Climate models produce skillful predictions of Indian summer monsoon rainfall

Timothy DelSole[1,2] and Jagadish Shukla[1,2]

Received 8 February 2012; revised 26 March 2012; accepted 27 March 2012; published 1 May 2012.

[1] After more than one hundred years of statistical forecasting and fifty years of climate model development, this paper shows that the skill of predicting Indian monsoon rainfall with coupled atmosphere-ocean models initialized in May is statistically significant, and much higher than can be predicted empirically from May sea surface temperatures (SSTs). The superior skill of dynamical models is attributed to the fact that slowly evolving sea surface temperatures are the primary source of predictability, and to the fact that climate models produce more skillful predictions of June-September sea surface temperatures. The recent apparent breakdown in SST-monsoon relation can be simulated in coupled models, even though the relation is significant and relatively constant on an ensemble mean basis, suggesting that the observed breakdown could be due, in large part, to sampling variability. Despite the observed breakdown, skillful predictions of monsoon rainfall can be constructed using sea surface temperatures *predicted* by dynamical models. This fact opens the possibility of using readily available seasonal predictions of sea surface temperatures to make real-time skillful predictions of Indian summer monsoon rainfall. In addition, predictors based on tendency of SST during spring information show skill during both the recent and historical periods and hence may provide more skillful predictions of monsoon rainfall than predictors based on a single month. **Citation:** DelSole, T., and J. Shukla (2012), Climate models produce skillful predictions of Indian summer monsoon rainfall, *Geophys. Res. Lett.*, *39*, L09703, doi:10.1029/2012GL051279.

1. Introduction

[2] Fluctuations in the Indian Summer Monsoon Rainfall (ISMR) can have devastating impacts on the economic and societal welfare of India. For over a hundred years, the Indian Meteorological Department has used statistical models to predict ISMR. Unfortunately, these models have shown little consistent skill [*Gadgil et al.*, 2005]. Similarly, dynamical models have advanced significantly over the past few decades, but continue to show substantial biases in simulating ISMR [*Drbohlav and Krishnamurthy*, 2010]. In this paper, we use statistical optimization techniques to show that no significant May sea surface temperature (SST) pattern for predicting ISMR can be statistically justified in the recent period. Despite the lack of significant SST predictors, we show that the current generation of coupled atmosphere-ocean models can predict ISMR with skill when initialized in May, as noted also by *Rajeevan et al.* [2012]. The skill arises from the fact that though May SSTs are not useful predictors, June-September SSTs are significantly related to ISMR and can be predicted more skillfully by dynamical models. This result provides the basis for making real-time skillful predictions of ISMR with dynamical models.

2. Data

[3] The observational rainfall data set used in this study is the monthly mean subdivisional rainfall from the Indian Institute of Tropical Meteorology (IITM). This data was downloaded from http://www.tropmet.res.in and is available for 1871–2010. The all-India estimates were computed by summing the appropriate area weighted total rainfall of each subdivision. Only the June-September (JJAS) total rainfall is considered in this study. This quantity will be called the Indian Summer Monsoon Rainfall (ISMR) index.

[4] The observational SST used in this study is the $2° \times 2°$ Extended Reconstruction SST (ERSSTV3, which includes satellite data) of *Smith et al.* [2008]. This data was downloaded from http://www.ncdc.noaa.gov/ersst/ and is available from 1871 to present. Only the tropical Indo-Pacific region 30°S–30°N and 30°E–60°W are analyzed.

[5] The model simulations used in this paper are the hindcasts from the ENSEMBLES project. This data set, described in *Weisheimer et al.* [2009], consists of nine-member ensemble hindcasts of five state-of-the-art coupled atmosphere-ocean general circulation models during the period 1960–2005. The five models are from the European Centre for Medium-Range Weather Forecasts (ECMWF), Météo France (MF), UK Met Office (UKMO), the Leibniz Institute of Marine Sciences at Kiel University (IFM-GEOMAR), and the Euro-Mediterranean Centre for Climate Change (CMCC-INGV) in Bologna. Only hindcasts initialized on May 1 of each year were considered, as these were most relevant to summer monsoon prediction. The model resolutions differ, as documented in *Weisheimer et al.* [2009], but the outputs have been interpolated onto a common $2.5° \times 2.5°$ grid prior to analysis. The model ISMR is estimated by aggregating precipitation over land within 70°E–90°E and 10°N–25°N. Major conclusions are not sensitive to domain boundaries.

[6] Model simulations overlap with observational data only during 1960–2005. Data during the earlier period

[1]Department of Atmospheric, Oceanic and Earth Sciences, George Mason University, Fairfax, Virginia, USA.

[2]Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland, USA.

