Supplementary material: Sensitivity analyses

Methods

In order to better understand the behavior of the prediction methods, a number of variants of the setup described in the main text were tested for which prediction models were fitted and skill scores computed. These variants were:

(i) Instead of the last 21 years being the test period, the last 11 or 31 years were used, with the previous years in each case being the training period. It was expected that a shorter test period and longer training period would result in more skill but greater variability because the skill scores are averaged over fewer test years. A longer test period was expected to result in more stable estimated skill scores, which may be lower because fewer years are available for training.

(ii) Instead of retaining enough SST modes (about 40) to explain 90% of the variability, the fraction retained f was set to 0.6 or 0.3, resulting in only about 10 and 2-3 modes retained respectively. These variants tested whether the less-pronounced modes contributed meaningfully to the prediction skill.

(iii) SSTs were detrended by either removing the global warming signal or the ENSO signal to see how these components affect prediction skill. Detrending was accomplished by least-squares linear regression of the signal series on each of the SST grid points and subtracting the fit to obtain detrended SST. A global warming time series was constructed by isotone regression [1] on the global-mean SST, to which locally weighted linear regression [2] with 10-year bandwidth was applied to obtain a smooth monotone increasing series. ENSO was represented by the Southern Oscillation Index (SOI) time series, obtained from the Australian Bureau of Meteorology (`http://www.bom.gov.au/climate/current/soi2.shtml`).

(iv) The SST fields (and the years input to RF) at zero lead time were randomly permuted across years to understand the behavior of the prediction methods where the predictors have no useful signal. 19 random permutations were carried out, so that their upper range provides a 95% bound [3] for the skill scores obtained when the predictors have no association with monsoon precipitation.

Results

Decreasing the test period from 21 to 11 years resulted in greater rRMSE for the linear regression method, but lower rRMSE at long lead times for the nonlinear regression method, associated with greater variability in rRMSE between adjacent lead times (Figure S1). Increasing the test period from 21 to 31 years resulted in only slightly lower rRMSE for the nonlinear regression method, but zero rRMSE for most lags for the linear regression method, as, with the shorter training period, cross-validation found a zero coefficient matrix to give the best performance (Figure S1).

