Measuring the potential utility of seasonal climate predictions

Michael K. Tippett
International Research Institute for Climate Prediction, Columbia University, Palisades, New York, USA

Richard Kleeman and Youmin Tang
Courant Institute of Mathematical Sciences, New York University, New York, New York, USA

Received 21 September 2004; accepted 19 October 2004; published 16 November 2004.

[1] Variation of sea surface temperature (SST) on seasonal-to-interannual time-scales leads to changes in seasonal weather statistics and seasonal climate anomalies. Relative entropy, an information theory measure of utility, is used to quantify the impact of SST variations on seasonal precipitation compared to natural variability. An ensemble of general circulation model (GCM) simulations is used to estimate this quantity in three regions where tropical SST has a large impact on precipitation: South Florida, the Nordeste of Brazil and Kenya. We find the yearly variation of relative entropy is strongly correlated with shifts in ensemble mean precipitation and weakly correlated with ensemble variance. Relative entropy is also found to be related to measures of the ability of the GCM to reproduce observations. *INDEX TERMS:* 1620 Global Change: Climate dynamics (3309); 3354 Meteorology and Atmospheric Dynamics: Precipitation (1854); 3339 Meteorology and Atmospheric Dynamics: Ocean/atmosphere interactions (0312, 4504); 4522 Oceanography: Physical: El Nino; 1869 Hydrology: Stochastic processes. **Citation:** Tippett, M. K., R. Kleeman, and Y. Tang (2004), Measuring the potential utility of seasonal climate predictions, *Geophys. Res. Lett.*, *31*, L22201, doi:10.1029/2004GL021575.

1. Introduction

[2] Seasonal variability of precipitation and associated extremes such as drought or flooding are of particular interest to society. Some seasonal climate anomalies are associated with variation of tropical sea surface temperature (SST) on seasonal-to-interannual time-scales. A notable example of such a connection between seasonal precipitation and SST are precipitation anomalies associated with ENSO [*Ropelewski and Halpert*, 1987; *Mason and Goddard*, 2001]. Information theory provides a useful framework for measuring the impact of SST forcing on climate variability [*Schneider and Griffies*, 1999; *Kleeman*, 2002; *DelSole*, 2004]. In this setting, the seasonal precipitation amount x is viewed as a random variable with climatological distribution q. This climatological distribution is then compared with the distribution p of precipitation amounts given a particular SST. The impact of SST on seasonal precipitation is measured by the extent to which the two distributions differ. If SST has no impact on precipitation, the two distributions will be identical. On the other hand, if SST has an impact on precipitation amounts, the two distributions will differ significantly.

0094-8276/04/2004GL021575$05.00

[3] There are various measures to quantify the difference between two distributions including the t and F tests for Gaussian distributions and the Kolmogorov-Smirnov distance for general distributions [*Sardeshmukh et al.*, 2000]. Relative entropy is sensitive to changes in mean, variance and higher order moments, and measures the informational inefficiency of using the climatological distribution instead of the SST-forced distribution [*Kleeman*, 2002]. Relative entropy can be used to detect when distributions are different as well as to measure the difficulty of detection. Relative entropy is invariant with respect to invertible transformations, meaning that it is unchanged when units are changed or when the quantity of interest is a nonlinear function of precipitation, for instance, in applications that are sensitive to extreme values [*Kleeman*, 2002; *Majda et al.*, 2002]. Other interpretations of this quantity include determining the financial advantage of a gambler knowing the SST-forced distribution when "fair-odds" come from climatology [*DelSole*, 2004].

