GEOPHYSICAL RESEARCH LETTERS, VOL. ???, XXXX, DOI:10.1029/,

Measuring the potential utility of seasonal climate predictions

Michael K. Tippett

International Research Institute for Climate Prediction, Columbia University,

Palisades, New York

Richard Kleeman and Youmin Tang

Courant Institute of Mathematical Sciences, NYU, New York, New York

M. K. Tippett, International Research Institute for Climate Prediction, The Earth Institute of Columbia University, Lamont Campus / 61 Route 9W, Palisades New York 10964, USA. (tippett@iri.columbia.edu)

Variation of sea surface temperature (SST) on seasonal-to-interannual time-scales leads to changes in seasonal weather statistics and seasonal climate anomalies. Here we use relative entropy, an information theory measure of utility, 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 compute the yearly variation of relative entropy and find that it 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.

1. Introduction

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 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 a climatological distribution q. This climatological distribution is then compared with the distribution p of precipitation amounts given a particular SST. The impact of the SST on seasonal precipitation is measured by the extent to which the two distributions differ. If the SST has no impact on precipitation, then the two distributions will be identical. On the other hand, if SST has an impact on precipitation amounts, then the two distributions will be significantly different. The *relative entropy* measures the informational inefficiency of using the climatological distribution instead of the SST forced distribution [*Kleeman*, 2002].

Some important properties of relative entropy are that it is invariant with respect to invertible transformations, it is sensitive to changes in distribution shape as well as changes in mean and variance, and it vanishes only when the two distributions are identical [*Kleeman*, 2002; *Majda et al.*, 2002]. 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]

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

where μ_p^2 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 form of (2) shows that relative entropy is increased when the SST forcing reduces the variance ($\sigma_p < \sigma_q$) or shifts the mean ($\mu_p^2 > 0$). In general, the relative importance of the contributions to relative entropy from changes in the mean and variance depends on dynamical properties of the system [*Kleeman*, 2002].

The calculation of relative entropy requires specifying the distribution p of precipitation amounts given a particular SST. However, since nature provides only a single realization of precipitation for a given SST, general circulation models (GCMs) are forced with observed SST conditions and used to estimate the distribution of precipitation amounts given a particular SST [*Kumar and Hoerling*, 1995; *Anderson and Stern*, 1996; *Rowell*, 1998; *Sardeshmukh et al.*, 2000]. Relative entropy 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.

Here we compute the relative entropy for three regions where SST has a large impact on precipitation: South Florida, the Nordeste of Brazil and Kenya. Goals of this work are: to quantify the yearly variation of potential information 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

2.1. Model and observational data

The three regions and seasons we examine are: southern Florida (Dec-Feb; DJF), the Nordeste of Brazil (Mar-May; MAM) and Kenya (Oct-Dec; OND). Model data is taken from a 24 member ensemble of T42 ECHAM 4.5 GCM simulations forced with observed SSTs for the period January 1950 to March 2004 [*Roeckner et al.*, 1996]. Precipitation observations are taken from the extended *New et al.* [2000] gridded dataset of monthly precipitation for the period of 1950 to 1998. Model and gridded observed precipitation are averaged over the spatial domains and seasons indicated in Table 1.

The precipitation response of the GCM to tropical SST anomalies is apparent either when the ensemble mean is compared to individual ensemble members or when the ensemble mean is compared to observations. The size of the SST forced response relative to the model's own internal variability is measured by the signal-to-noise ratio $S \equiv \sigma_{\text{mean}}/\sigma_{\text{ens}}$; σ_{mean} is the interanual standard deviation of the ensemble mean and σ_{ens} is the standard deviation of the ensemble about its mean agrregrated over all cases. The signal-to-noise ratio determines the "perfect model" correlation r_{perfect} of the ensemble mean with any ensemble member [*Kleeman and Moore*, 1999; *Sardeshmukh et al.*, 2000]:

$$r_{\text{perfect}} = \frac{S}{\sqrt{1+S^2}}\,. \tag{3}$$

The correlation r_{obs} between ensemble mean and observations is typically less than r_{perfect} due to model deficiencies, though in practice both quantities are affected by sampling error. Both the perfect model correlation r_{perfect} and the observed correlation r_{obs} are high (Table 1) for these regions, indicating that the model is able to reproduce itself and observations.

