An Upper Limit to Seasonal Rainfall Predictability?

SETH WESTRA AND ASHISH SHARMA

School of Civil and Environmental Engineering, University of New South Wales, Sydney, New South Wales, Australia

(Manuscript received 20 April 2009, in final form 6 January 2010)

ABSTRACT

The asymptotic predictability of global land surface precipitation is estimated empirically at the seasonal time scale with lead times from 0 to 12 months. Predictability is defined as the unbiased estimate of predictive skill using a given model structure assuming that all relevant predictors are included, thus representing an upper bound to the predictive skill for seasonal forecasting applications. To estimate predictability, a simple linear regression model is formulated based on the assumption that land surface precipitation variability can be divided into a component forced by low-frequency variability in the global sea surface temperature anomaly (SSTA) field and that can theoretically be predicted one or more seasons into the future, and a ''weather noise'' component that originates from nonlinear dynamical instabilities in the atmosphere and is not predictable beyond ~10 days.

Asymptotic predictability of global precipitation was found to be 14.7% of total precipitation variance using 1900–2007 data, with only minor increases in predictability using shorter and presumably less error-prone records. This estimate was derived based on concurrent SSTA–precipitation relationships and therefore constitutes the maximum skill achievable assuming perfect forecasts of the evolution of the SSTA field. Imparting lags on the SSTA–precipitation relationship, the 3-, 6-, 9-, and 12-month predictability of global precipitation was estimated to be 7.3%, 5.4%, 4.2%, and 3.7%, respectively, demonstrating the comparative gains that can be achieved by developing improved SSTA forecasts compared to developing improved SSTA–precipitation relationships. Finally, the actual average cross-validated predictive skill was found to be 2.1% of the total precipitation variance using the full 1900–2007 dataset and was dominated by the El Niño–Southern Oscillation (ENSO) phenomenon. This indicates that there is still significant potential for increases in predictive skill through improved parameter estimates, the use of longer and/or more reliable datasets, and the use of larger spatial fields to substitute for limited temporal records.

1. Introduction

Seasonal climate forecasting has been an active area of research since Sir Gilbert Walker first discovered a relationship between large-scale atmospheric variability in the tropics and rainfall in many parts of the world (e.g., Walker 1923). Since then, there have been tremendous developments in conceptual understanding of the climate system (e.g., Lorenz 1963), availability of large climate datasets obtained from in situ and remotely sensed sources, and computational resources that enable the analysis of large multivariate datasets or the simulation of the dynamical equations that drive the various elements of the climate system. Despite all these developments, improvements in the predictive skill for precipitation, probably the most important climate variable from a human impact perspective, has been frustratingly slow with the most sophisticated dynamical models often still unable to outperform linear regression relationships between regional precipitation and one or several indices that describe relevant modes of variability such as the El Niño–Southern Oscillation (ENSO) phenomenon (e.g., Anderson et al. 1999; Quan et al. 2006; Rajeevan et al. 2007; Van den Dool 2007; Wilks 2008).

It therefore seems appropriate to ask the question: to what extent is the global precipitation field predictable? It is well known that individual weather patterns are not predictable beyond a period of about 10 days because of the nonlinear internal dynamics of the atmosphere that effectively limits the predictability beyond individual synoptic systems (e.g., Palmer and Anderson 1994; Goddard et al. 2001; Van den Dool 2007). Although in certain cases the atmospheric general circulation is predictable beyond

Corresponding author address: Dr. Seth P. Westra, School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW, 2052 Australia.
E-mail: s.westra@unsw.edu.au

DOI: 10.1175/2010JCLI3212.1

individual weather systems, at the seasonal time scale the majority of predictability is derived from lower-boundary forcing, which evolves on a much lower time scale than these weather systems (Horel and Wallace 1981; Palmer and Anderson 1994). These boundary conditions do not allow specification of the exact timing of transitions between weather regimes that result largely out of internal atmospheric variability. However, they can influence the probability of their occurrence (Palmer and Anderson 1994), thereby allowing specification of the probability of below- or above-average precipitation for longer lead times as long as the relevant boundary forcing can be predicted. This forms the basis for developing precipitation forecasts for lead times up to a year (Goddard et al. 2001) or even longer (Ruiz et al. 2005).

The first step in seasonal prediction therefore is to identify a set of external boundary forcing variables relevant to a particular precipitation field and attempt to describe the future evolution of these variables. A range of variables relevant to seasonal forecasting have been proposed, including sea surface temperature (SST), soil moisture, vegetation, and snow and sea ice cover, although these are not equally important. In particular, SST anomalies have long been regarded as the principal forcing variable of atmospheric circulation (Barnston et al. 2005; Quan et al. 2006) and have been shown to influence the probabilities of below- and above-average precipitation in many parts of the world (e.g., Nicholls 1989; Ward and Folland 1991; Barnston 1994; Drosdowsky and Chambers 2001; Hoerling and Kumar 2003; Rajeevan et al. 2007).

Complete knowledge of the future evolution of all relevant external boundary conditions does not, however, imply a perfect precipitation forecast. In particular, a conceptual breakdown of atmospheric circulation into a predictable component driven largely by low-frequency variations in external boundary conditions, and a "weather noise" component that is unpredictable at the seasonal time scale, implies that there exists some upper limit to the seasonal predictability that cannot be improved upon even with a perfect mathematical representation of the global climate system and with perfect forecasts of the future evolution of all the relevant boundary conditions. To this end, Barnston et al. (2005) described an approach to quantifying the upper limit to atmospheric predictability by generating ensembles of different atmospheric general circulation models (AGCMs) forced to historical boundary conditions but with different initial conditions to isolate the relative influences of (potentially predictable) boundary forcing and (largely unpredictable) internal atmospheric dynamics on response variables such as global precipitation. Although in many ways this represents a conceptually attractive approach able to capture the full nonlinear dynamical relationship between all the relevant external boundary variables and precipitation, the difficulty in accurately representing the fine temporal- and spatial-scale precipitation processes often leads to significant biases (e.g., Tippett et al. 2005), which may remain even after developing ensembles of multiple AGCMs (Barnston et al. 2005).

