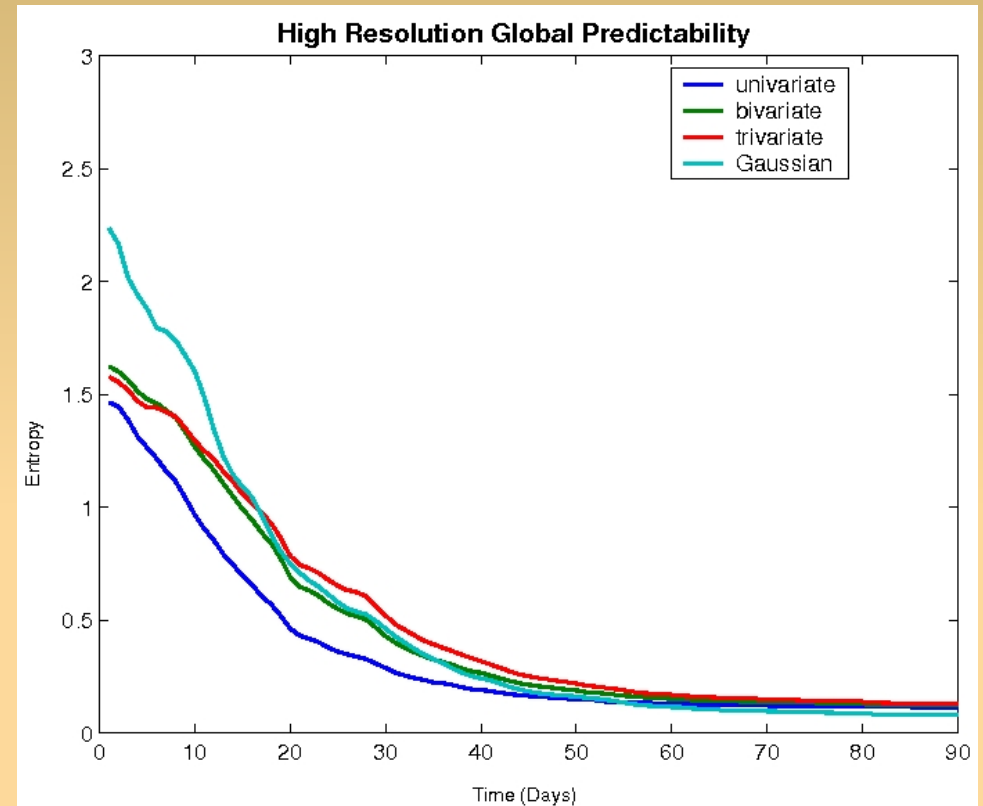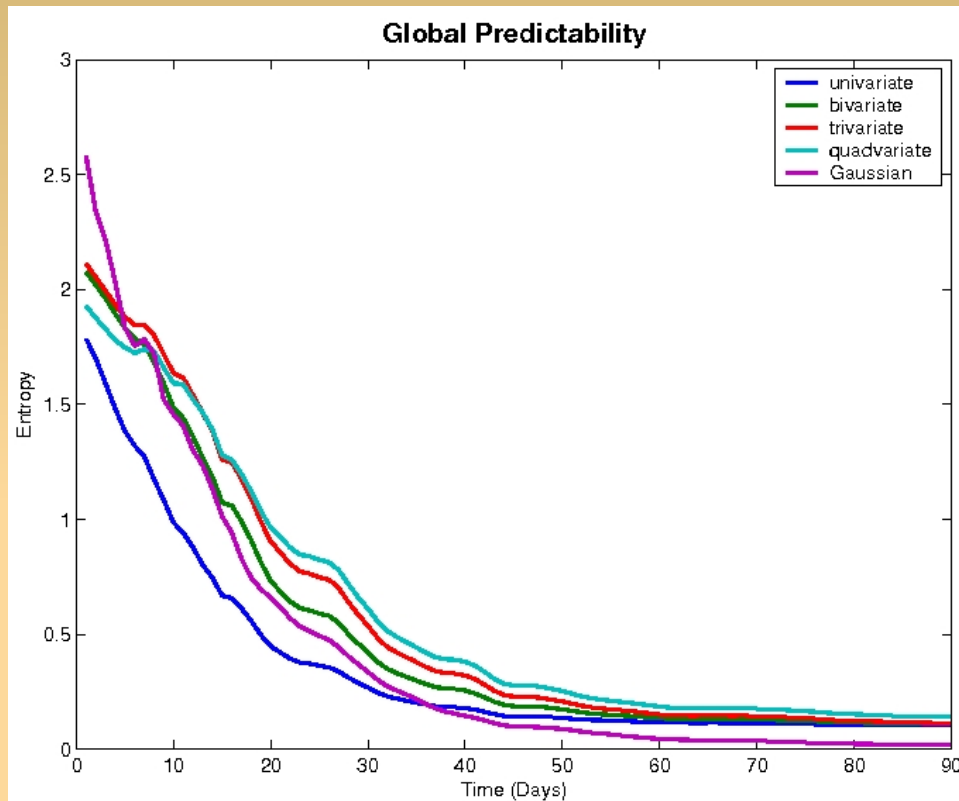
Qualitative (but not quantitative) behaviour does not vary much from initial condition to initial condition; with increased model horizontal resolution (T85) or with reduced state-spaces of increased dimension. Initial condition means of prognostic variables here are drawn at random from a climatological run of the model.