Corresponding Author: T. DelSole, Department of Atmospheric, Oceanic and Earth Sciences, George Mason University, 4400 University Dr., Fairfax, VA 22030, USA. (tdel-sole@mason.gmu.edu)

Copyright 2012 by the American Geophysical Union.
0094-8276/12/2012GL051279

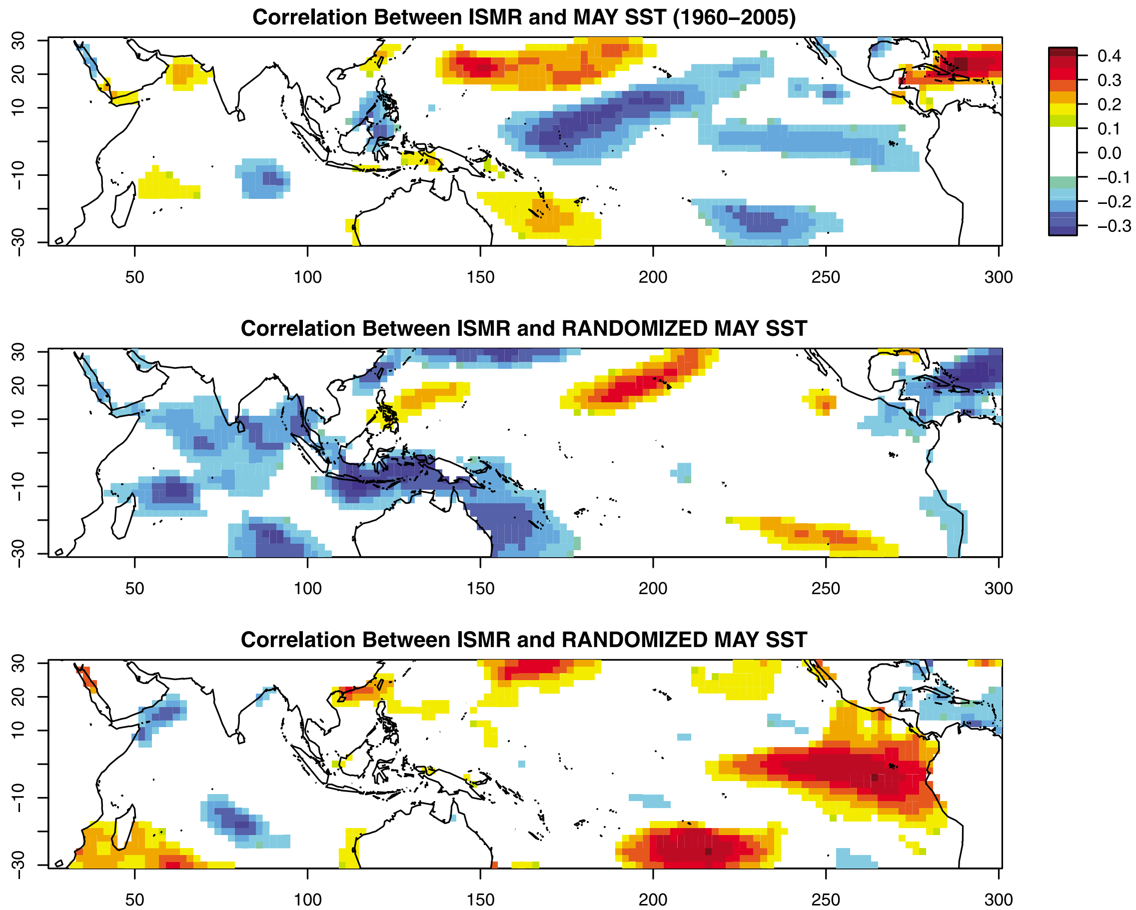

Figure 1. (top) Correlation between JJAS all-Indian Monsoon Rainfall and May tropical sea surface temperature during the period 1960–2005 and (middle and bottom) analogous correlation maps derived from *randomized* May SST fields. The same color scale is used in each panel. Insignificant correlations have been masked out.

1880–1959 will be used for supplemental observational analysis.

3. Results

3.1. Statistical Predictions

[7] The search for predictors of ISMR based on antecedent climate variables goes back at least to *Walker* [1910]. Though it was not known at the time, several predictors used by Walker were proxies for SST. The possibility that SSTs could affect monsoons and could be predicted beyond the two-week limit of atmospheric predictability received theoretical support from *Charney and Shukla* [1981]. A widely used approach to identifying predictors of ISMR is to construct correlation maps between ISMR and antecedent SST, isolate local regions with large correlations, compute spatial averages over these regions, and then use the resulting indices as predictors. Unfortunately, this methodology is a form of screening, also called data-fishing. *DelSole and Shukla* [2009] showed that this kind of screening typically leads to such large bias in skill measures that many prediction models derived from screened predictors would be indistinguishable from a no-skill model constructed from random numbers if the prior screening were taken into account.

[8] To illustrate the problem with screening, we show in Figure 1 the correlation map between observed ISMR and May SST during the period 1960–2005, as well as two further correlation maps derived from the same data but with *randomized* SSTs (i.e., SST fields drawn with replacement from randomly selected years). Since SST fields in the latter cases are randomized, there can be no real relation between them and ISMR. Despite this independence, the figure shows that correlation maps computed from randomized SST give local correlations as large as those derived from the original data. Such spurious correlations are an anticipated consequence of computing a large number of correlations. In our particular case, there are 3,424 ocean grid cells available for computing correlations, so we would expect about 171 grid cells to exceed the 5% significance level even if SST were independent of ISMR. Moreover, SST is spatially correlated, which means that if a particular grid cell is correlated with a predictand, then neighboring points also tend to be correlated with the predictand, giving a false impression of a physical relation.