In this paper we propose a simple alternative empirical approach based on linear regression methods to directly estimate the upper limit of predictive skill for global precipitation. The upper limit of predictive skill, which we will henceforth refer to as predictability, is defined within the context of our regression modeling approach as the unbiased estimate of predictive skill using a given model structure assuming all relevant predictors are included. This is a statistically tractable equivalent to that proposed by Madden (1989; see also Kalnay 2003) who defined potential predictability beyond the limit of deterministic weather predictability to be "the total variance of the anomalies averaged over a month or season, minus the variance that can be attributed to weather noise." We make the assumption that the pool of relevant predictors is contained within the global SSTA field, which can be justified by the importance of this field as the dominant lower-boundary forcing for precipitation (Barnston et al. 2005; Quan et al. 2006). A second assumption—that relationship between the SSTA field and precipitation can be represented by a linear statistical model—is more difficult to justify on theoretical grounds (e.g., see Hoerling et al. 1997) and probably will require the type of dynamically based analysis suggested by Barnston et al. (2005) to confirm. As we will show, however, the linear assumption does have some grounding, and even if it proves to be theoretically unjustified, it may well provide a practical limit to seasonal prediction because of the high dimensionality of the climate datasets and comparatively short observational records (see discussion in Van den Dool 2007).

The remainder of this paper is structured as follows. In the next section we introduce the datasets to be used in this paper and conduct a preliminary statistical analysis of global precipitation using a representation of the ENSO phenomenon to highlight aspects of statistical forecasting relevant to the discussion that follows. In section 3, we describe our proposed methodology and include a synthetic example to illustrate the approach. Our estimates of precipitation predictability are then described in section 4, using both concurrent and lagged relationships between the SSTA field and global precipitation. Section 5 contains a summary of our results and a description of associated implications on seasonal forecasting as well as some potential avenues for developing improved seasonal rainfall predictions.

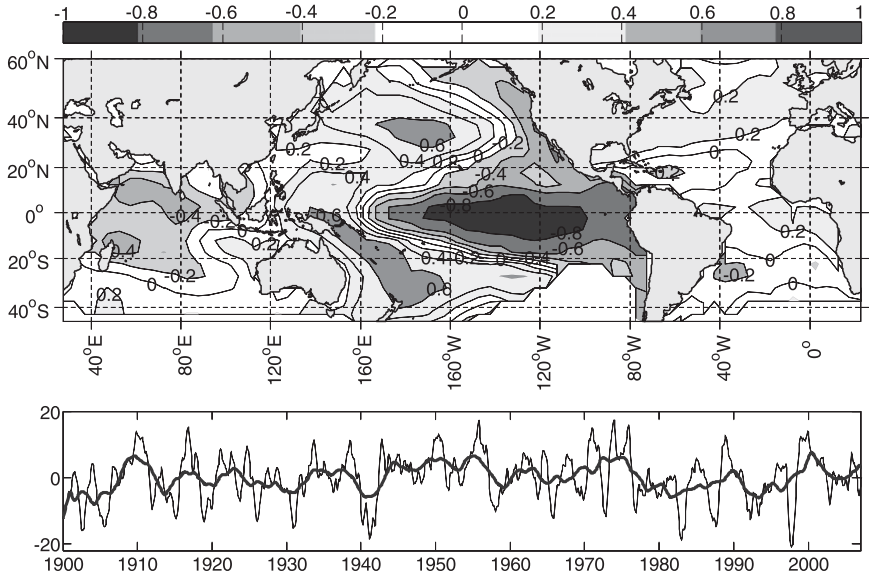

FIG. 1. Principal component 1 calculated over all months of record from 1900 to 2007 (1296 months), representing 18.8% of the global SSTA dataset variance. (top) Correlation coefficients between PC1 and SSTAs at individual grid points; (bottom) time series from 1900 to 2007 (thin black line) and 4-yr moving average (thick black line).

2. Data and preliminary analysis

a. Global SST anomalies

A global sea surface temperature anomaly (SSTA) dataset was obtained from the reconstruction of raw SST values using an optimal smoother, as described in Kaplan et al. (1998; available online at http://iridl.ldeo.columbia.edu/SOURCES/.KAPLAN/). The data are available on a 5° longitude by 5° latitude grid across the global ocean, totaling 1207 grid points. In the temporal dimension, the data comprise monthly data that we converted to seasonal data by calculating overlapping three-month averages [i.e., December–February (DJF), January–March (JFM), February–April (FMA), etc.]. We use data from 1900 to 2007, such that we have a record of 1296 overlapping seasons.

To facilitate linear regression modeling, the data were converted to a subset of orthogonal components using principal component analysis (PCA; see Preisendorfer 1988 or Wilks 2006 for details). Unless indicated otherwise in the paper, the global temperature trend was removed by

1) deriving a global temperature trend series by calculating the weighted average temperature of the 1207 grid points separately for each month of record from 1900 to 2007, with the weighting based on the relative area of each grid point (as a 5° longitude by 5° latitude grid box has a larger surface area at lower latitudes than at higher latitudes);
2) subtracting this series from the SSTA dataset separately at each grid point.

Thus, the sea surface temperature data used here represent both anomaly (in the sense that the climatological mean is removed from each grid point separately) and detrended (in the sense that the global average trend is removed from each grid point) data. This ensured that the principal components (PCs) (time series) and eigenvectors (''loading vectors'' or empirical orthogonal functions representing spatial patterns) were both mutually orthogonal. We note that the sensitivity of our results to this data preprocessing approach was tested at length and found to be minimal, and the implications of keeping the global temperature trend in the data are discussed in section 4f.