2.2. Computing relative entropy

The climatological distribution q is estimated from all ensemble members and years, and the sample size is 1296 (54×24). The distribution p of precipitation amounts in response to SST forcing is estimated for each year from a sample of size 24. Both distributions are approximated with a kernel density estimate using a normal kernel function [*Bowman and Azzalini*, 1997]. The integral in (1) is computed using the estimated distributions evaluated at 100 equally-spaced points covering the range of the model climatology. We note that this kernel density estimate would be inappropriate for a quantity like daily rainfall whose distribution is far from Gaussian. However, distributions of seasonal totals are much closer to being Gaussian. We also computed relative entropy using the Box-Cox transformed data (the estimated Box-Cox parameter λ was 1/3, 1 and 0 for the Florida, Nordeste and Kenya data respectively) and found negligible (< 0.03) differences.

Although the relative entropy should be zero when the simulation and climatology distributions are identical, finite ensemble size introduces sampling error. *Kleeman and Majda* [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 50,000 times, and the sorted results indicate the likelihood that relative entropy exceeds a given value by chance. We consider values above the 95th percentile as significant.

3. Results

The time-series of relative entropy in Fig. 1 shows that the relative entropy of the simulation with respect to climatology is significant in 59% (32/54) of the years for Florida, 78% (41/54) of the years for the Nordeste and 54% (29/54) of the years for Kenya. Relative entropy is very large for Florida and Kenya in only a handful of years. In the case of Florida, the three largest years, 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; in OND 1996 the model predicted a shift toward below-normal precipitation while the observed precipitation was close to normal. ENSO is an important factor, and the correlation of relative entropy with the Niño 3.4 index is 0.76, 0.67 and 0.39, 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.

Scatter plots of relative entropy with ensemble mean and variance in Fig. 2 show that relative entropy is highly 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. In the case of Kenya, the correlation between ensemble variance and relative entropy is approximately zero, though the scatter plot of ensemble variance and relative entropy shows some of the same qualitative features.

The weak relation between ensemble variance and relative entropy suggests that the dominant contribution to relative entropy is from shifts of the ensemble mean. The relatively small interannual variability of ensemble variance and the 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* whose mean is the same as that of the simulation ensemble but whose distribution about that mean is fixed and is estimated from the climatological distribution of ensemble members about their mean. A parametric description would be an alternative construction when the distributions are normal [*Kharin and Zwiers*, 2003]. We now use relative entropy to compare the simulation and and constructed ensembles. That is, 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. Significance levels for the difference are constructed in a similar manner as before with a Monte Carlo computation. Figure 3 shows that the relative entropy between the simulation ensemble and constructed ensemble is small with few years being significant. We expect 2-3 years would appear significant at the 95% level by chance in a time-series of this length. These results suggest that with this ensemble size utility of higher order moments is seldom significant.

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. 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, in those years the utility was small because the ensemble 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 observed correlation r_{obs} [*Tang et al.*, 2004]

$$r_{\text{obs}} = \frac{1}{\sigma_o \sigma_{\text{mean}}} \sum_i O_i \mu_i \,, \tag{4}$$

where O_i and μ_i are the observations and ensemble mean respectively at time-step i, and σ_o and σ_{mean} are their standard deviations. The time-correlation of the terms in (4) with R is high (0.87, 0.8 and 0.85 for 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

We have used relative entropy to measure the impact of SST on precipitation simulated by a 24-member GCM ensemble. The impact is statistically significant in more than half of the years. However, the time-series of relative entropy is dominated by a handful of years associated with substantial shifts in the precipitation amount distribution. 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.