[9] Recently, *DelSole and Yang* [2011] developed a rigorous, multivariate procedure for testing the hypothesis that a field of regression coefficients vanish simultaneously. This *field significance* test effectively implies that a collection of regression coefficients differs significantly from zero when the coefficient of determination between the index and the other variables is statistically significant. To reduce overfitting, the predictors are replaced by a few leading principal components, the precise number being determined from

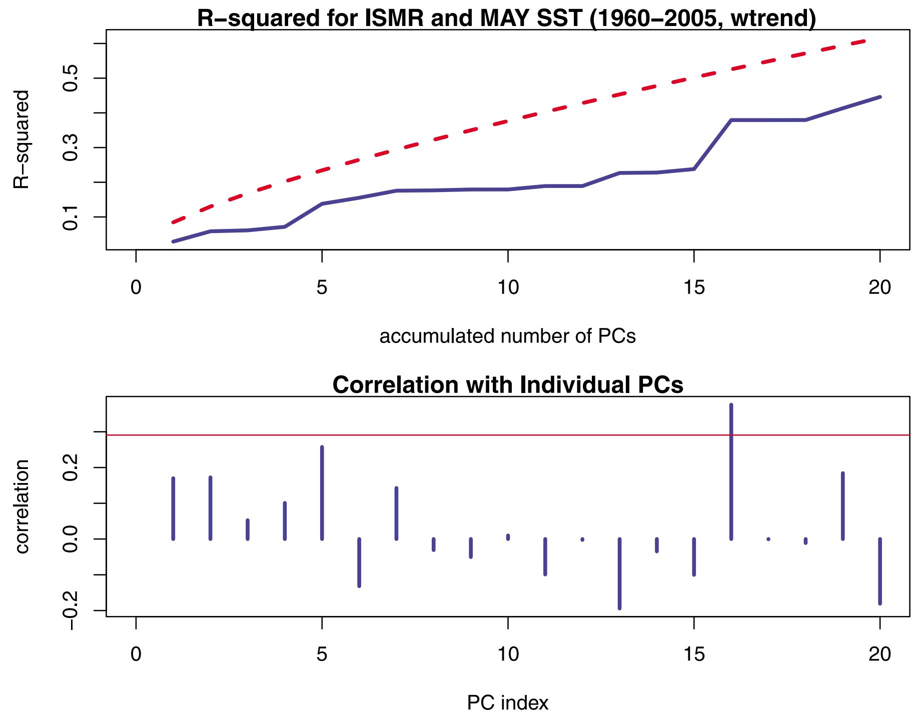

Figure 2. Results of predicting ISMR using a linear combination of the leading PCs of May SST during 1960–2005. (top) The coefficient of determination, R-squared, of a regression model for predicting ISMR based on the leading May SST PCs (blue solid), and the associated 5% significance level (red dashed). (bottom) The pair-wise correlation between ISMR and each individual PC (blue bar), and the 5% significance level (red line).

cross validation experiments. The procedure turns out to be equivalent to CCA, with the coefficient of determination equalling the squared canonical correlation. As a result, this procedure yields the linear combination of principal components of SST that is most correlated with ISMR. In this way, the best linear SST predictor of ISMR is obtained.

[10] The coefficient of determination for predicting ISMR as a function of the number of leading principal component predictors is shown as the blue curve in Figure 2. Because the coefficient is consistently less than the 5% significance level (red dashed curve), the figure shows that there is *no set of leading PCs of May SSTs that gives statistically significant predictions of ISMR during* 1960–2005. Cross validation experiments (not shown) also indicate that no set of leading PCs give mean square errors less than those of predictions based on climatology. Finally, even pair-wise correlations between ISMR and individual PCs of May SST, shown in the bottom of Figure 2, reveal that only one out of 20 exceeds the 5% significance level, exactly the number expected by random chance. These results imply that the correlation map shown in Figure 1 is not significant, and that justifying predictors of ISMR based on correlations with May SSTs will be difficult for 1960–2005.

[11] A similar analysis for the (longer) independent period 1880–1959 yields very different results: the cross validated skill is maximized at three PCs of May SST, and the