The first two principal component time series (lower panels) and maps representing the correlation coefficients between the principal components and the original gridded data (upper panels) are presented in Figs. 1 and 2. Correlation coefficients, rather than eigenvectors, are presented because of the ease of interpretation and the direct relationship between correlation coefficients and the variance accounted for by each component at each grid point, and it should be noted that visual inspection of the eigenvector maps show qualitatively similar features. The physical interpretability of the PCs of the global SSTA field is generally well understood and reported in numerous other publications (e.g., Richman 1986; Van den Dool 2007; Westra et al. 2009); however, the first two

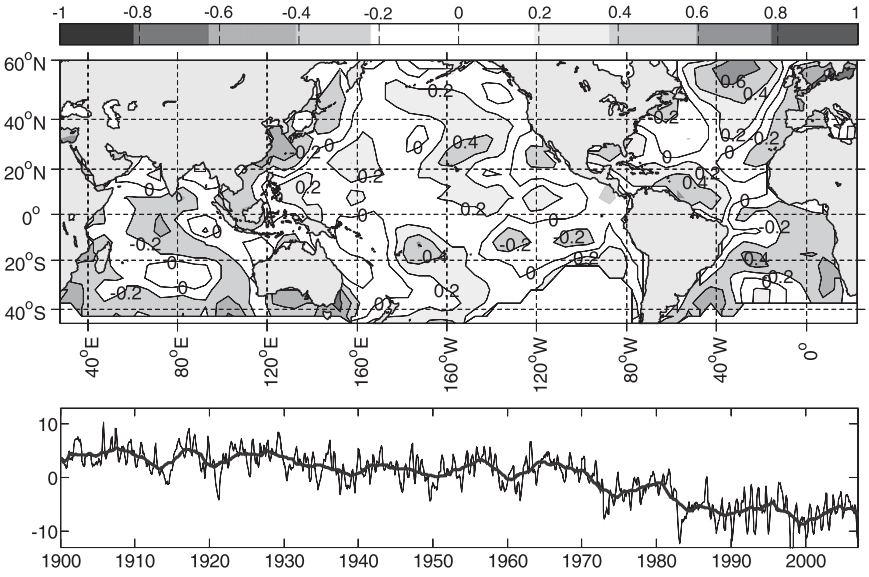

FIG. 2. Principal component 2 calculated over all months of record from 1900 to 2007 (1296 months), representing 7.8% of the global SSTA dataset variance. (top) Correlation coefficients between PC2 and SSTAs at individual grid points; (bottom) time series from 1900 to 2007 (thin black line) and 4-yr moving average (thick black line).

PCs contain interesting attributes relevant to the remaining analysis, so we therefore provide a brief review here.

As can be seen from Fig. 1, the first principal component is representative of the ENSO phenomenon and is in fact correlated with a well-known index of the ENSO phenomenon, the Niño-3.4 time series (defined as the seasonally averaged SSTA over the central Pacific Ocean defined by the region 5°N–5°S and 170°–120°W; see Trenberth 1997) with a correlation coefficient of −0.91. The negative sign is due to the high negative weightings in the central and eastern equatorial region of the Pacific Ocean commonly associated with ENSO. This component accounts for 18.8% of the global SSTA variance, making it by far the single dominant global mode of SST variability.

The second principal component presented in Fig. 2 accounts for 7.8% of global SSTA variance and indicates a strong negative trend, with positive correlation coefficients in the Pacific and North Atlantic Oceans and negative coefficients in the Indian and South Atlantic Ocean. Because of the detrending step conducted prior to applying PCA, the trend represented by this PC can be viewed as the warming of the Indian and South Atlantic Oceans relative to the Pacific and North Atlantic Oceans, with this feature being broadly consistent with the analysis of Casey and Cornillon (2001) who find higher warming trends in the Indian and Southern Oceans compared to the North Atlantic and North Pacific basins for the period 1960–90.

For both PCs, a 4-yr moving average is calculated to demonstrate the strong season-to-season persistence of these representations of the SSTA field. As we will show, this persistence is the basis for developing one or more season-ahead statistical forecasts of the global precipitation field.

b. Global precipitation anomalies

The Global Historical Climate Network version 2 (GHCN) gridded monthly precipitation dataset was selected as the global precipitation field to be used for this analysis (Peterson and Vose 1997; Peterson et al. 1997; available online at http://www.ncdc.noaa.gov/pub/data/ghcn/v2) and represents the most comprehensive available dataset of monthly precipitation covering the entire twentieth century (New et al. 2001). The dataset provides an extended coverage of global precipitation from 1900 to 2007 on a 5° longitude by 5° latitude grid and is derived from 2064 homogeneity adjusted precipitation stations from the United States, Canada, and former Soviet Union, together with 20 590 raw precipitation stations throughout the world. Prior to averaging over a 5° by 5° grid, the raw precipitation data were converted to anomaly data with respect to the 1961–90 base period. The final gridded data comprised 819 individual grid points covering the majority of the global land surface area.

We converted the monthly precipitation data into seasonally averaged data using the same approach that

was adopted for the SSTA data. There were numerous grid points, particularly over arid regions, in which a significant portion of the record between 1900 and 2007 was not reported because of insufficient data to estimate the anomaly over that grid box. To ensure consistency in the analysis and ensure that only high-quality data were included, all grid points that had more than 15% of the record missing for a particular season were excluded. Furthermore, to avoid statistical difficulties associated with cases where a large number of zero values were reported at a particular location and season, grid points that contained zeros for more than 5% of the record were removed. Although this process does not ensure normality in the resulting precipitation record, the prevalence of highly skewed distributions is nonetheless reduced. In all cases the removal of grid points from the dataset was conducted on a season-by-season basis, such that if a location had more than 5% of zeros for JFM but not for June–August (JJA), we only removed that location for the analysis for JFM. The result of this filtering is that, for each season, approximately 435 spatial gridded locations were included in the response dataset, indicating a reduction of around 50% (i.e., ~435/819) of the original dataset.