Regional Variation in Predictability
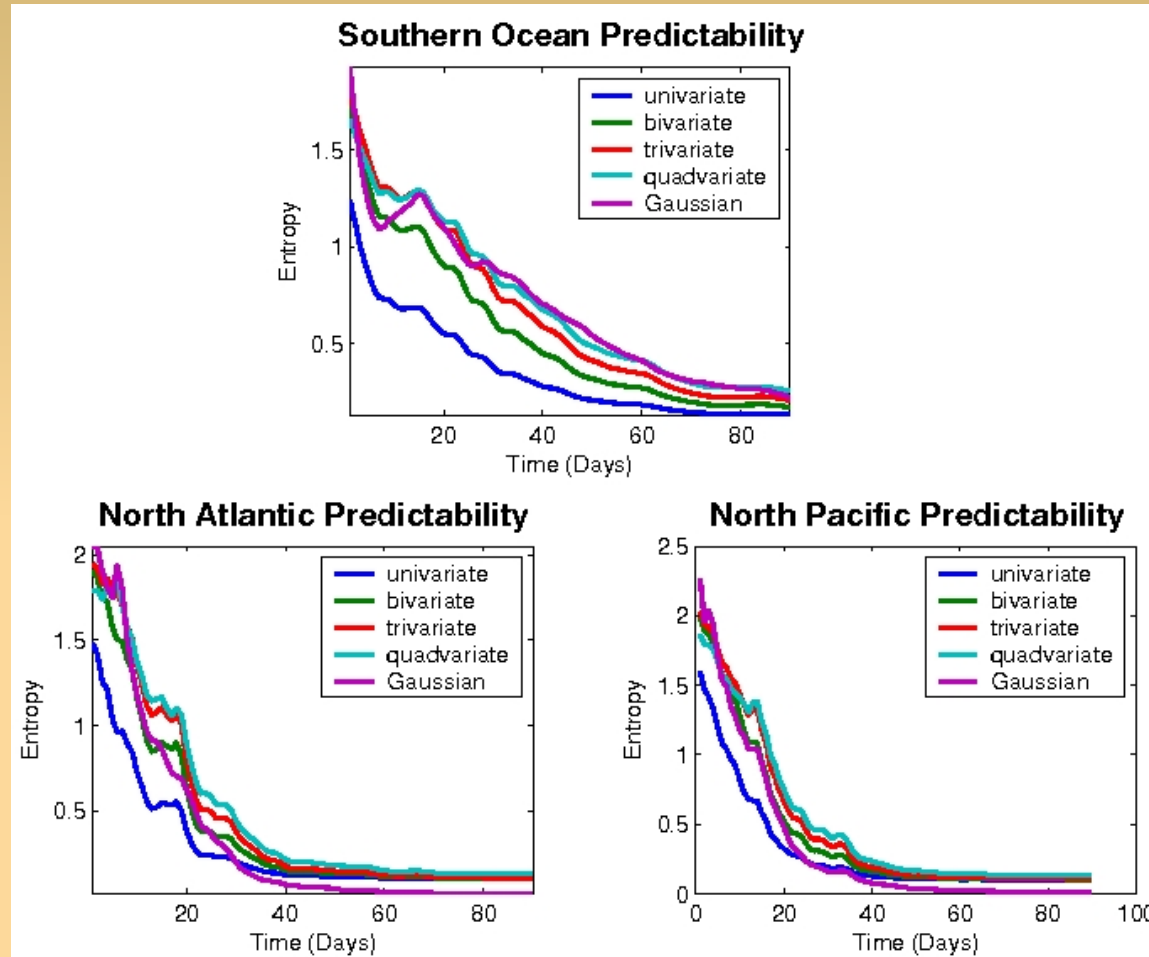
Southern Ocean in the summer hemisphere shows much enhanced predictability over the northern winter storm track regions. Note that reduced EOF spaces here are different for each case.
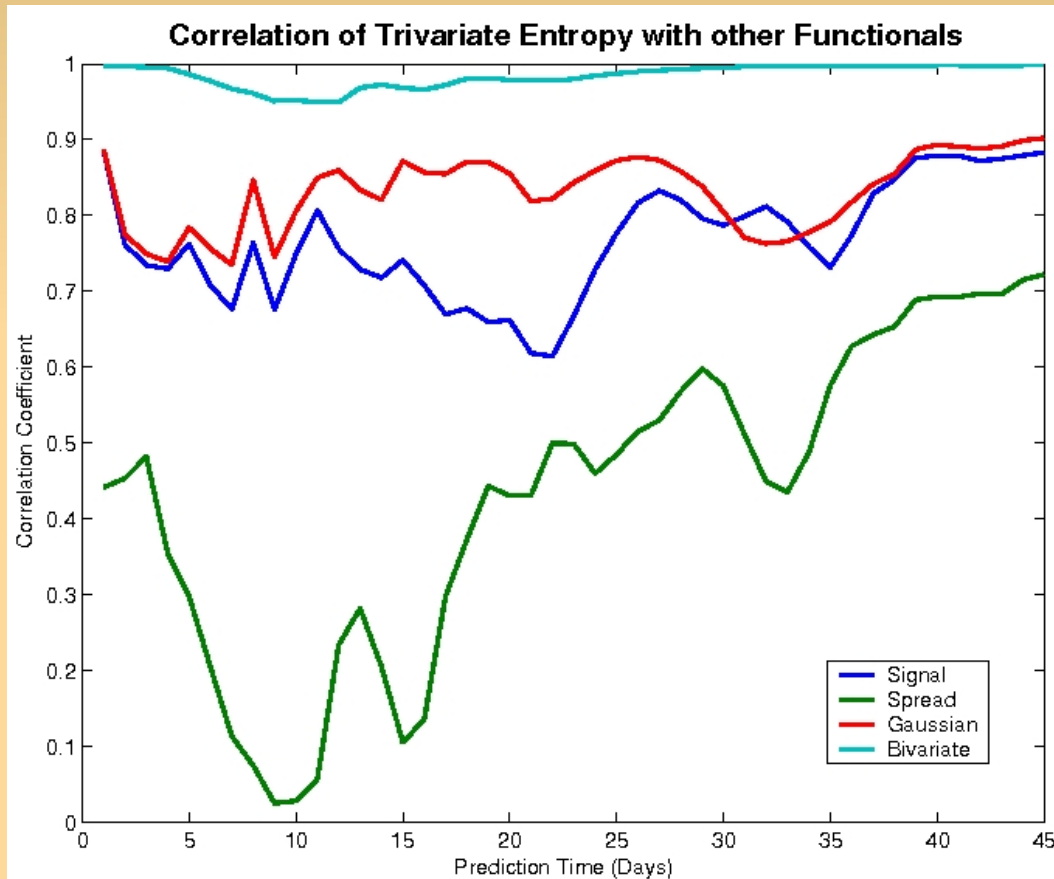
Gaussian and marginal entropies are very highly correlated because of the approximately Gaussian behaviour of most ensembles. Signal dominates over dispersion particularly at "short" predictions leads. Dispersion is probably underestimated for very short predictions because our initial condition distribution is not model dependent unlike practical NWP where the model is used during data assimilation as a dynamical interpolator.