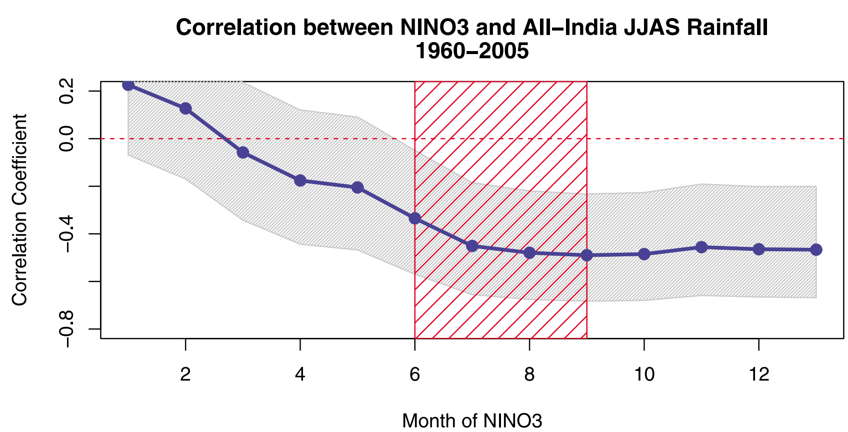

Figure 3. Time lagged correlation between all-India JJAS Monsoon Rainfall and monthly mean NINO3 index during the period 1960–2005. The red hatching indicates the JJAS period, the horizontal red dashed line indicates zero, and the grey shading indicates the 95% confidence interval for the time lagged correlation.

Table 1. Correlation Between JJAS All-India Monsoon Rainfall and Various Indices of SST[a]

SST Predictor in Indo-Pacific	Trained	1880–1959 April	1880–1959 May	1880–1959 JJAS	1960–2005 April	1960–2005 May	1960–2005 JJAS
Canonical	1960–2005	–	–	–	NA	NA	0.39*
Canonical	1880–1959	0.47*	**0.54**	**0.66**	0.10	0.15	**0.34**
NINO3	–	−0.21	**−0.40**	**−0.60**	−0.18	−0.21	**−0.46**
Canonical time-lagged (w/ Jan)	1960–2005	–	–	–	0.46*	0.54*	0.6
NINO3 time-lagged (w/ Jan)	1880–1959	**0.24**	**0.43**	**0.60**	**0.37**	**0.30**	**0.47**
NINO3 time-lagged (w/ Jan)	1960–2005	**0.44**	**0.38**	**0.49**	0.21	**0.30**	**0.56**

[a]"NA" indicates that no linear combination of leading principal components of SST had positive cross validated skill. Asterisks indicate the canonical predictor had positive cross validated skill score and statistically significant R-squared, but is not significant according to the criterion of *DelSole and Yang* [2011]. Bold indicates correlations that are significant at the 5% level.

correlation skill of this model is 0.54, implying that about 30% of the variance of ISMR can be predicted based on May SSTs. Nearly the same correlation skill is obtained for the 46-year period 1880–1925, implying that the difference is not caused solely by sample size. This dramatic difference could reflect a change in the SST-monsoon relation, or be due to sampling variability. This point will be discussed in more detail in section 4.

[12] In regards to the earlier 1880–1959 period, the optimal May SST predictor has most of its loading in the eastern tropical Pacific (not shown), and projects strongly on the NINO3 index. For comparison, the correlation between ISMR and May NINO3 during 1880–1959 is −0.40, which is weaker than the canonical correlation (as expected), but statistically significant. In many other cases discussed below, the NINO3 index has a similar correlation with ISMR as other predictors derived from CCA. Based on this, we adopt the NINO3 index as a convenient (suboptimal) SST index for predicting ISMR.

[13] The time-lagged correlation between ISMR and NINO3, shown in Figure 3, shows that the lag correlations are statistically insignificant during antecedent months, but strengthen and become significant during the monsoon season. Unfortunately, these latter NINO3 values are not available in practical prediction situations (but see section 4).

[14] Various other SST predictors were investigated, as summarized in Table 1, but no SST predictor based on a single antecedent month could be statistically justified during 1960–2005. However, predictors based on *time-lagged* antecedent months were found to be statistically significant. For instance, a regression model based on a linear combination of antecedent May and January NINO3 values had significant R-squared values in 1960–2010, and the same model trained in 1960–2010 had significance correlation