The methodology discussed in the subsequent section assumes that records at each precipitation grid point are temporally uncorrelated. All analyses are conducted for individual seasons in isolation, such that we are only concerned with temporal correlation between a season in a given year and the same season of the preceding and following years. We estimated the autocorrelation coefficients for each season separately at all grid points and found that the globally averaged autocorrelation coefficient was 0.05 and is therefore unlikely to significantly impact on our estimates of predictive skill.

The spatial correlation between successive grid points was also examined. We calculated the spatial correlation in precipitation time series between all adjacent grid points and found about 86% of adjacent sequences reported correlation coefficients that were statistically significant at the 5% significance level, with a globally averaged spatial correlation coefficient of 0.44. Unlike temporal correlation, we do not assume spatial independence; rather, it is only necessary that the spatial correlation is sufficiently low to ensure a high effective dimension of the global precipitation dataset.

The regions retained for the season JFM are shown as dots in Fig. 3, and as can be seen there is generally good spatial coverage, with exceptions in northern North America, Central America and Amazonia, parts of Africa, central Asia, and Siberia. These largely comprise locations that are either arid such that the record comprises a large number of months with no precipitation or, alternatively, do not have complete precipitation records over the period 1900 to 2007. For this reason we suggest that the exclusion of precipitation at these grid points is likely to provide more robust results in the sections that follow.

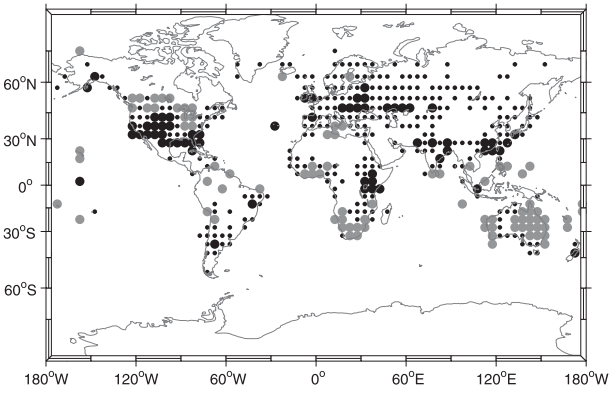

FIG. 3. Location of 5° longitude by 5° latitude gridded precipitation data points used in the analysis (all dots). Only grid points having greater than 85% of months of data between 1900 and 2007 were included. The large gray (black) dots represent locations with significant positive (negative) correlation with PC1 (ENSO) at the 5% level for the season JFM, with remaining small black dots representing locations that do not have statistically significant correlation with PC1.

c. *Preliminary analysis*

As a preliminary analysis, we evaluated the correlation between precipitation time series at each location and the first principal component, which corresponds to the ENSO phenomenon. The large gray (black) dots in Fig. 3 comprise those grid points for which positive (negative) statistically significant correlation with PC1 is observed at the 5% significance level, which is equivalent to a correlation coefficient of approximately ± 0.19. In total, 30% of the global precipitation grid locations exhibited statistically significant correlation. At this significance level, one would expect on average 5% of locations (or approximately 22 grid points) to report statistically significant correlation by random chance; however, because of the observed spatial correlation a more rigorous field significance test is required to ensure that spurious correlations are not reported (Wilks 2006).

We conducted a field significance test using a bootstrapping with replacement procedure in which data were randomly drawn from the PC1 time series to construct a new series PC1* of the same length as PC1. This was repeated to generate 10 000 samples of PC1* having the same length and distributional characteristics as the original PC1 but with each data point now occurring randomly in time (see Efron and Tibshirani 1993 for more details on the bootstrap). Each PC1* was regressed against the global

precipitation data to obtain 10 000 bootstrapped correlation coefficients, and this was used to estimate the value of the correlation coefficient with given probability of being exceeded. In this case we found that the 5% field significance level occurred when 6.7% of stations reported statistically significant correlation, which as expected was slightly higher than the estimate using a Binomial distribution assuming spatial independence (Wilks 2006). The field significance of the precipitation–ENSO relationship, with 30% of the grid points exhibiting statistically significant correlation coefficients, therefore cannot be attributed to random chance.

In addition to the strong field significance score, the individual regions exhibiting statistically significant correlations is also consistent with what is presently understood about ENSO–precipitation relationships. Specifically, regions with statistically significant correlation comprise North, Central, and South America, southern Africa, Southeast Asia, and Australasia. The results are consistent with other findings that anomalously dry conditions are typically found during warm ENSO periods in the tropical regions bordering the eastern Indian Ocean such as Australia, Indonesia, and South Asia (large gray dots in Fig. 3), and anomalously wet conditions over much of North America (large black dots in Fig. 3; see also Dai and Wigley 2000; Diaz et al. 2001). Interestingly, we examined other skill scores, notably the Spearman rank correlation (Wilks 2006) and the linear error in probability space (LEPS; Potts et al. 1996) and found almost identical results, both in terms of the statistical significance of the scores (with significance levels estimated by bootstrapping PC1 with replacement 10 000 times and estimating the 5-percentile skill score that would be expected by random chance) and in terms of the geographic distribution of the statistically significant scores. This lends qualified support to the assumption that the relationship between PC1 and global precipitation can be represented linearly, as significant nonlinearity in the predictor–response relationship would result in rank correlations being noticeably higher than the mean squared error skill score (MSESS).

Results such as these form the basis of the assertion that ENSO is the dominant source of predictability for global precipitation (e.g., New et al. 2001). This is an assertion that we will now examine more closely.