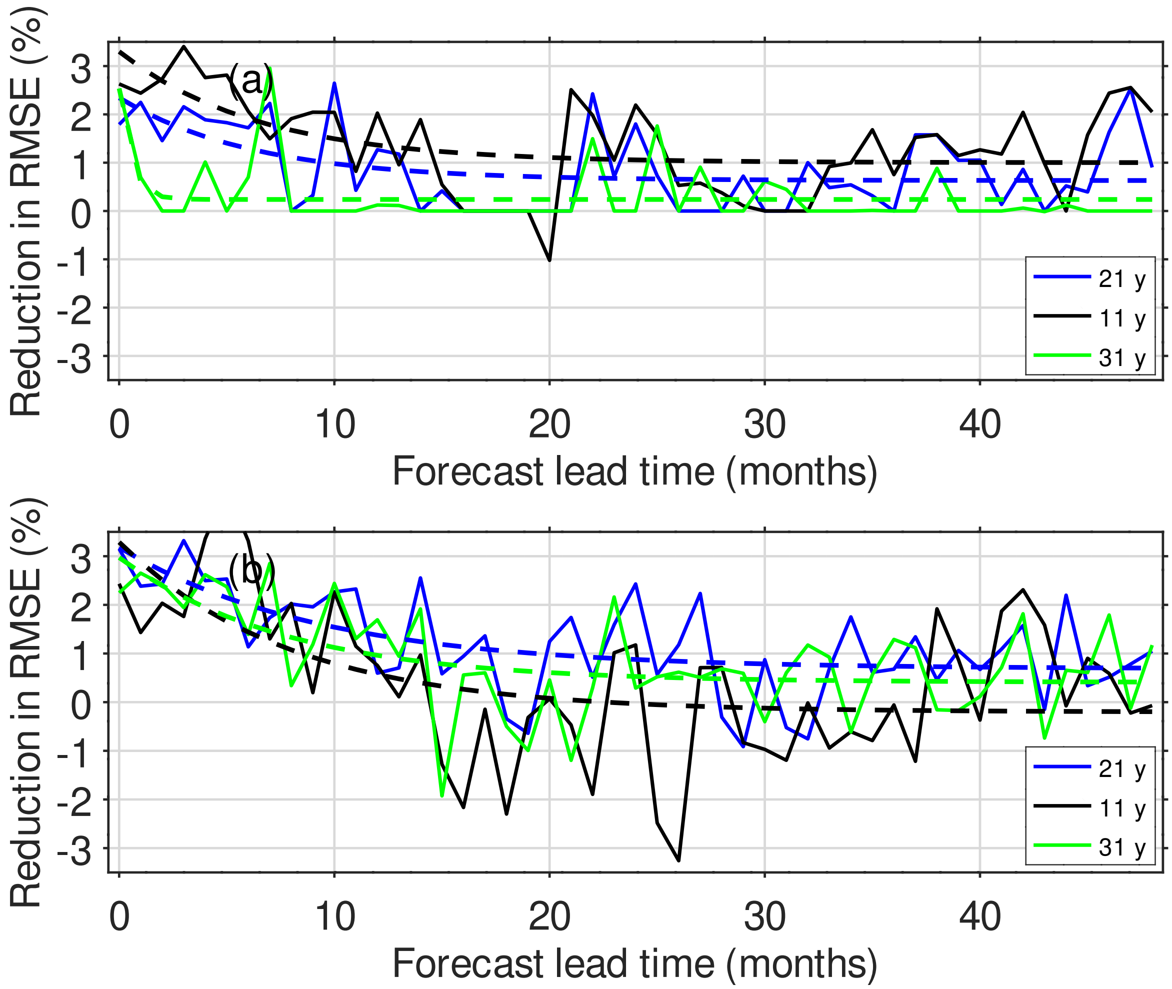

Figure S1 . Prediction skill for South Asia monsoon-season precipitation, quantified as reduction in root mean square error (rRMSE) relative to a forecast based on climatology, for different configurations of test and training years (test years 1997-2017, 2007-2017, or 1987-2017, with the previous years, starting in 1901, used for training). The thick dashed lines are exponential fits to the RMSE reduction as a function of lead time. (a) Linear regression with lasso. (b) Nonlinear random forest regression.

Decreasing the fraction f of SST variability retained from 0.9 to 0.6 resulted in a moderate increase in rRMSE for the linear regression method and a moderate decrease for the nonlinear regression method. Decreasing f further to 0.3 resulted in even higher rRMSE for the linear regression method, particularly at long lags, but consistently negative rRMSE for the nonlinear method (Figure S2).

Removing the global warming signal removes skill at long lag times (past about 18 months), suggesting that this skill is only due to extrapolating the impacts of long-term climate trends on the monsoon (rather than to, for example, decadal variability). At short lead times, nonlinear regression has only somewhat reduced rRMSE compared to