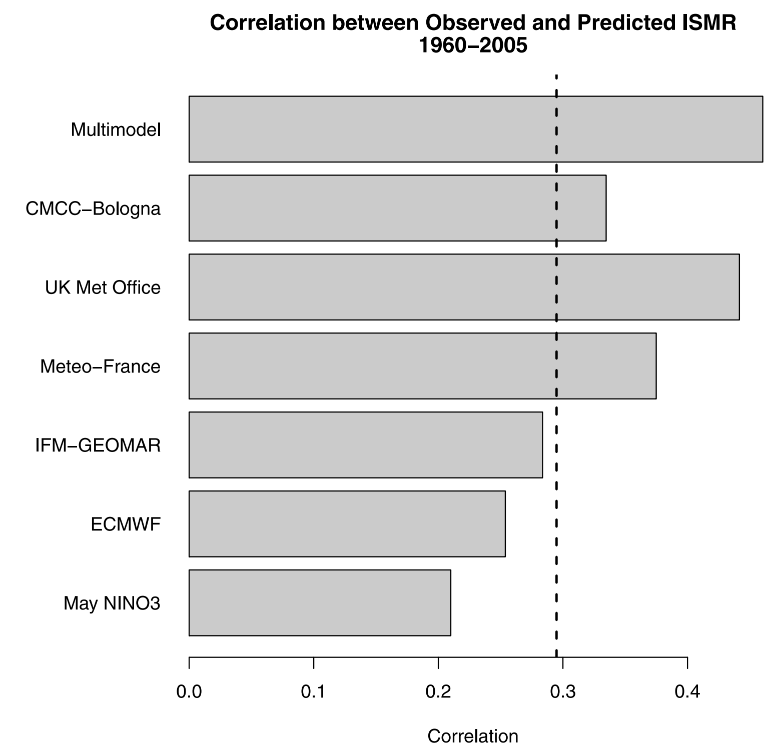

Figure 4. Correlation between observed and predicted JJAS all-India rainfall for hindcasts in the ENSEMBLES data set for the period 1960–2005, and for a linear regression model based on May NINO3. All-India rainfall in dynamical models is defined as the total land precipitation within 70°E–90°E and 10°N–25°N. The 5% significance threshold is indicated by the vertical dashed line. The ensemble mean ISMR was used from each model, while the mean of all model ISMR values was used to compute the "multimodel" value.

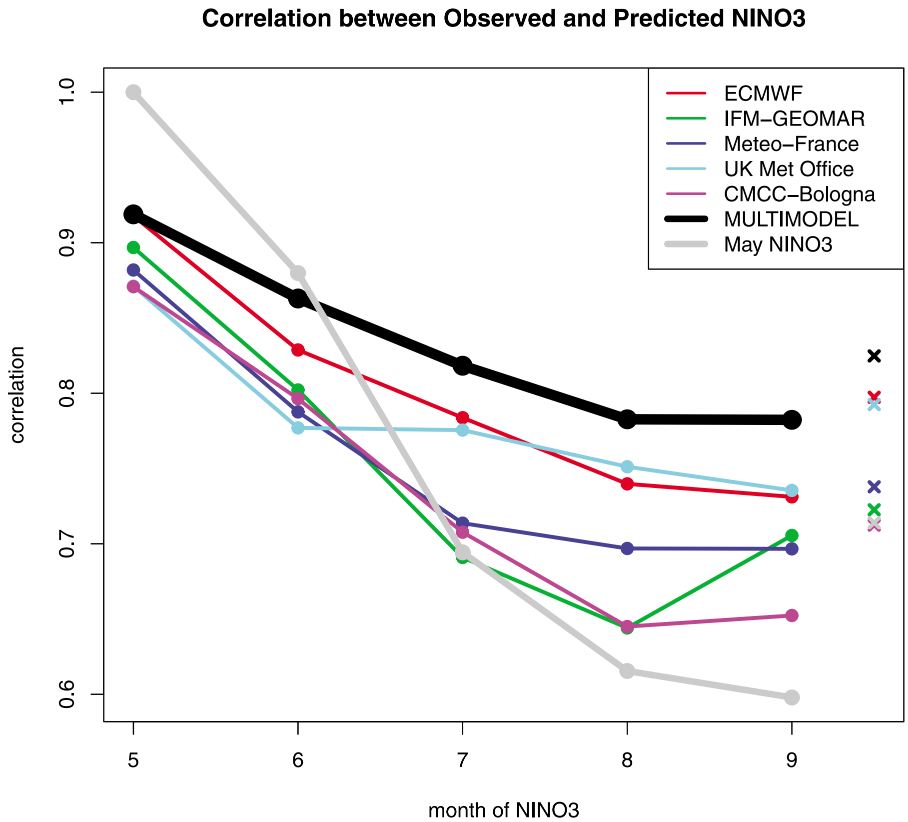

Figure 5. Correlation between observed NINO3, and ensemble mean NINO3 predicted by the ENSEMBLES models, for hindcasts in the period 1960–2005, as a function of calendar month. Also shown is the correlation between observed NINO3 and the least squares prediction of NINO3 based on the observed May NINO3 value (thick grey). The 'x'-symbols on the far right give the correlations between the observed and predicted JJAS NINO3 index.

skill in the independent period 1880–1959. Skill was maintained even if the training and verification periods were swapped. Correlation skill could be improved further with CCA, using PCs from January SST in addition to the first three PCs of May SST. The coefficients of the time-lagged SSTs tend to be opposite sign, consistent with the suggestion by *Shukla and Paolino* [1983] that the *tendency* of ENSO indices may be a good predictor of ISMR. These results suggest that predictors based on time-lagged SST indices may provide more skillful predictions of ISMR than SST predictors based on a single month.

3.2. Dynamical Predictions

[15] We now consider hindcasts of JJAS all-Indian monsoon rainfall by five state-of-the-art coupled atmosphere-ocean models. The correlation between observed and ensemble mean ISMR for each model is shown in Figure 4. The figure shows that three models predict all-India monsoon rainfall with statistically significant skill, with the multi-model mean prediction giving the strongest correlation of all, namely 0.46, as has been noted previously [*Rajeevan et al.*, 2012]. This correlation skill is well above the 0.15–0.21 values obtained from the best statistical models based on May SST for this period (see Table 1). Thus, dynamical models initialized in May give superior predictions of ISMR compared to statistical models using May predictors. These results imply that dynamical models extract more predictive information from May initial conditions than statistical models.