[4] The relative entropy R is defined mathematically by

$$R = \int p \ln \frac{p}{q} \, dx \,. \tag{1}$$

When the distributions are Gaussian, (1) has the simple form [*Kleeman*, 2002]

$$R = \frac{1}{2}\left[\ln\left(\frac{\sigma_q^2}{\sigma_p^2}\right) + \frac{\sigma_p^2}{\sigma_q^2} + \frac{\mu_p^2}{\sigma_q^2} - 1\right], \tag{2}$$

where μ_p and σ_p^2 are the mean and variance of p, and σ_q^2 is the climatological variance; the climatological mean is assumed without loss of generality to be zero. The relative importance of the contributions to R from changes in mean and variance depends on dynamical properties of the system [*Kleeman*, 2002].

[5] Calculating relative entropy requires specifying the SST-forced precipitation distribution p given a particular SST. Since nature only provides a single precipitation realization for a given SST, the SST-forced precipitation distribution is estimated from an ensemble of general circulation model (GCM) simulations forced with observed SST conditions [*Kumar and Hoerling*, 1995; *Rowell*, 1998; *Sardeshmukh et al.*, 2000]. Relative entropy, like signal-to-noise, is a perfect model measure of utility, and model deficiencies can limit its usefulness. However, one may expect that for good models its variations may be an indication of real variations in prediction utility.

[6] We compute the relative entropy for three regions where SST has a large impact on precipitation: South

Table 1. Domains and Seasons

Region	Domain	Season	r_{perfect}	r_{obs}
Florida	85W–75W, 22N–28N	Dec–Feb	0.78	0.75
Nordeste	45W–35W, 10S–EQ	Mar–May	0.77	0.63
Kenya	33E–43E, 5S–5N	Oct–Dec	0.67	0.84

Florida (including Cuba), the Nordeste of Brazil and Kenya [*Ropelewski and Halpert*, 1987]. Our goals are to quantify the yearly variation of potential utility as measured by relative entropy, to characterize the relative importance of changes in mean and higher order moments and to relate relative entropy with skill in reproducing observations.

2. Data and Methods

[7] Model data come from a 24 member ensemble of T42 ECHAM 4.5 GCM simulations forced with observed SST for the period January 1950 to March 2004 [*Roeckner et al.*, 1996]. Precipitation observations come from the extended *New et al.* [2000] gridded data set of monthly precipitation for the period 1950 to 1998. Model and observation data are averaged over the spatial domains and seasons indicated in Table 1.

[8] The sensitivity of precipitation in these three regions to SST anomalies is apparent either when the ensemble mean is compared to individual ensemble members or to observations. The size of the SST-forced response relative to the model's internal variability determines the "perfect model" correlation r_{perfect} of the ensemble mean with any ensemble member [*Kleeman and Moore*, 1999; *Sardeshmukh et al.*, 2000]. Both the perfect model correlation r_{perfect} and the observed correlation r_{obs} are high (Table 1) for these regions.

[9] The climatological and SST-forced distributions are approximated with a kernel density estimate using a normal kernel function [*Bowman and Azzalini*, 1997]. The climatological distribution q is estimated from all ensemble members and years (sample size is 1296); alternatively q could be estimated from a simulation forced by climatological SST. The SST-forced distribution p is estimated each year from the 24 member ensemble. The integral definition of relative entropy in (1) is computed using the estimated distributions evaluated at 120 equally spaced abscissa points whose range exceeds that of the model climatology distribution by 10% on either side. This kernel density estimate would likely be inappropriate for a quantity like daily rainfall whose distribution is far from Gaussian. However, the seasonal total distributions here are closer to being Gaussian than are those of daily values (Figure 1). The Gaussian approximation in (2) is reasonably accurate though it gives generally larger values, particularly when R itself is large (Figure 1).

[10] Although the relative entropy is zero when the simulation and climatology distributions are identical, finite ensemble size introduces sampling error. R. Kleeman and A. J. Majda (Predictability in a model of geostrophic turbulence, submitted to *Journal of Atmospheric Sciences*, 2004) discuss this issue in detail. In particular, a 24 member ensemble drawn from the model climatological distribution will generally not have zero relative entropy. We quantify the effect of sampling with a Monte Carlo method. 24 samples are drawn from the entire model climatology and their relative entropy is computed with respect to the climatological distribution. This process is repeated 100,000 times, and the sorted results indicate the likelihood that relative entropy exceeds a given value by chance. Values above the 95th percentile are considered significant.