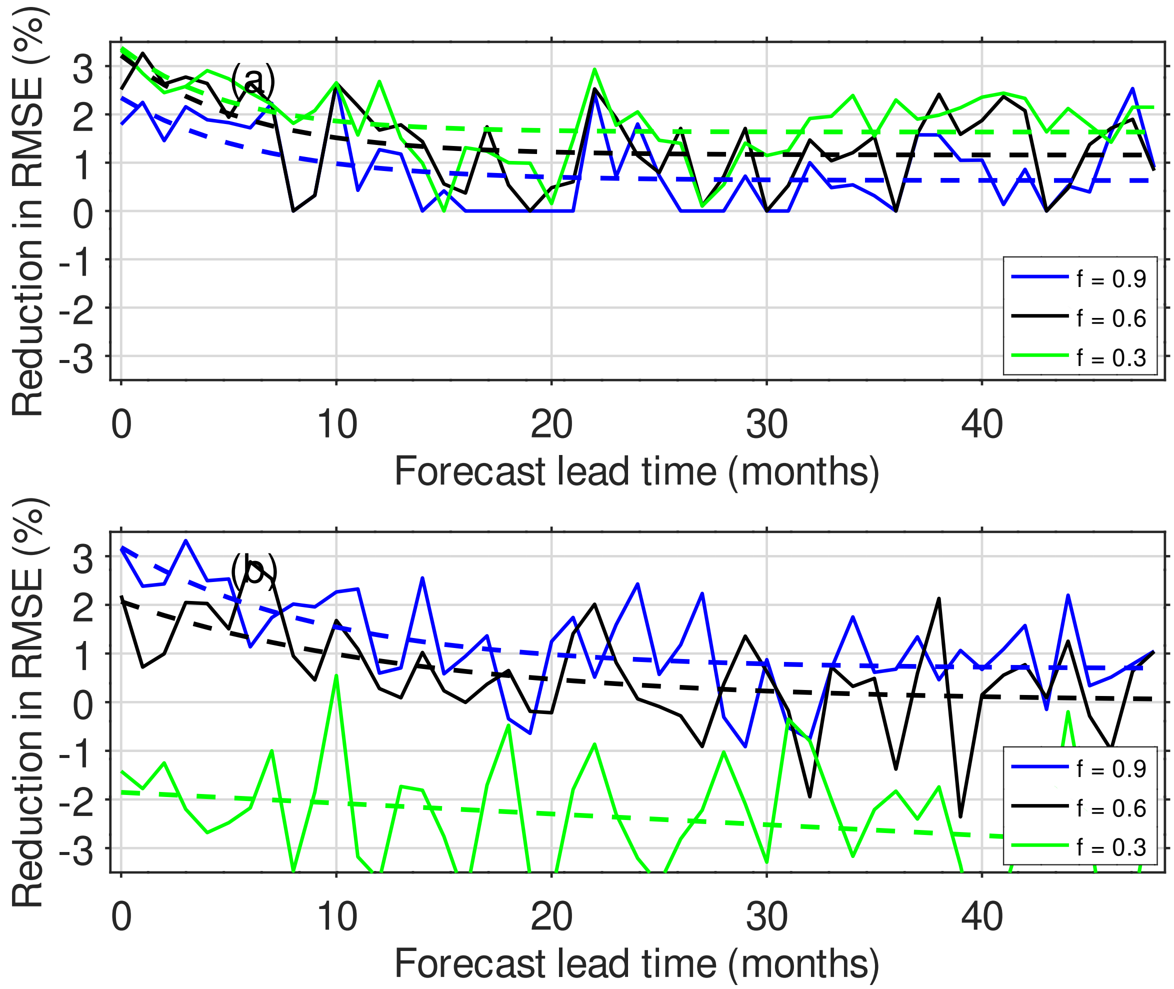

Figure S2 . Prediction skill for South Asia monsoon-season precipitation, quantified as reduction in root mean square error (rRMSE), as a function of the fraction of SST variability in the modes retained as possible predictors (0.9, 0.6, or 0.3). The thick dashed lines are exponential fits to the RMSE reduction as a function of lead time. (a) Linear regression with lasso. (b) Nonlinear random forest regression.

using the actual SST, but linear regression still has zero or negative rRMSE (Figure S3). Removing the ENSO signal leads to moderately lower rRMSE for both forecast methods at lead times up to around 12 months, and no impact on rRMSE at longer lead times (Figure S3), suggesting that ENSO only contributes part of the predictive skill of SST-based SASM precipitation forecasts.

Random permutation of the SST field's years leads to rRMSE that are indistinguishable from zero on average for both methods. The linear method usually sets the coefficient matrix to all zeros, giving exactly zero rRMSE, whereas the nonlinear method has more scatter in rRMSE (Figure S4).