4. The Basis of Dynamical Prediction Skill

[16] Although antecedent SSTs are weakly related to ISMR, *concurrent* SSTs are significantly related to ISMR (see Figure 3 and Table 1). Presumably, then, the skill of the dynamical models arises from the significant relation between ISMR and June-September SSTs, and the fact that dynamical models can predict these SST more skillfully than statistical models. To investigate this hypothesis, we show in Figure 5 a comparison of NINO3 predictions by dynamical and statistical models, where "statistical" refers to linear regression predictions based on May NINO3. The figure shows that while the statistical model is more skillful than dynamical models in June, all dynamical models are more skillful the statistical model in August and September– half the monsoon season. In some cases, the difference in skill is not very large (e.g., 0.60 vs. 0.65 in September), but inferior NINO3 predictions in consecutive months likely produce accumulating inferior predictions of rainfall. This reasoning suggests a possible shortcoming with analyzing only JJAS quantities. Note that ECMWF has the highest skill in predicting NINO3 but the lowest skill in predicting ISMR, suggesting that the SST-monsoon relation is not as accurate in this model as the others. In any case, the results show that

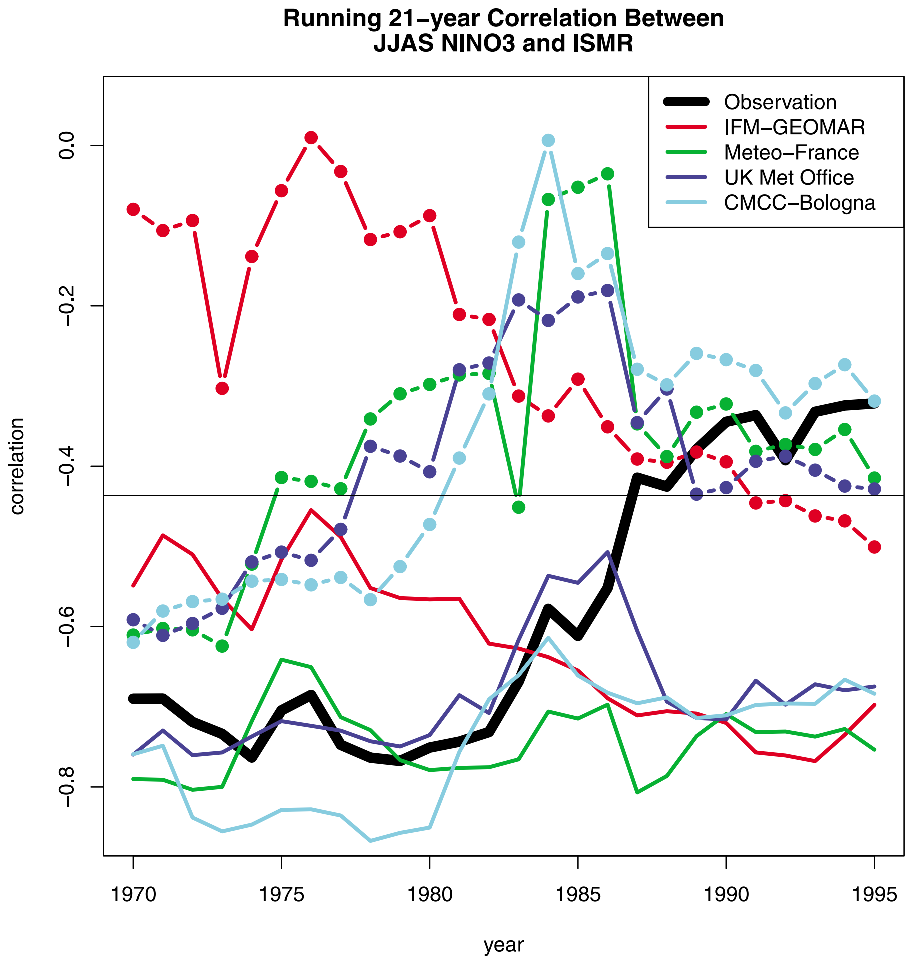

Figure 6. Running 21-year correlations between observed ISMR and observed JJAS NINO3 (thick black), between ensemble mean ISMR and ensemble mean JJAS NINO3 (colored solid lines), and between ISMR and JJAS NINO3 of a selected ensemble member from each model (dots connected by solid lines). All-India rainfall in dynamical models is defined as the total land precipitation within 70°E–90°E and 10°N–25°N. The center year of the 21-year period is plotted on the horizontal axis. The 5% significance threshold for a 21-year correlation is shown as a thin horizontal black line.

most dynamical models give better SST predictions than statistical models beyond two months.