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. This result indicates that the utility of higher order moments (e.g. spread, shape) of the distribution is small with an ensemble of this size. The weakness of the SST impact on higher order moments of the precipitation distribution is similar to the conclusion of *Kumar et al.* [2000] who found little correlation between SST anomalies and height distribution variance in the Pacific-North America region. Though compositing did show warm events reduced height distribution variance, the impact on categorical probabilities was relatively small.

The modest ensemble size in this study means that higher order moments are poorly estimated. Larger ensembles may allow detection of changes in relative entropy related to ensemble spread and shape. This issue may be particularly important when the ensemble mean shift is small but changes in ensemble spread or shape significantly change the probabilities of extreme events. *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.

Acknowledgments. We thank Benno Blumenthal for the IRI Data Library and Tony Barnston for his comments and suggestions. IRI is supported by its sponsors and NOAA Office of Global Programs Grant number NA07GP0213.

References

Anderson, J. L., and W. F. Stern (1996), Evaluating the potential predictive utility of ensemble forecasts, *J. Climate*, *9*, 260–269.

Bowman, A., and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis*, Oxford University Press.

DelSole, T. (2004), Predictability and Information Theory Part I: Measures of Predictability, *J. Atmos. Sci.*, in press.

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.

Kharin, V. V., and F. W. Zwiers (2003), Improved seasonal probability forecasts, *J. Climate*, *16*, 1684–1701.

Kleeman, R. (2002), Measuring dynamical prediction utility using relative entropy, *J. Atmos. Sci.*, *59*, 2057–2072.

Kleeman, R., and A. J. Majda (2004), Predictability in a model of geostrophic turbulence, *J. Atmos. Sci.*, submitted.

Kleeman, R., and A. M. Moore (1999), A new method for determining the reliability of dynamical ENSO predictions, *Mon. Wea. Rev.*, *127*, 694–705.

Kumar, A., and M. P. Hoerling (1995), Prospects and limitations of seasonal atmospheric GCM predictions, *Bull. Amer. Met. 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. Climate*, *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. Amer. Met. 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. Climate*, *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*, Max-Planck Institute for Meteorology, Hamburg, Germany, 90 pp.

Ropelewski, C., and M. Halpert (1987), Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation., *Mon. Wea. Rev.*, *115*, 1606–1626.

Rowell, D. P. (1998), Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations, *J. Climate*, *11*, 109–120.

Sardeshmukh, P. D., G. P. Compo, and C. Penland (2000), Changes of probability associated with El Niño, *J. Climate*, *13*, 4268–4286.

Schneider, T., and S. Griffies (1999), A Conceptual Framework for Predictability Studies, *J. Climate*, *12*, 3133–3155.

Tang, Y., R. Kleeman, and A. M. Moore (2004), On the reliability of ENSO dynamical predictions, *J. Atmos. Sci*, submitted.

Whitaker, J. S., and A. F. Loughe (1998), The relationship between ensemble spread and ensemble mean skill, *Mon. Wea. Rev.*, *126*, 3292–3302.

Table 1. Domains and seasons.

Region	Domain	Season	r_{perfect}	r_{obs}
Florida	85W-75W, 22N-28N	DJF	0.76	0.75
Nordeste	45W-35W, 10S-EQ	MAM	0.81	0.69
Kenya	33E-43E, 5S-5N	OND	0.67	0.84