Lagged Correlation of Signal

Lagged Correlation of Dispersion
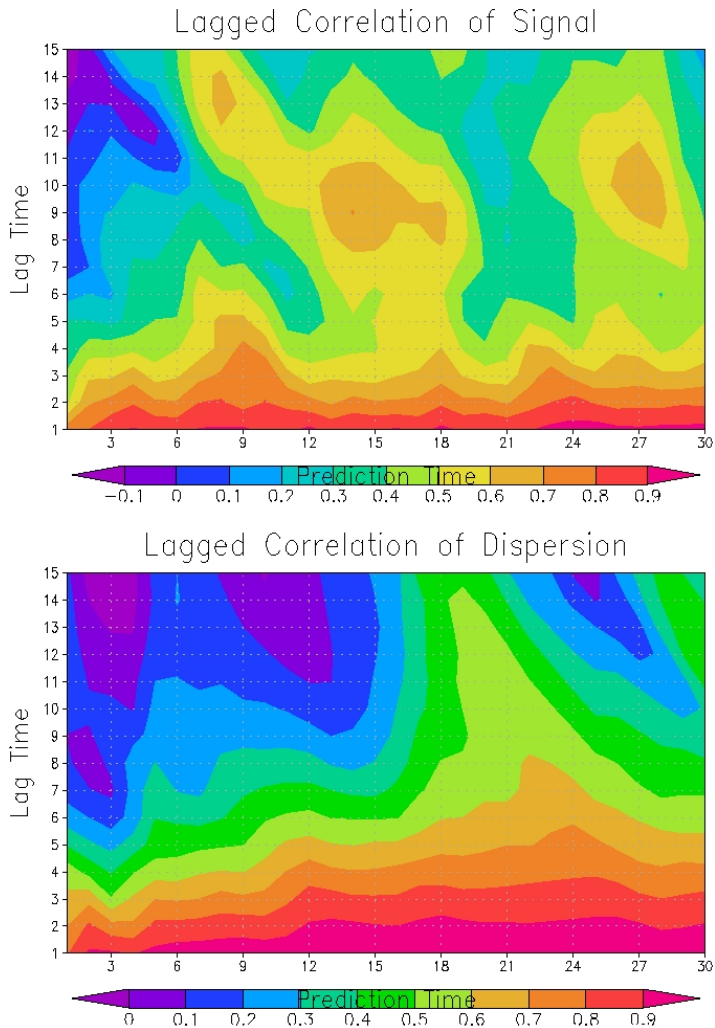
Signal shows little preference for any particular EOF and seems to be simply a generic anomalous pattern present in the prediction ensemble. Dispersion shows more such preference but this has not been thoroughly analysed yet. Both dispersion and signal show strong temporal decorrelation over 2-5 days suggesting that the patterns responsible for high or low predictability "dissipate" rather rapidly. Note that longer lag correlations appear stronger for long range predictions. The general kind of behaviour reported here seems different to climate prediction where two spatial modes appear responsible for signal (and predictability).