3. Methodology

As discussed in the introduction, the purpose of this analysis is not to estimate the predictive skill by considering the influence of individual predictors, such as those attributable to the ENSO phenomenon described above, but to develop an unbiased estimate of the global precipitation predictability. As we will show, provided that certain necessary simplifying assumptions are made, such an estimate is possible using very simple linear regression techniques.

a. Overview of approach

The objective of this analysis is to develop an estimate of the upper limit of predictability of global precipitation at the seasonal time scale. To achieve this we conceptually divide the seasonal precipitation variability into two components: the first associated with variability attributable to external boundary conditions and the second associated with internal variability in the atmosphere. This second quantity is generally considered to be unpredictable beyond a period of approximately 10 days (Van den Dool 2007), such that at the seasonal time scale it can be considered to be random weather noise (Barnston et al. 2005).

We propose that, at any location or grid point, the rainfall time series can be partitioned into these two components, with the relationship represented by the equation

$$\mathbf{y} = \beta_0 + f(\mathbf{X}) + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathbf{y}$ represents the time series of precipitation of length n at any location or grid point, β_0 represents the sample mean precipitation at that location and can be viewed as an estimate for the location climatology (i.e., the long-term mean), and $\boldsymbol{\varepsilon}$ represents the random weather noise component. Beyond climatology, the predictable component of $\mathbf{y}$ is contained in the term $f(\mathbf{X})$, where $\mathbf{X}$ represents an $n \times p$ matrix, p is the dimension of the predictor matrix and is assumed to contain all the relevant information concerning the variability in the external boundary conditions, and f represents some function relating $\mathbf{X}$ to $\mathbf{y}$, such that the expected value of $f(\mathbf{X})$ equals zero.

In addition to the assumption that precipitation at any location can be separated into a potentially predictable component related to fluctuations in external boundary conditions and a nonpredictable component related to internal atmospheric variability, we make two further assumptions. These assumptions are 1) that the function f is linear and 2) that the matrix representing the external boundary conditions can be represented as a zero-mean orthogonal representation of the SSTA dataset. Thus, Eq. (1) can be simplified to yield

$$\mathbf{y} = \beta_0 + \sum_{i=1}^{p} \beta_i \mathbf{x}_i + \boldsymbol{\varepsilon}, \tag{2}$$

where β_i represents the regression coefficients, and $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_p]$ represents the PCA-transformed time series representation of the SSTA dataset. The PCA transform

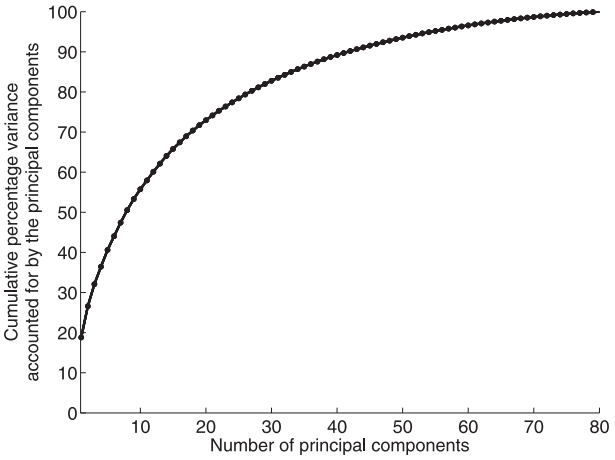

FIG. 4. Cumulative percentage variance accounted for by each principal component of the global SSTA dataset. Principal components were calculated over the full monthly dataset from 1900 to 2007, totaling 1296 data points.

ensures that **X** is orthogonal and that successive components maximize variance of the data in a least squares sense. The cumulative variance accounted for by the principal components is shown in Fig. 4 for up to 80 components, at which exactly 100% of the variance of the global SSTA field is accounted for. The upper limit of 80 components is due to the approach adopted by Kaplan et al. (1998) in deriving the reconstructed SSTA dataset, as this was based on an optimal interpolation in the subspace of the first 80 empirical orthogonal functions. As can be seen, PCA efficiently represents the variance of the data, with more than a quarter of the variance of the full 1207-dimensional global SSTA dataset accounted for by just the first two components, approximately half of the variance accounted for by the first eight components, and 90% accounted for by the first 42 components.

To estimate the maximum predictability of the precipitation data, we use the MSESS defined as (Wilks 2006)

$$\text{MSESS} = 1 - \frac{\text{MSE}_{\text{pred}}}{\text{MSE}_{\text{clim}}}, \quad \text{where} \tag{3a}$$

$$\text{MSE}_{\text{pred}} = \frac{1}{(n-1-p)}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad \text{and} \tag{3b}$$

$$\text{MSE}_{\text{clim}} = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \overline{y}_i)^2, \tag{3c}$$

where n represents sample length indexed by i, $\hat{y}_i$ represents the least squares estimate for **y** in Eq. (2), and $\overline{y}_i$ represents the sample mean of **y** and is used to represent climatology. The MSESS can be interpreted as the fraction of the observed variance accounted for by the forecasts using a particular model compared to a reference climatology model, with a score of 1 representing a perfect model [i.e., the error term ε in Eq. (2) is reduced to zero] and a score of 0 representing no improvement over climatology. Furthermore, a negative score indicates an inferior result compared to the reference climatology "forecast," with the variance of residuals in Eq. (3b) being greater than the variance of residuals from the climatology model in Eq. (3c). Dividing the model and climatology residual squared error terms in Eqs. (3b) and (3c) by $(n - 1 - p)$ and $(n - 1)$ ensures that the estimates are unbiased, which means that adding a random predictor to the model will not change the expected value of MSE_{pred}. In the remainder of this paper we express the MSESS as a percentage, by multiplying the MSESS in Eq. (3a) by 100.

We have now presented all the theory we intend to use for estimating the predictability for global precipitation. The model in Eq. (2) is fitted to estimate **y** from **X**, where **X** is of dimension p and constitutes a dimension-reduced version of the full 80-dimensional orthogonal representation of the global SSTA dataset. This allows the MSESS to be estimated separately for each **y** using Eq. (3). To estimate global predictability, this process is repeated separately for each precipitation grid point, and the globally averaged MSESS is calculated.