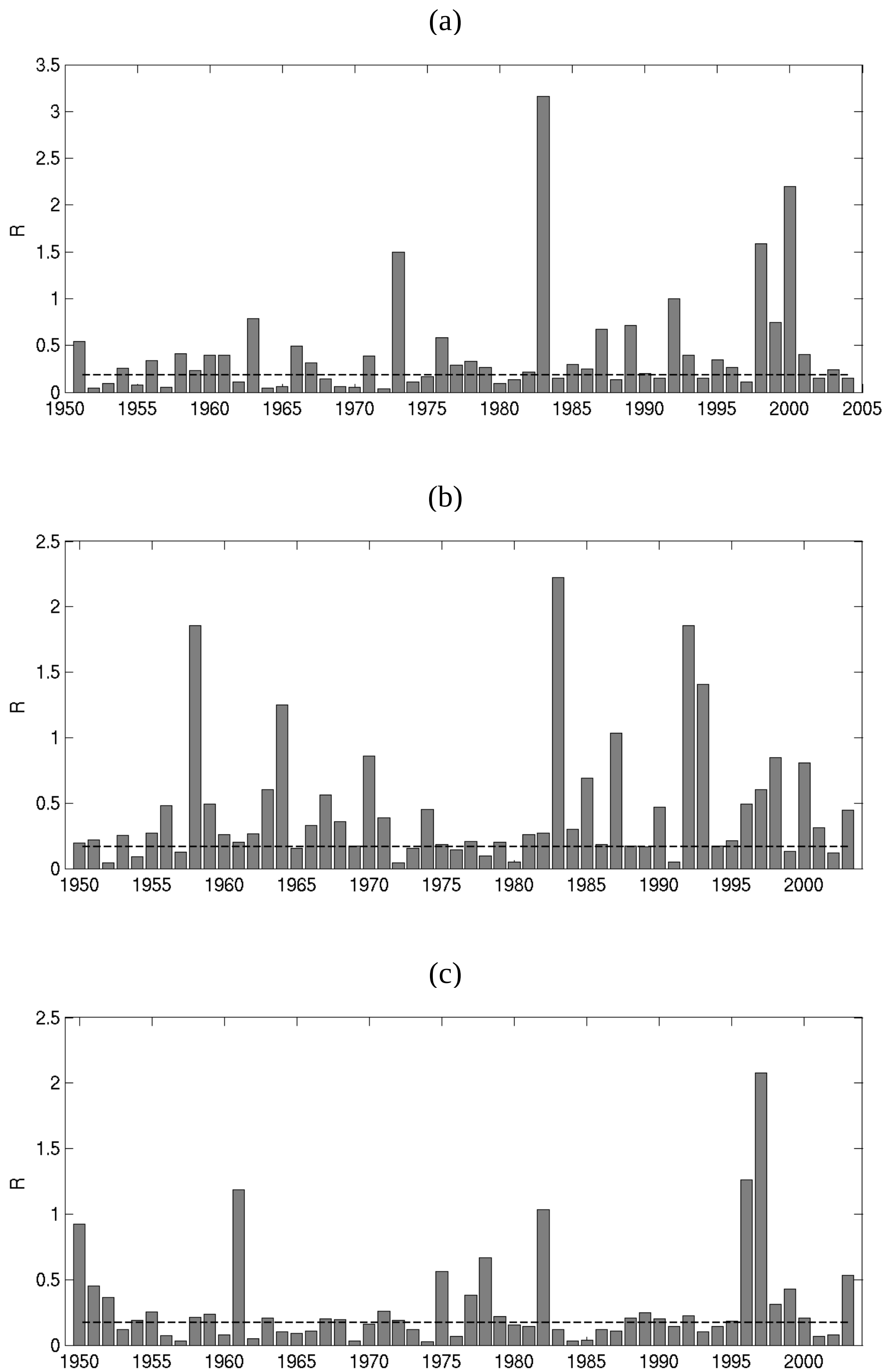


Figure 1. Time-series of relative entropy for (a) Florida, (b) the Nordeste and (c) Kenya. Dashed line shows the 95% confidence level.