3. Results

[11] The relative entropy of the SST-forced simulation with respect to climatology has mostly modest values; for Gaussian distributions a shift of one standard deviation without a change in variance corresponds to a relative entropy value of 0.5. Relative entropy values are statistically significant in 59% (32/54) of the years for Florida, 70% (38/54) of the years for the Nordeste and 50% (27/54) of the years for Kenya. The time-series in Figure 1 shows that relative entropy is very large for Florida and Kenya in only a handful of years. In the case of Florida, the three years with highest relative entropy, 1983, 1998 and 1973 are all warm ENSO events. In the case of Kenya, the three years with highest relative entropy, 1997, 1996, 1961 are warm, neutral and cold events respectively. ENSO is an important factor, and the correlation of relative entropy with the square of the Niño 3.4 index is 0.76, 0.67 and 0.38 for Florida, the Nordeste and Kenya, respectively; the low correlation in the case of Kenya may be due to the role of the Indian Ocean [*Goddard and Graham*, 1999]. We comment later about the relation of relative entropy with skill in reproducing observations.

[12] Scatter plots of relative entropy with ensemble mean and variance in Figure 2 show that relative entropy is highly

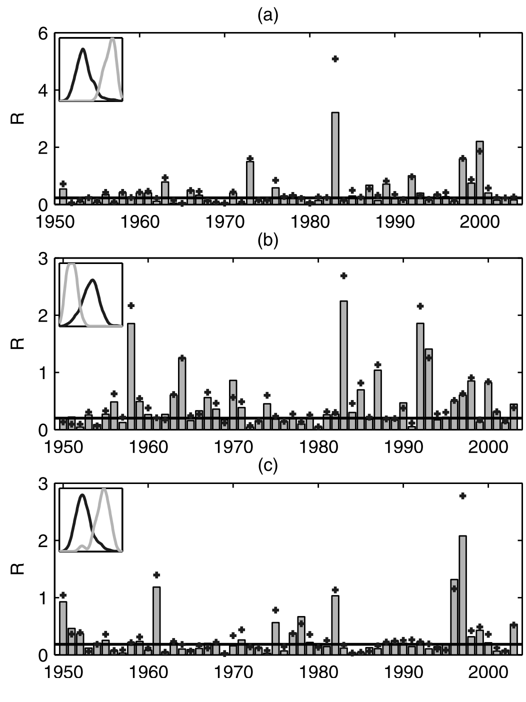

Figure 1. Time series of relative entropy (bars) for (a) Florida, (b) the Nordeste and (c) Kenya; plus signs show the Gaussian approximation in (2). Solid line shows the 95% confidence level. Insets show climatological distributions (black) and forecast distribution (gray) of the year with largest relative entropy.