[17] A fundamental assumption in the above argument is that SSTs influence monsoon rainfall. To quantify this assumption, we re-calculated CCA for ISMR and JJAS SST for 1880–1959 and obtained 0.66 for the canonical correlation. This result implies that about $0.66^2 \approx 44\%$ of the variance of ISMR could be predicted if the concurrent SSTs were known. Similar correlations were found using monthly mean SSTs instead of JJAS mean. However, when this calculation is done for 1960–2005, cross validation experiments indicate that the best model has correlation of 0.39. This apparent weakening of the simultaneous ENSO-monsoon relation has been noted previously and suggested to be linked to the midlatitude continental warming trend [*Krishna Kumar et al.*, 1999]. For reference, we note that repeating all calculations with detrended SSTs yields virtually identical correlations, as expected since ISMR has negligible trend.

[18] If the ENSO-monsoon relation has weakened dramatically, then it would be difficult to understand how dynamical models could make skillful predictions of monsoon rainfall. An alternative hypothesis that avoids this problem is the following: the ENSO-monsoon relation has not in fact weakened, but only appears to have weakened due to sampling variability. *Gershunov et al.* [2001] show that the observed change in ENSO-monsoon relation could be explained easily by sampling variability. To further illustrate the plausibility of this hypothesis, we show in Figure 6 a running 21-year correlation between ISMR and JJAS NINO3 index for various cases. In the case of observations (thick black curve), we see that the correlations have weakened and become insignificant after 1985, consistent with the CCA results mentioned above, and consistent with *Rajeevan et al.* [2012]. If, however, the correlations are computed with the ensemble mean ISMR and ensemble mean NINO3 from each dynamical model (solid lines), the correlations remain significant throughout the period– no weakening occurs. (Results from ECMWF are not shown because its correlations are insignificant throughout most of the period.) The correlations also remain significant (with no breakdown) throughout the period even if individual ensemble members are concatenated rather than averaged. Finally, if the correlations are computed from individual ensemble members, substantial variability is found. Figure 6 shows examples (dots connected by lines) that were specifically chosen to illustrate the same kind of variability as seen

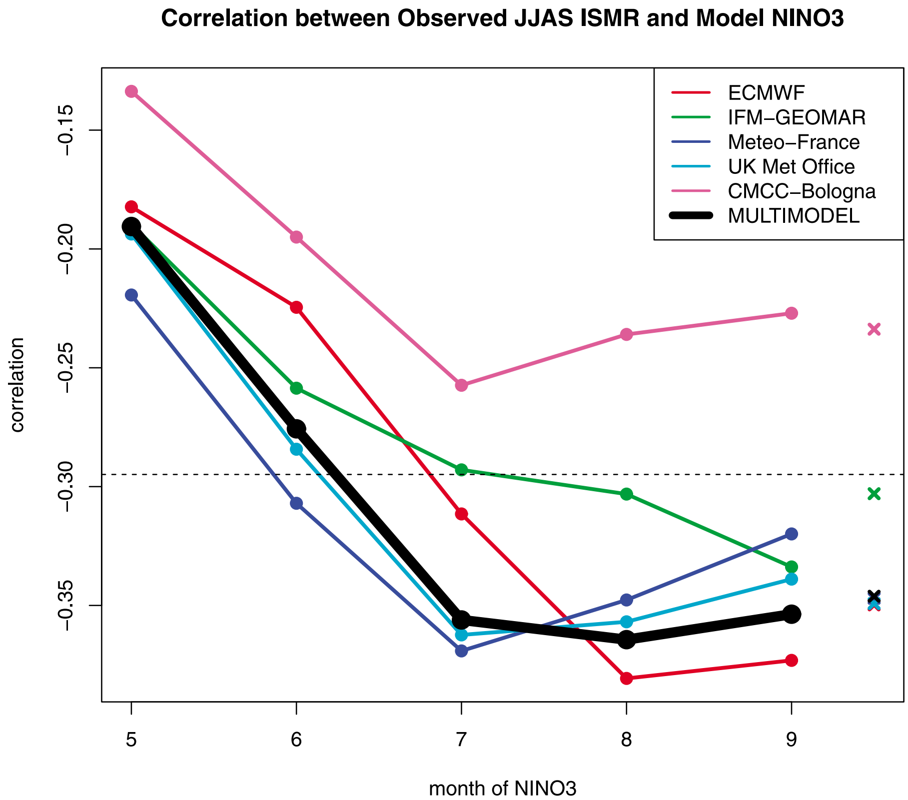

Figure 7. Correlation between observed JJAS all-India rainfall, and ensemble mean model predicted NINO3 index, for hindcasts from the ENSEMBLES data set during 1960–2005. The 'x'-symbols on the far right give the correlations for the JJAS NINO3 index in each model. The correlation with the mean of all models is shown as "multimodel."

in observations. The fact that at least one out of nine ensemble members could be found from each model to have similar variability (including the ECMWF model, not shown), and that the variability is much larger than seen in observations, and yet the ensemble means yield significant and relatively steady correlations, suggests that sampling variability can explain both the apparent weakening of the ENSO-monsoon relation and the significant skill of dynamical predictions of ISMR.