Madden Julian Oscillation in Models
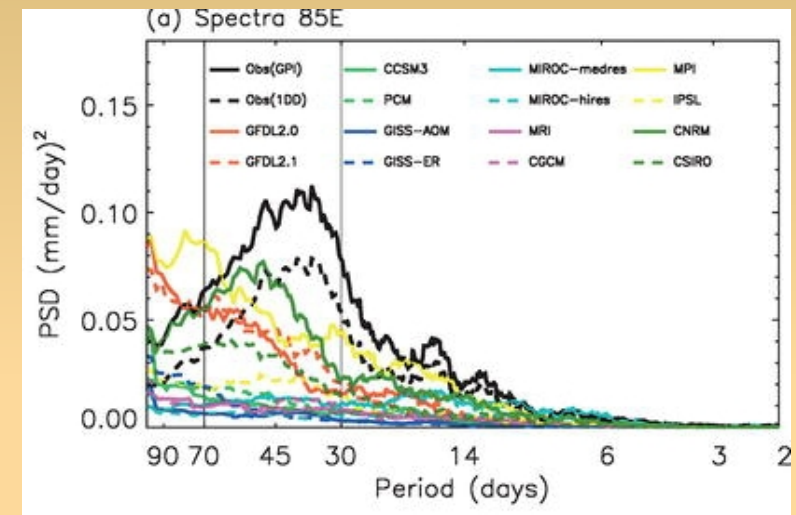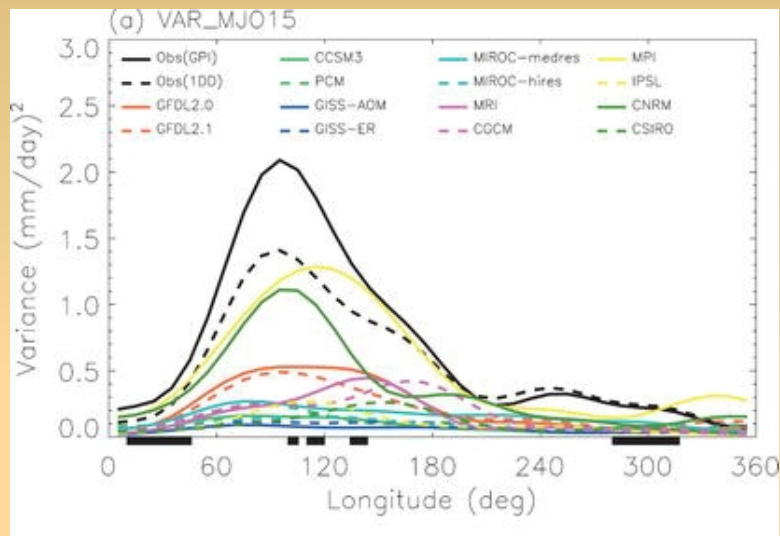
J. Lin, G. N. Kiladis, B. E. Mapes, K. M. Weickmann, K. R. Sperber, W. Lin, M. C. Wheeler, S. D. Schubert, A. Del Genio, L. J. Donner, S. Emori, J-F Gueremy, F. Hourdin, P. J. Rasch, E. Roeckner, and J. F. Scinocca
Tropical Intraseasonal Variability in 14 IPCC AR4 Climate Models. Part I: Convective Signals *J. Clim.* **19**, (June 2006) pp. 2665–2690

There are clearly major problems still with simulated variance and spectrum. Similar problems are evident in NWP models. Might this have an effect on mid-latitude weather prediction?

Background and Motivation

- The MJO is a quasi-regular large scale convective disturbance which has a less than perfect representation in present generation general circulation models.

- Better simulation of the MJO may allow improvements in medium to long range weather predictions since it is of significant amplitude; has some degree of regularity and is sufficiently large scale to induce global teleconnections

- Study of the potential effect on predictability requires an adequate global atmospheric model; a reasonably realistic representation of the influence of the MJO on the atmosphere and a careful analysis of its effect on atmospheric ensemble predictions.

Models

- Statistical model of the MJO diabatic heating. Consists of a stochastically forced damped oscillator for the amplitudes of the first two observed band-passed OLR EOFs. Many previous observational studies have shown these two EOFs to adequately represent the dominant component of the MJO. Realistic time-lag relations and quasi-regularity of the MJO result from suitable parameter choices for the stochastic damped oscillator. Model diabatic heating variance was adjusted to be close to local warm pool observations (Arkin's formula was used on observational OLR data). A sinusoidal tropospheric vertical heating profile was assumed.

- Diabatic heating from statistical model above inserted into the atmospheric models described previously.

Prediction Experiments

- 1000 member atmospheric atmospheric ensembles for the initial conditions were constructed as described earlier.

- Predictability was assessed by measuring as before by looking at the convergence of the prediction ensemble to the climatological ensemble. The Gaussian relative entropy of the two ensembles was chosen to measure this convergence but results with marginal entropy were qualitatively similar.