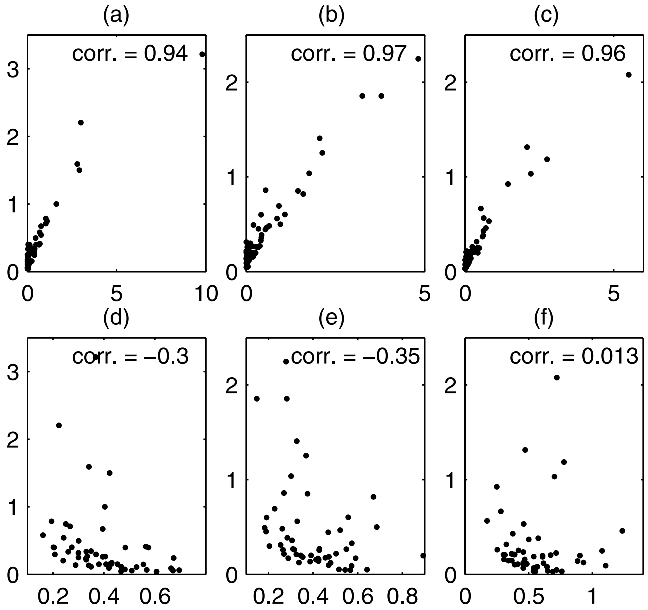

Figure 2. Scatter plots of relative entropy (ordinate) with (a)–(c) the square of the normalized ensemble mean shift (abscissa) for Florida, the Nordeste and Kenya respectively, and with (d)–(f) the normalized ensemble variance.

correlated with the simulation ensemble mean in all three regions. Florida and the Nordeste show a negative correlation (~ -0.3) between ensemble variance and relative entropy. Large ensemble variance is associated with low relative entropy but low ensemble variance is not a good indicator of high relative entropy. For Kenya, the correlation between ensemble variance and relative entropy is approximately zero, though the scatter plot shows some of the same qualitative features seen in the other regions.

[13] The weak relation between ensemble variance and relative entropy suggests that here the dominant contribution to relative entropy is from ensemble mean shifts. The relatively small interannual variability of ensemble variance and modest ensemble size may be factors in this result. *Whitaker and Loughe* [1998] found in several settings that the relation between spread and skill is strong when the variability of ensemble variance is large. To explore the value of higher order moments of the simulation ensemble, we define a constructed ensemble with the same mean as the simulation ensemble but whose distribution about that mean is fixed and is estimated from the climatological distribution of ensemble members about their mean. We use relative entropy to compare the simulation and constructed ensembles; the reference distribution q in (1) is now the constructed ensemble distribution rather the climatological one, and the relative entropy tells how much the simulation and constructed ensemble distributions differ. Monte Carlo significance levels for the difference are constructed in a similar manner as before. Figure 3 shows that the relative entropy between the simulation ensemble and constructed ensemble is small with few years being significant; there are fewer years (3, 5 and 4 for Florida, Nordeste and Kenya respectively) where both the relative entropy between the simulation ensemble and constructed ensemble, and the relative entropy itself are significant. This comparison between the constructed and simulation ensembles is equivalent to computing the relative entropy between the climatological and SST-forced distributions with their means removed.

[14] We now briefly examine the relation between relative entropy and the ability of the model to reproduce observations. Figure 4 shows the ensemble mean, standard deviation and observed anomaly for the five years with highest relative entropy and the five years with lowest relative entropy. Years with high relative entropy show large shifts in the ensemble mean, while years with small relative entropy show small shifts in the ensemble mean and some expansion of the ensemble spread relative to the model climatology. High relative entropy is a perfect model measure and does not guarantee skill; note the large prediction errors for Florida 1992 and Kenya 1961. Model performance in many of the years with small relative entropy was "good" in the sense that the observations were within a standard deviation of the ensemble mean. However, the utility as measured by relative entropy was small in those years because the SST-forced distribution was little different from climatology. Those years also contribute little to the observed correlation r_{obs}. Consider the terms that appear in the expression for the correlation r_{obs} (Y. Tang et al., On the reliability of ENSO dynamical predictions, submitted to *Journal of Atmospheric Sciences*, 2004)

$$r_{\mathrm{obs}} = \frac{1}{\sigma_o \sigma_{\mathrm{mean}}} \sum_i O_i \mu_i \,, \tag{3}$$

where O_i and μ_i are the observations and ensemble mean respectively for year i, and σ_o and σ_{mean} are their standard deviation. The time-correlation of the terms in (3) with R is