At first glance, the empirical approach described here appears overly simplistic, in particular when compared against the approach proposed by Barnston et al. (2005), which involves using a suite of GCMs that represent the dynamical equations relating slow-moving boundary conditions to global precipitation. Much of the simplicity in the approach stems from necessity. In particular, even though the SSTA dataset may only represent one of the boundary conditions that are likely to drive long-term precipitation variability, with other drivers including soil moisture, vegetation, snow cover, sea ice extent, and so on, to our knowledge it is the only such variable for which a long-term global dataset is available. Furthermore, as we will discuss in more detail in section 4a, we intend to apply the model in Eq. (2) such that the majority of variance of the global SSTA is accounted for, which makes the use of a nonlinear model for the application proposed in this paper difficult. This is because adopting a nonlinear model will significantly increase the effective dimension of the predictor pool, which for the sample sizes considered here (i.e., slightly over 100 years of data) would prevent the development of a statistical model that explains a sufficiently large portion of the variance of the SSTA dataset to estimate predictability.

Despite the simplicity, we believe each of the assumptions is justified to an approximation, and we will

return to this during our discussion in section 5. To illustrate the proposed approach, we have developed the following simple synthetic example.

b. Synthetic example

The objective of this research is to estimate the maximum prediction performance achievable from a particular model structure. The focus on *predictability* is therefore distinct from the actual predictive skill available from a finite sample and represents the unbiased estimate of predictive skill assuming all relevant predictors are included in the model.

We construct a simple synthetic model with p predictors, $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_p]$, by generating independently p samples of $\mathbf{x}$ each of length n by sampling from a normal (Gaussian) distribution with mean zero and unit variance [$\sim N(0, 1)$]. We assume that the predictor pool $\mathbf{X}$ comprises the full set of predictors available to estimate $\mathbf{y}$ and is therefore analogous to the orthogonal SSTA dataset in which each principal component plausibly influences the variance of any given precipitation grid point. We then generate the response variable $\mathbf{y}$ as a linear function of the first two of these predictors:

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \boldsymbol{\varepsilon}, \tag{4}$$

where $\beta_0 = 0$, $\beta_1 = 0.2$, $\beta_2 = 0.05$, and $\boldsymbol{\varepsilon} \sim N(0, 1)$. The value of β_0 has been set to zero to ensure that the expectation of $\mathbf{y}$ is also zero, and the values of β_1 and β_2 have been set to small values such that the variance of $\mathbf{y}$ is dominated by the variance of the error term $\boldsymbol{\varepsilon}$. Furthermore, the remaining predictors $\mathbf{x}_3, \mathbf{x}_4, \ldots, \mathbf{x}_p$ have not been included in the generative model and are therefore spurious predictors.

The formulation of the model in Eq. (4) is analogous to Eq. (2) for the case where there are only two predictors and all the model coefficients are known. The objective of statistical seasonal forecasting in this linear framework is to develop an empirical relationship between $\mathbf{X}$ and $\mathbf{y}$ by identifying the predictors that are relevant for the model—in this case $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$, together with the model parameters $\boldsymbol{\beta}$. Finally, the error term $\boldsymbol{\varepsilon}$ is analogous to the weather noise component described in the introduction and is by definition unpredictable. It should be emphasized that we are using a linear model to keep the analysis as simple as possible, and we temporarily ignore the issue of lagged predictor–response relationships as would be necessary in a true forecast setting.

As in this case we know the structure and parameters of the model in Eq. (4), we can estimate the theoretical predictive skill using the MSESS. Given that each of the predictors $\mathbf{x}$ and the error term $\boldsymbol{\varepsilon}$ were generated from a normal distribution with zero mean and unit variance, we can estimate the theoretical MSESS provided in Eq. (3) as

$$\text{MSESS} = 1 - \frac{\sigma_\varepsilon^2}{\beta_1^2 + \beta_2^2 + \sigma_\varepsilon^2}, \tag{5}$$

where $\sigma_\varepsilon^2 = 1$ by construction. Thus, for $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$, the MSESS $= 1 - 1^2/(1^2 + 0.2^2 + 0.05^2) = 4.08\%$. Furthermore, in the case where we fit the model in Eq. (4) using only a single predictor $\mathbf{X} = [\mathbf{x}_1]$, the residual term becomes $\boldsymbol{\varepsilon}' = \beta_2 \mathbf{x}_2 + \boldsymbol{\varepsilon}$ and the variance associated with this term becomes $\sigma_{\varepsilon'}^2 = \beta_2^2 + \sigma_\varepsilon^2$. Thus, the theoretical MSESS for the case where we only use $\mathbf{x}_1$ is $1 - (1^2 + 0.05^2)/(1^2 + 0.2^2 + 0.05^2) = 3.84\%$, slightly lower than when both predictors are used. Addition of the remaining possible predictors $[\mathbf{x}_3, \ldots, \mathbf{x}_p]$ will not affect the theoretical predictive skill as by construction the response $\mathbf{y}$ is generated using only the first two predictors.

We now wish to estimate the true skill score empirically from the data. To prevent reporting artificial skill, the usual approach is to cross validate or apply some other complexity penalty criterion to estimate forecast skill. We illustrate the implications of using a range of complexity penalty approaches, including leave-one-out cross validation and parametric approaches such as the adjusted R^2, Mallow's Cp statistic, and the Bayesian information criterion (BIC) (see Hastie et al. 2001 for details). Each of these complexity penalty criteria affect the MSESS via the MSE_{pred}, by providing an additional penalty to the MSE to account for increasing p.

To see how this affects the estimate of the MSESS, we add predictors sequentially such that for the one-parameter case the predictor vector becomes $\mathbf{X} = [\mathbf{x}_1]$, and the p-parameter case the predictor matrix becomes $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p]$. A total of 100 000 samples of $\mathbf{X}$, $\boldsymbol{\varepsilon}$, and $\mathbf{y}$ were then randomly generated for different values of p up to $p = 20$, and for sample lengths $n = 100$ and $n = 1000$. For each sample, the $\boldsymbol{\beta}$ parameters in Eq. (4) were estimated using linear regression, and the MSESS calculated using the bias-adjusted version of the MSE_{pred} in Eq. (3b), together with each of the complexity penalty versions described above. Thus we have 100 000 values of the MSESS for each n and p. The results of this analysis are presented in Fig. 5, with the MSESS presented in this figure representing the mean value over the 100 000 samples.