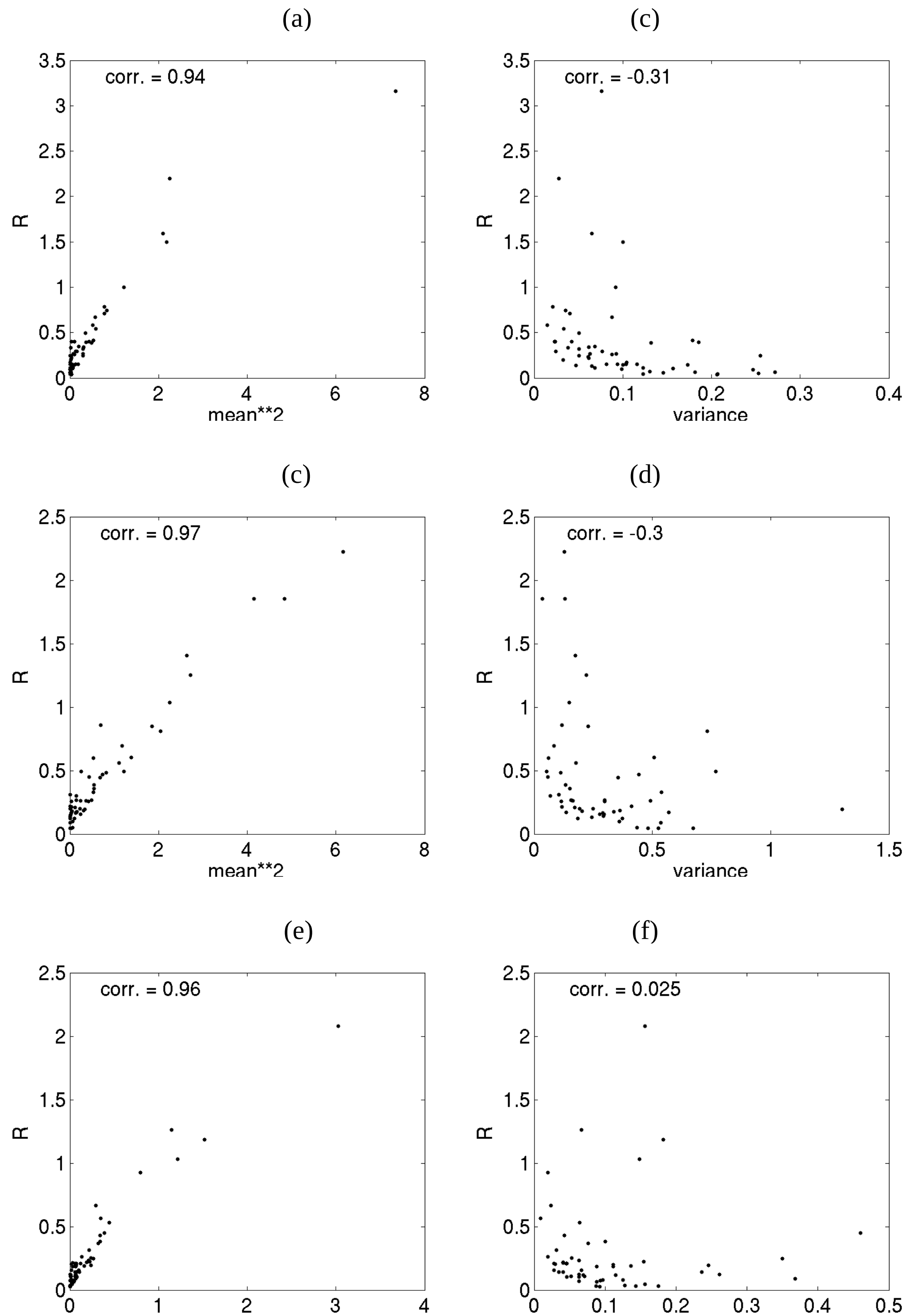
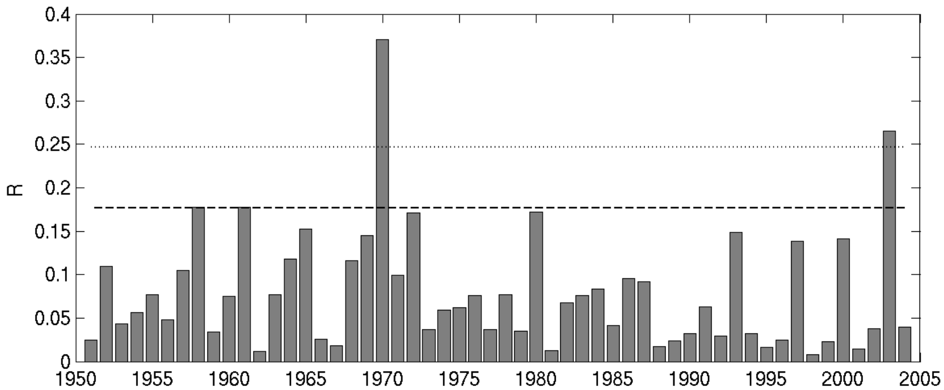
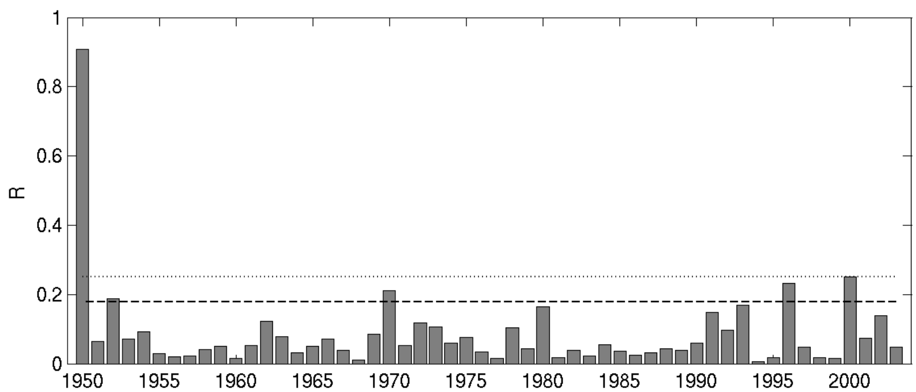