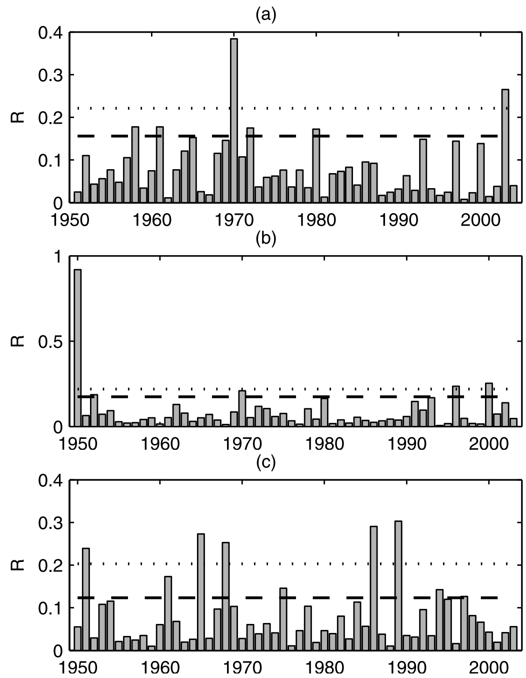

Figure 3. Relative entropy of the simulation ensemble with respect to the constructed ensemble for (a) Florida, (b) the Nordeste and (c) Kenya. Dashed and dotted lines show respectively the 95% and 99% confidence levels.

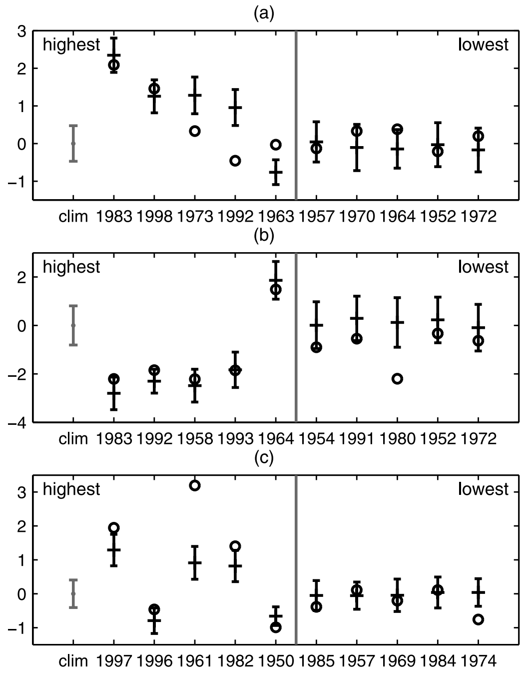

Figure 4. . The years with the highest (left to right) and lowest (right to left) potential relative entropy for (a) Florida, (b) the Nordeste and (c) Kenya. Error bars mark the ensemble mean precipitation anomaly plus and minus the standard deviation of the ensemble; the climatological standard deviation is in gray. The observed precipitation anomaly is marked with a circle. Units are mm/day.

high (0.94, 0.87 and 0.89 for Florida, the Nordeste and Kenya, respectively) indicating that relative entropy is large (small) in those years that contribute most (least) to the observed correlation.

4. Summary and Discussion

[15] We have used relative entropy to measure the impact of SST on GCM simulated seasonal precipitation in three regions. The impact is statistically significant in half or more of the years. However, large values of relative entropy are observed in only a handful of years. This behavior is likely due to relative entropy depending on the square of the normalized ensemble mean anomaly. Relative entropy is highly correlated with shifts in the ensemble mean precipitation. The relation between relative entropy and ensemble variance is weak, although large ensemble variance generally indicates low utility.

[16] We compared the simulation ensemble with a constructed ensemble having the same mean but with a fixed distribution and found little difference as measured by relative entropy, indicating little detectable year-to-year variation of higher order distribution moments (e.g., spread, shape) with this size ensemble. This conclusion is similar to that of *Kumar et al.* [2000] who found that SST-forced changes in height distribution variance in the Pacific-North America region were modest and had a relatively small impact on the associated categorical probabilities.