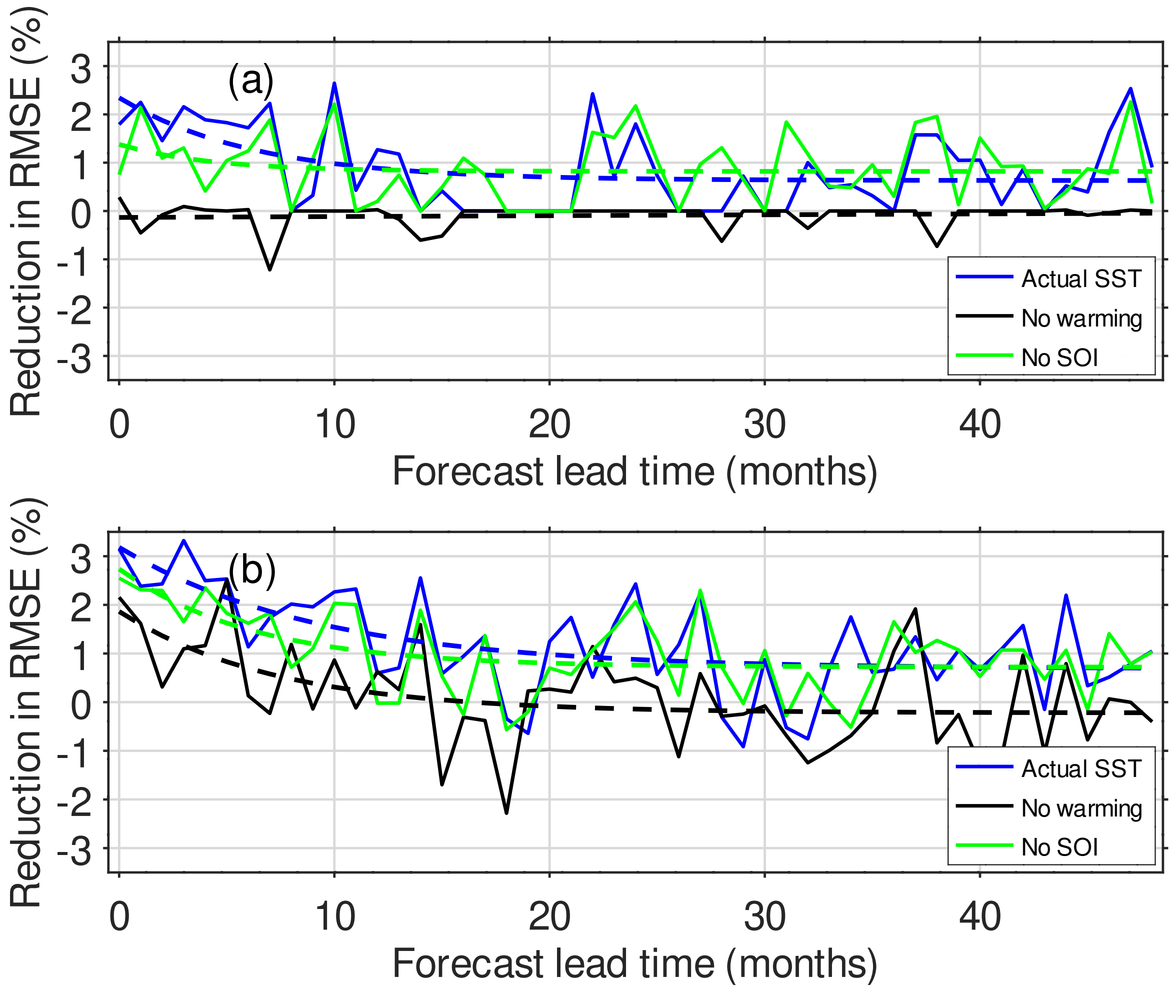

Figure S3 . Prediction skill for South Asia monsoon-season precipitation, quantified as reduction in root mean square error (rRMSE), either using actual SSTs as predictors or detrending the SST first by removing a global warming trend or the component correlated with the Southern Oscillation Index. The thick dashed lines are exponential fits to the RMSE reduction as a function of lead time. (a) Linear regression with lasso. (b) Nonlinear random forest regression.