- Two experiments were run. One in which it was assumed the MJO amplitudes were known perfectly and hence did not change among ensemble members. The second assumed ignorance of the MJO amplitudes which were therefore chosen randomly according to climatology.   The difference was taken as a measure of potential predictability. This is a crude test of the influence of a very poor model on predictability.
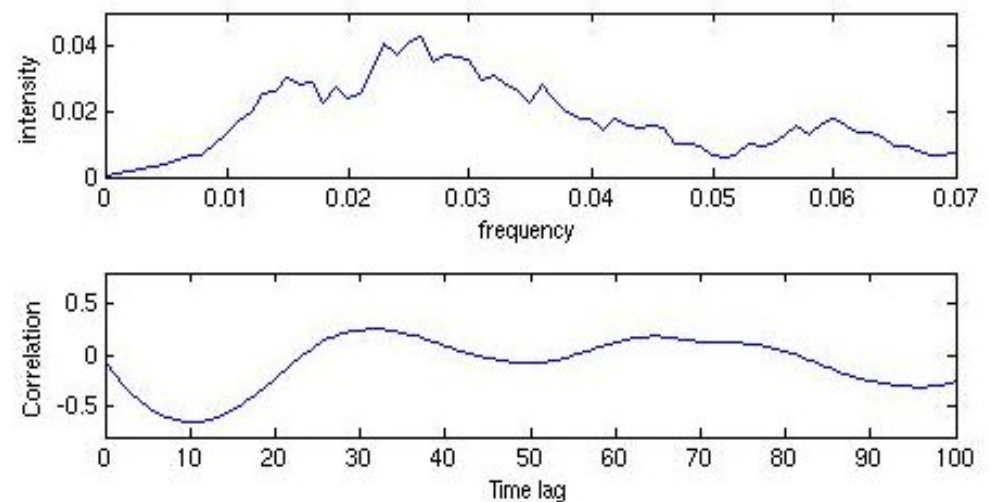
MJO Model Performance

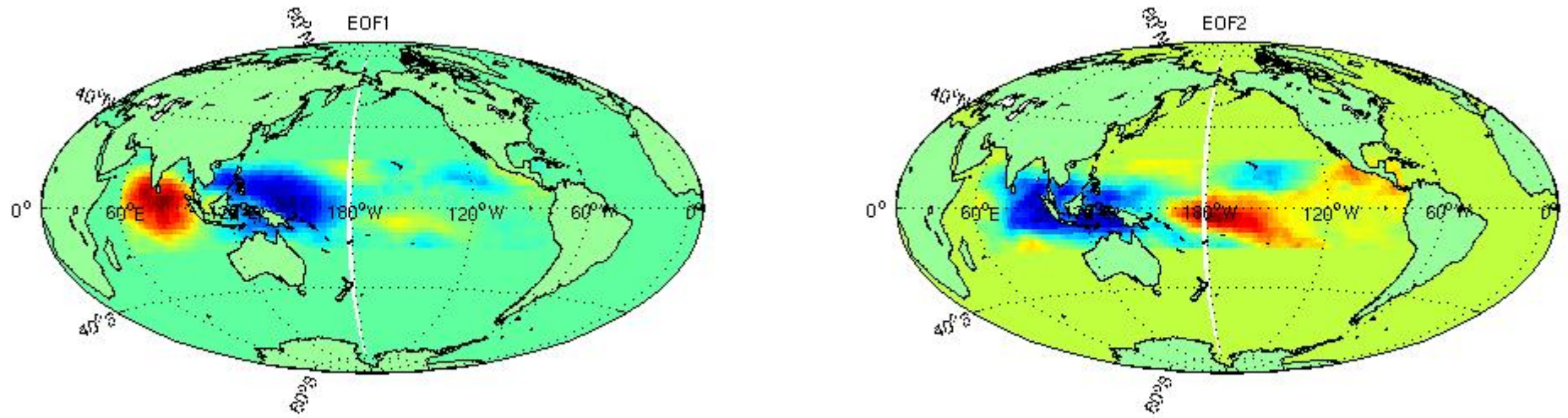**Spectral Intensity and Time Lagged Correlation between EOF amplitudes**