[17] Larger ensembles allow better estimation of higher order moments and may permit more robust detection of relative entropy changes related to ensemble spread and shape, though the changes themselves may still be small. This issue may be particularly important when changes in ensemble spread or shape significantly change the probabilities of extreme events; while changes in the probability of extreme events are measured by the relative entropy functional, they are balanced against other changes in the distribution. *Sardeshmukh et al.* [2000] using the NCEP MRF9 GCM found regions where the ENSO-induced change of variability makes as large a contribution to the change in the probability of extreme events as does the ENSO-induced shift of the mean. However, requiring dynamical models to simulate higher order moments of distributions accurately is a significant challenge, and the utility of large ensembles to reproduce observed distributions remains to be established.

[18] **Acknowledgments.** We thank Benno Blumenthal for the IRI Data Library, and Tony Barnston, Arun Kumar and anonymous reviewers for their comments and suggestions. IRI is supported by its sponsors and NOAA Office of Global Programs Grant number NA07GP0213.

References

Bowman, A., and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis*, Oxford Univ. Press, New York.
DelSole, T. (2004), Predictability and information theory, part I: Measures of predictability, *J. Atmos. Sci.*, *61*, 2425–2440.
Goddard, L., and N. E. Graham (1999), The importance of the Indian Ocean for simulating precipitation anomalies over eastern and southern Africa, *J. Geophys. Res.*, *104*, 19,099–19,116.
Kleeman, R. (2002), Measuring dynamical prediction utility using relative entropy, *J. Atmos. Sci.*, *59*, 2057–2072.
Kleeman, R., and A. M. Moore (1999), A new method for determining the reliability of dynamical ENSO predictions, *Mon. Weather Rev.*, *127*, 694–705.
Kumar, A., and M. P. Hoerling (1995), Prospects and limitations of seasonal atmospheric GCM predictions, *Bull. Am. Meteorol. Soc.*, *76*, 335–345.
Kumar, A., A. G. Barnston, P. Peng, M. P. Hoerling, and L. Goddard (2000), Changes in the spread of the variability of the seasonal mean atmospheric states associated with ENSO, *J. Clim.*, *13*, 3139–3151.
Majda, A. J., R. Kleeman, and D. Cai (2002), A mathematical framework for quantifying predictability through relative entropy, *Meth. Appl. Anal.*, *9*, 425–444.
Mason, S. J., and L. Goddard (2001), Probabilistic precipitation anomalies associated with ENSO, *Bull. Am. Meteorol. Soc.*, *82*, 619–638.
New, M. G., M. Hulme, and P. D. Jones (2000), Representing 20th century space-time climate variability. II: Development of 1901–1996 monthly terrestrial climate fields, *J. Clim.*, *13*, 2217–2238.
Roeckner, E., K. Arpe, L. Bengtsson, M. Christoph, M. Claussen, L. Dümenil, M. Esch, M. Giorgetta, U. Schlese, and U. Schulzweida (1996), The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate, *Tech. Rep. 218*, 90 pp., Max-Planck Inst. for Meteorol., Hamburg, Germany.
Ropelewski, C., and M. Halpert (1987), Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation, *Mon. Weather Rev.*, *115*, 1606–1626.
Rowell, D. P. (1998), Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations, *J. Clim.*, *11*, 109–120.
Sardeshmukh, P. D., G. P. Compo, and C. Penland (2000), Changes of probability associated with El Niño, *J. Clim.*, *13*, 4268–4286.
Schneider, T., and S. Griffies (1999), A conceptual framework for predictability studies, *J. Clim.*, *12*, 3133–3155.
Whitaker, J. S., and A. F. Loughe (1998), The relationship between ensemble spread and ensemble mean skill, *Mon. Weather Rev.*, *126*, 3292–3302.

R. Kleeman and Y. Tang, Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA.
M. K. Tippett, International Research Institute for Climate Prediction, Earth Institute of Columbia University, Lamont Campus/61 Route 9W, Palisades, NY 10964, USA. (tippett@iri.columbia.edu)