The upper line in Fig. 5 represents the theoretical MSESS as calculated by Eq. (5) above, with a value of 3.84% for $p = 1$, and 4.08% for the remaining p [for $p \geq 2$ both predictors in Eq. (4) are included in $\mathbf{X}$]. As can be seen, each complexity penalty criterion results in an underestimation of the true MSESS, with the magnitude of the underestimation being a function of both n and p. Interestingly, for $n = 100$, the maximum MSESS for

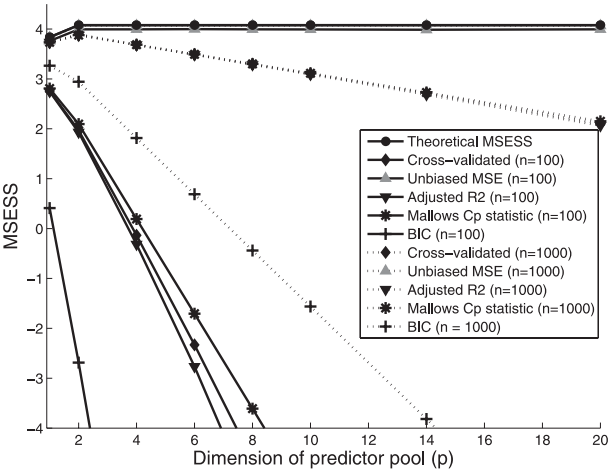

FIG. 5. Implications of different values of sample length n and predictor dimension p on the MSESS. The theoretical (true) MSESS (upper black line, circles), the MSESS calculated using the unbiased estimator MSEpred (located just below upper black line; gray triangles), and four alternative complexity penalty approaches (cross validation, adjusted R^2, Mallows Cp statistic and the BIC) are shown.

each of the complexity penalty criterion occurs for the one predictor model, such that $\mathbf{x}_2$ generally would be excluded from the forecast model. This is important, as complexity penalty approaches such as cross validation are commonly used in predictor selection for statistical seasonal forecasting of precipitation (e.g., Wilks 2006). As such, the application of such approaches for the estimation of asymptotic predictive skill would most likely result in the underestimation of the number of climate predictors (e.g., principal components of the SSTA field) that influence precipitation variability.

For $n = 1000$, the maximum MSESS occurs for the two predictor model for each of the complexity penalty criteria except for the BIC, which applies the most severe penalty. It is therefore clear that an approach for estimating forecast skill based on some complexity penalty criteria, including a cross-validation approach in which part of the sample is withheld, cannot be used to estimate the theoretical MSESS, with the magnitude of the underestimation particularly notable for relatively short, high-dimensional datasets such as is the case for most statistical seasonal forecasting applications.

Finally, we turn to the unbiased estimators for $\mathrm{MSE}_{\mathrm{pred}}$ and $\mathrm{MSE}_{\mathrm{clim}}$ and compute the MSESS as defined in Eq. (3), represented as gray triangles in Fig. 5 for both $n = 100$ (solid line) and $n = 1000$ (dotted line). These lines are difficult to visualize in the figure, as they lie just below the theoretical MSESS, thereby indicating that they provide a good estimate of the true MSESS. In particular, the use of the unbiased $\mathrm{MSE}_{\mathrm{pred}}$ shows an increase in the MSESS up to $p = 2$, followed by a constant MSESS for higher values of p. This is desirable for the subsequent analysis using the SSTA field to forecast seasonal precipitation, as in reality we do not know which predictors should be included, and therefore we want to know what the asymptotic predictive skill might be if we maximize the pool of plausible predictors.

It needs to be emphasized that although the MSESS provides a good approximation, it is not exact, since even though it is possible to provide an unbiased representation of $\mathrm{MSE}_{\mathrm{pred}}$ and $\mathrm{MSE}_{\mathrm{clim}}$, the MSESS calculated via Eq. (3a) is slightly biased. This is because assuming we are drawing independent observations from a normal distribution, Cochran's theory shows that s^2 (the sample variance) follows a chi-squared distribution, which approaches a normal distribution only when n is large (Stuart and Ord 1987). This causes an underestimation of the true MSESS, with the magnitude of the underestimation increasing as n decreases. Importantly, the bias is only a function of n and the magnitude of the theoretical MSESS (since if $\mathrm{MSE}_{\mathrm{pred}} = \mathrm{MSE}_{\mathrm{clim}}$, the distributional form of either is of no concern and MSESS $= 1$) and not a function of the dimensionality p of the predictor dataset.

We estimated the bias as a function of both n and the value of the true MSESS using only a single predictor $\mathbf{x}$ and present this in Fig. 6. To allow for different values of the true MSESS we have modified the model formulation given in Eq. (4) to $\mathbf{y} = \sqrt{m}\mathbf{x} + \sqrt{(1-m)}\boldsymbol{\varepsilon}$, where $\mathbf{x}$ and $\boldsymbol{\varepsilon}$ are $N(0, 1)$ as before, as this will ensure that the theoretical (true) MSESS becomes equal to m. This is because, by substituting into Eq. (5), one has

$$\mathrm{MSESS} = 1 - \frac{[\sqrt{(1-m)}]^2}{[\sqrt{(m)}]^2 + [\sqrt{(1-m)}]^2} = m. \quad (6)$$

The results are given as a percentage of the true MSESS and show underestimation in all cases, with the most significant bias occurring for small sample sizes and low values of the true MSESS. Returning to our synthetic example, given the true MSESS for $p = 2$ predictors is 4.08% and $n = 100$, the expected value of the MSESS is 3.92%, which constitutes a downward bias of approximately 2%. This bias can be observed in Fig. 5 by comparing the theoretical MSESS (black circles) with the MSESS calculated using the unbiased $\mathrm{MSE}_{\mathrm{pred}}$ (gray triangles) for sample size $n = 100$. For the application described in this paper we consider this bias to provide a negligible influence on the final predictability estimates.