Band Passed Observations
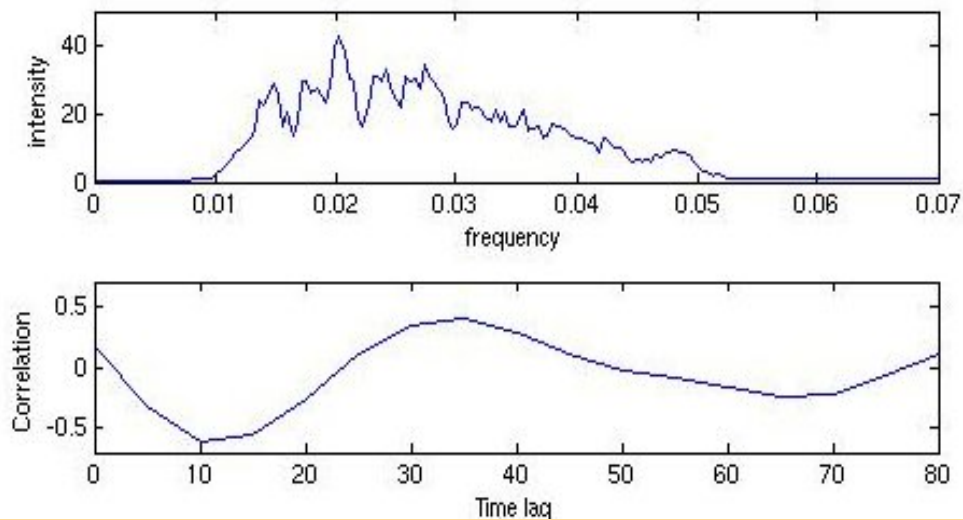
Stochastic Model