Figure 2. Scatter plots of relative entropy with (a) the square of the normalized ensemble mean shift and (b) the normalized ensemble variance for Florida. (c) and (d) as in (a) and (b) but for the Nordeste. (e) and (f) as in (a) and (b) but for Kenya.

(a)



(b)



(c)

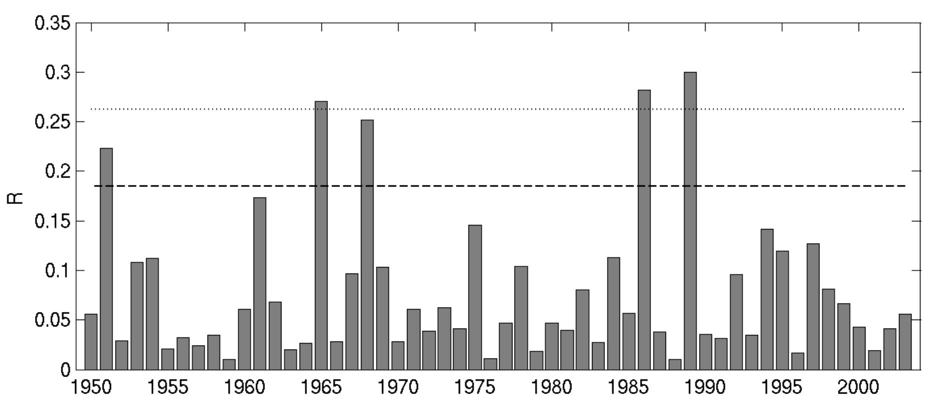
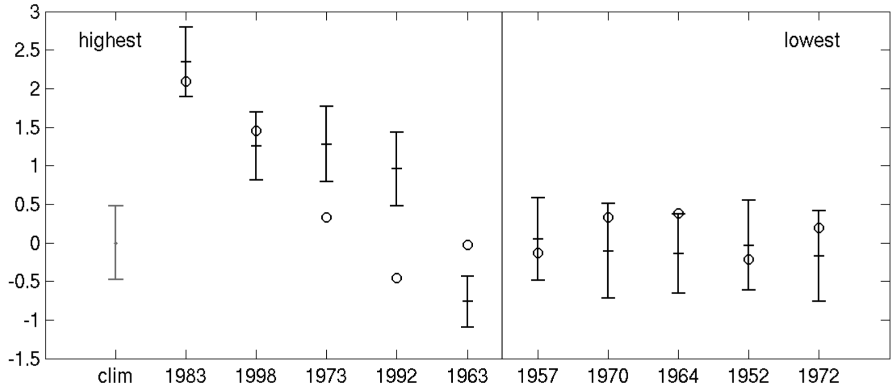
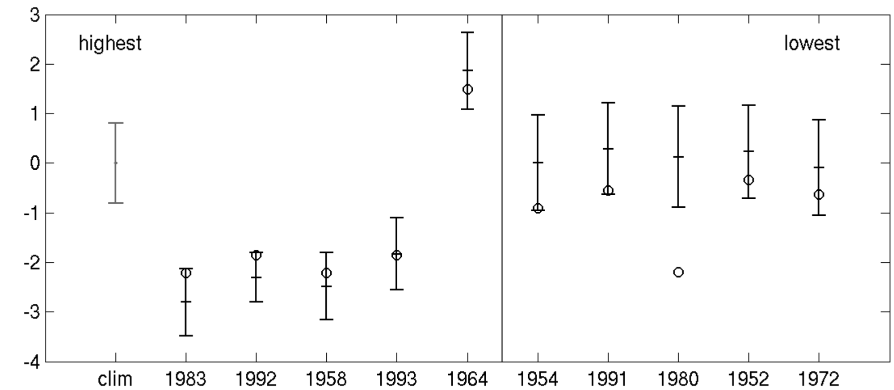


Figure 3. Time-series of the simulation ensemble relative entropy with respect to the constructed ensemble for (a) Florida, (b) the Nordeste and (c) Kenya. Dashed and dotted lines show respectively the 95% and 99% confidence levels.

(a)



(b)



(c)

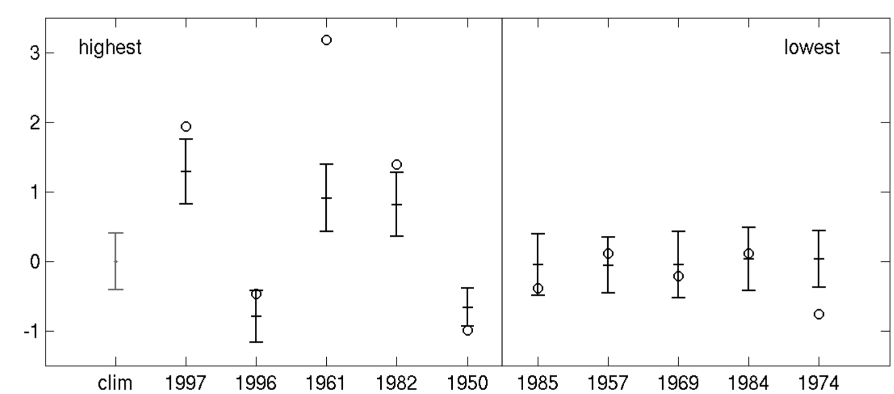


Figure 4. The years with the highest (left to right) and lowest (right to left) potential relative entropy for (a) Florida, (b) the Nordeste and (c) Kenya. Error bars mark the ensemble mean precipitation anomaly plus and minus the standard deviation of the ensemble; the climatological standard deviation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