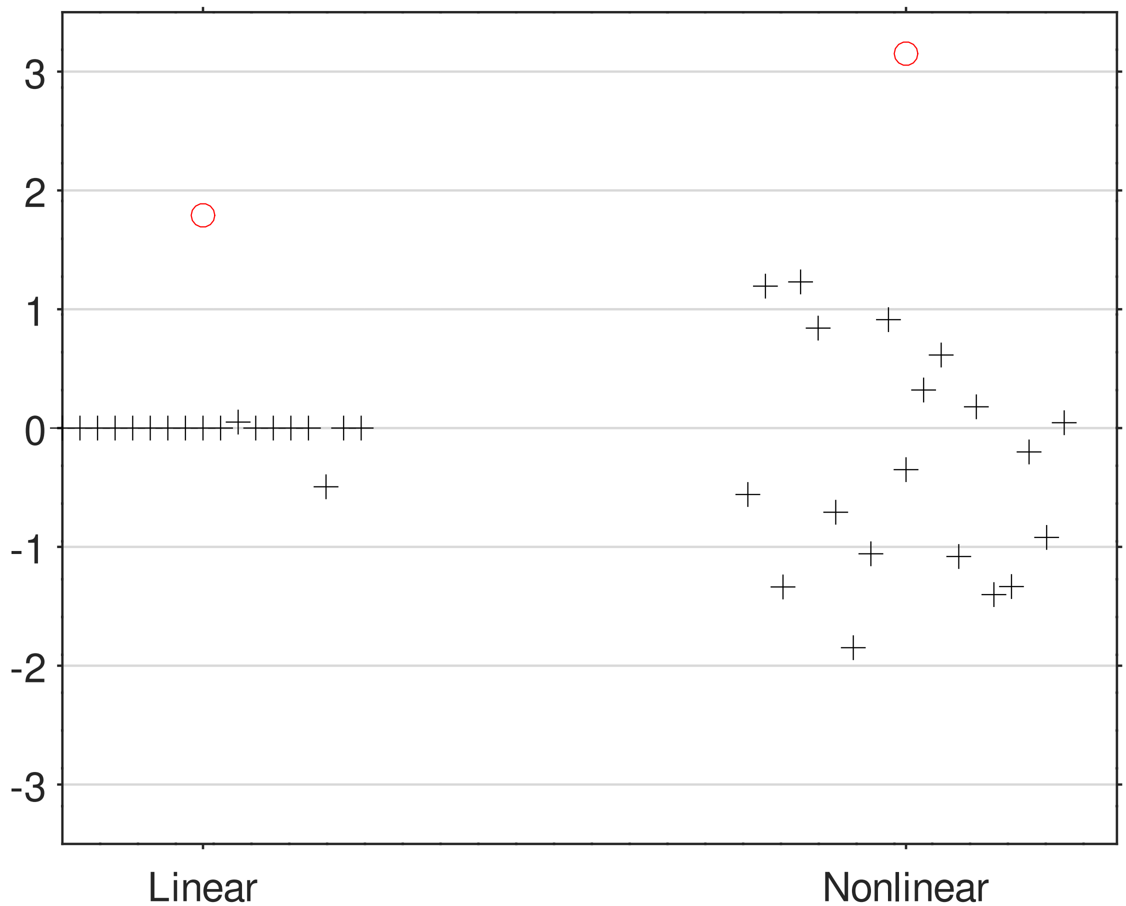

Figure S4 . Prediction skill for South Asia monsoon-season precipitation using actual SSTs as predictors (at 0 month lead time; red circles) or permuting the SSTs across years first (19 different random permutations; black crosses). Results for linear regression with lasso and nonlinear random forest regression are both shown.

References

[1] Richard E Barlow, Daniel J Bartholomew, James M Bremner, and H D Brunk. *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*. Wiley, 1972.

[2] William S. Cleveland and Susan J. Devlin. Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403):596–610, 1988.

[3] Thomas Schreiber and Andreas Schmitz. Surrogate time series. *Physica D*, 142(3-4):346–382, 2000.