[19] The above results are quite encouraging for the prospects of predicting all-India monsoon rainfall, but unfortunately requires forecasts from coupled atmosphere-ocean models. However, since the predictability of ISMR comes mainly from SSTs, the question arises as to whether the *predicted* NINO3 index from coupled models can be used to derive skillful forecasts of ISMR. To test this possibility, we show in Figure 7 the correlation between observed ISMR and *dynamically-predicted* NINO3. The figure shows that indeed all but one model has significant correlation for July through September. The one model with insignificant correlation, CMCC, also has the smallest correlation skill of NINO3 (see Figure 5). Nevertheless, since the best statistical predictions of ISMR based on May SST have correlations around 0.15–0.21, these results suggest that routine dynamical model forecasts of NINO3 may serve as better predictors of ISMR than observed May NINO3.

5. Summary

[20] This paper used statistical optimization techniques to show that no significant May SST pattern for predicting ISMR can be statistically justified during the past four decades. Despite the lack of significant relation between ISMR and May SST, the current generation of coupled atmosphere-ocean models can predict ISMR with skill when initialized in May. The superior skill of dynamical models is attributed to the fact that there exists a significant relation between ISMR and June-September SST, and to the fact that dynamical models can predict these SSTs more skillfully than statistical models. This reasoning is supported by the demonstration that skillful predictions of ISMR can be derived from *dynamically predicted SSTs*, a fact that opens up the possibility of constructing real-time skillful predictions of ISMR from routinely available dynamical model predictions of SST.

[21] Although May SST was found to be a weak predictor of ISMR, at least during 1960–2005, the *joint* January and May SST patterns were found to provide skillful predictors of ISMR during this same period. Moreover, models trained in the recent period had skill in the independent 1880–1959 period, and vice versa. Similar skill was found using February instead of January, and using April instead of May. These robust results suggest that SST predictors based on time-lagged information may provide more skillful predictions of ISMR than predictors based on a single month or season.

[22] Consistent with previous studies, we find that the simultaneous SST-monsoon relation has weakened in the last couple of decades. However, we show that similar changes in the SST-monsoon relation can be found in individual ensemble members of coupled atmosphere-ocean models, even though the SST-monsoon relation in ensemble mean

quantities varies relatively little. This result suggests that the apparent breakdown of SST-monsoon correlation could be due, in large part, to sampling variability, consistent with the conclusions of some previous studies [*Gershunov et al.*, 2001].

[23] **Acknowledgments.** We appreciate constructive comments from the reviewers. Support is gratefully acknowledged from grants from the NSF (ATM0830068, ATM0830062), the National Oceanic and Atmospheric Administration (NA09OAR4310058), and the National Aeronautics and Space Administration (NNX09AN50G). The editor thanks the two anonymous reviewers.

[24] The Editor thanks the two anonymous reviewers for their assistance in evaluating this paper.

References

Charney, J. G., and J. Shukla (1981), Predictability of monsoons, in *Monsoon Dynamics*, edited by J. Lighthill and R. P. Pearce, pp. 99–109, Cambridge Univ. Press, Cambridge, U. K.

DelSole, T., and J. Shukla (2009), Artificial skill due to predictor screening, *J. Clim.*, *22*, 331–345.

DelSole, T., and X. Yang (2011), Field significance of regression patterns, *J. Clim.*, *24*, 5094–5107.

Drbohlav, H.-K. L., and V. Krishnamurthy (2010), Spatial structure, forecast errors, and predictability of the South Asian monsoon in CFS monthly retrospective forecasts, *J. Clim.*, *23*, 4750–4769.

Gadgil, S., M. Rajeevan, and R. Nanjundiah (2005), Monsoon prediction—Why yet another failure?, *Curr. Sci.*, *88*, 1389–1400.

Gershunov, A., N. Schneider, and T. Barnett (2001), Low-frequency modulation of the ENSO-Indian monsoon rainfall relationship: Signal or noise?, *J. Clim.*, *14*, 2486–2492.

Krishna Kumar, K., B. Rajagopalan, and M. A. Cane (1999), On the weakening relationship between the Indian monsoon and ENSO, *Science*, *284*, 2156–2159.

Rajeevan, M., C. K. Unnikrishnan, and B. Preethi (2012), Evaluation of the ENSEMBLES multi-model seasonal forecasts of Indian summer monsoon variability, *Clim. Dyn.*, doi:10.1007/s00382-011-1061-x, in press.

Shukla, J., and D. A. Paolino (1983), The Southern Oscillation and long-range forecasting of the summer monsoon rainfall over India, *Mon. Weather Rev.*, *111*, 1830–1837.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore (2008), Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006), *J. Clim.*, *21*, 2283–2296.

Walker, G. T. (1910), Correlation in seasonal variations of weather, *Mem. Indian Meteorol. Dep.*, *21*, 22–45.

Weisheimer, A., F. J. Doblas-Reyes, T. N. Palmer, A. Alessandri, A. Arribas, M. Déqué, N. Keenlyside, M. MacVean, A. Navarra, and P. Rogel (2009), ENSEMBLES: A new multi-model ensemble for seasonal-to-annual prediction—Skill and progress beyond DEMETER forecasting tropical Pacific SSTs, *Geophys. Res. Lett.*, *36*, L21711, doi:10.1029/2009GL040896.