Dynamical Evolution
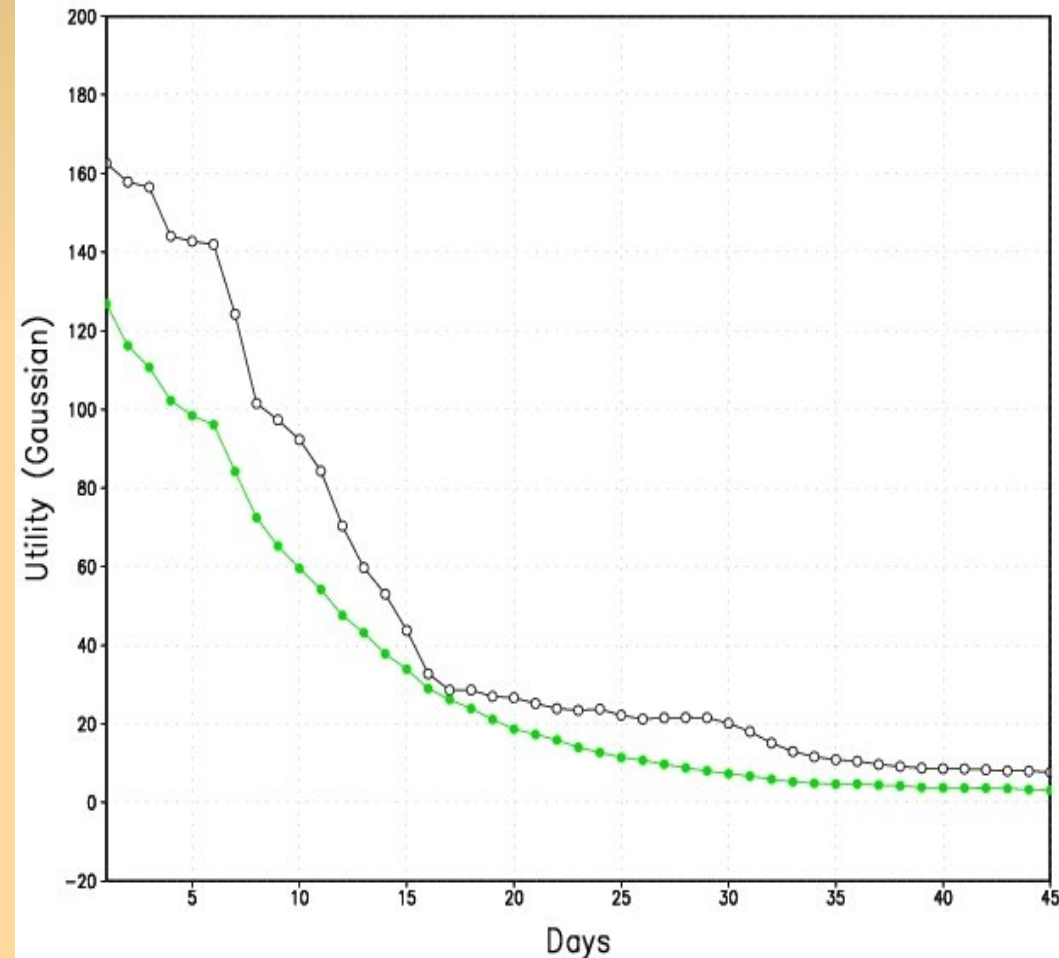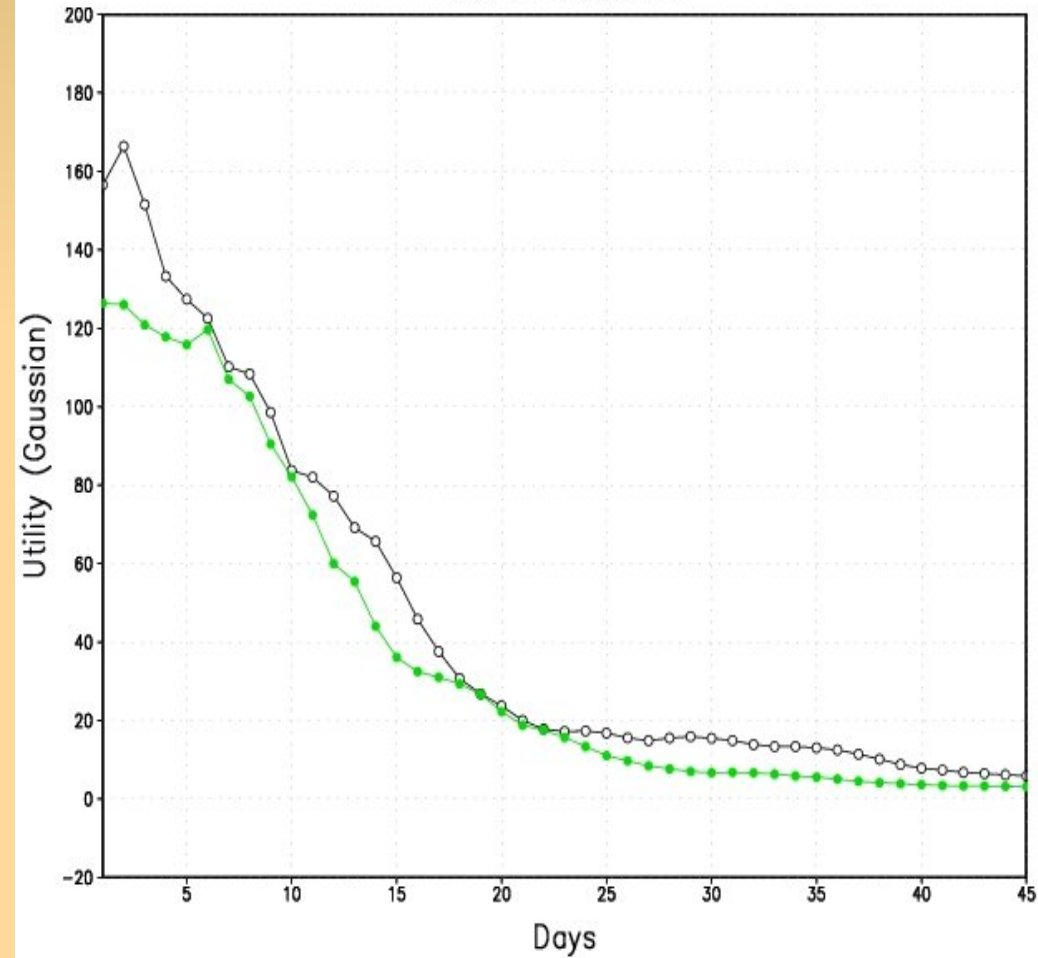EOF1 ⟶ -EOF2

MJO Potential Predictability Effect

Northern Winter Storm Track Regions

Black: Perfect MJO ics.   Green: Random climatological MJO ics.

Conclusions

- Mid-latitude statistical potential predictability from initial conditions appears to be around 40 days in the winter hemisphere storm tracks and twice as long in the summer hemisphere. Beyond these times climatological and prediction ensembles are statistically identical.

- Variations in predictability with initial conditions seem controlled primarily by variations in the amplitude of anomalous fields. This "signal" shows a very short temporal decorrelation scale.

- The MJO appears to be potentially able to increase mid-latitude predictability by several days. The effect is more pronounced for shorter predictions and is confined almost entirely to the first two weeks. Regions closer geographically to the MJO centre of action around the maritime continent show a bigger effect.

- These results support the case for improving the representation and increasing the understanding of MJO convection in NWP models.