4. Results

In the previous section we described a method for developing an estimate of the predictive skill (measured using the MSESS) that would be achievable assuming

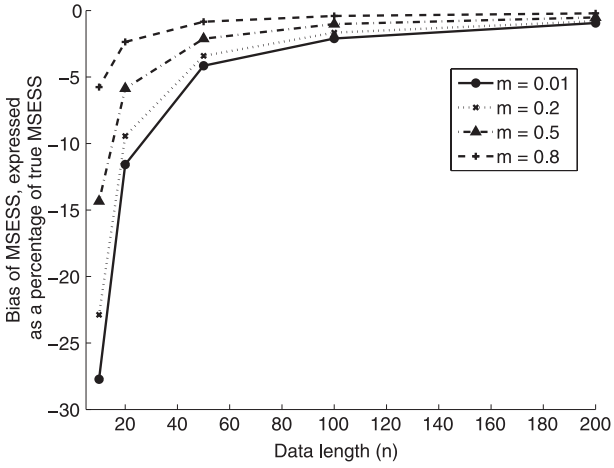

FIG. 6. Estimating the bias of the MSESS, using a simple model given by $\mathbf{y} = \sqrt{m}\mathbf{x} + \sqrt{(1-m)}\boldsymbol{\varepsilon}$ where $\mathbf{x}$ and $\boldsymbol{\varepsilon}$ are both generated random sequences of length n with 0 mean and a unit variance. The model is equivalent to the model represented by Eq. (4) for a single predictor, except that the coefficient m now can be interpreted as the fraction of the variance of $\mathbf{y}$ accounted for in $\mathbf{x}$ (with the remaining fraction accounted for by the error term $\boldsymbol{\varepsilon}$ necessarily being $1 - m$).

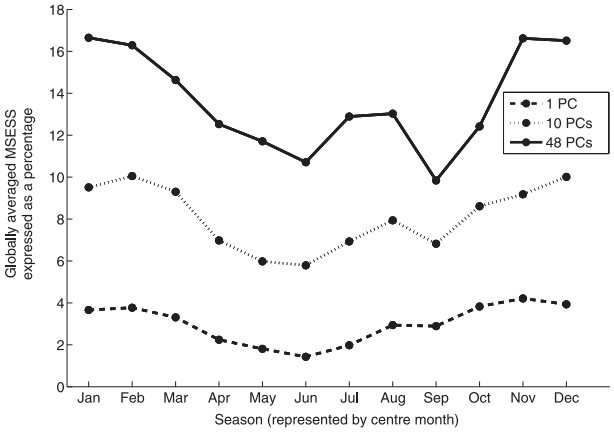

FIG. 7. Globally averaged MSESS by season for 1, 10, and 48 PCs. Each season is represented by the center month, such that "Jan" represents the season DJF.

a linear predictor–response relationship and that all possible predictors are included in **X**. In particular we pointed out that although the variance associated with the MSESS calculated by regressing a high-dimensional predictor matrix **X** against a single precipitation grid point **y** is likely to be high, the MSESS is approximately unbiased and therefore on average would yield the true value of predictive skill. Calculation of the global average MSESS after regressing **X** separately against a large number of response variables (i.e., separately for each precipitation grid point and each season) reduces the variance of this estimate, so that the best estimate of asymptotic MSESS can be achieved by considering the complete global precipitation record. Unfortunately, the focus on the globally averaged skill score in the ensuing analysis obscures interesting regional details; however, as we shall see, the consideration of global precipitation allows for some robust conclusions that would not be possible by taking a more regional approach.

a. Estimating the global asymptotic MSESS

The asymptotic MSESS can be estimated by fitting the model described in Eq. (2) separately for each precipitation grid point **y** using a common p-dimensional predictor pool **X** comprising the principal components of the SSTA field. The globally averaged MSESS is derived by computing the weighted average MSESS calculated at each grid point via

$$\text{MSESS}_{\text{global}} = \sum_{j=1}^{q} w_i \text{MSESS}_j \Big/ \sum_{j=1}^{q} w_i, \qquad (7)$$

where w_i represents a weight that accounts for decreasing grid size with increasing latitude calculated as cos(latitude). Furthermore, j indexes the grid point and q represents the total number of grid points, which as described in section 2b is approximately equal to 435 depending on the season. Similar to our synthetic example, we construct the p-dimensional predictor pool **X** comprising the principal components of the SSTA field by adding components sequentially in order of the variance accounted for by the PCs. Thus, for $p = 1$, **X** only contains the first PC and so on.

The globally averaged MSESS for $p = 1$, 10, and 48 is presented in Fig. 7, with the predictor pool accounting for 18.8%, 55.7%, and 92.8% of the variance of the global SSTA dataset, respectively. The MSESS was calculated separately for each season (consisting of 3 months referenced by the center month) using the full record from 1900 to 2007, such that n ranged from 90 to 108 (with the lower limit of 90 based on the filter that only seasons with more than 85% of the full record available are permissible). The maximum value of p was therefore selected to be 48, which is approximately half of n while representing the vast majority of the variance of the global SSTA dataset.

These results show that, despite the high level of emphasis of ENSO in the climate literature, this mode represented by the first PC only appears to result in a global annually averaged MSESS of 3.0%. Increasing the dimension of **X** results in large increases in the MSESS, up to an annually averaged MSESS of 13.6% for $p = 48$. Furthermore, there is clear seasonality in the MSESS, with maximum predictability occurring in the northern winter (approximately spanning November through February) and lowest predictability in the summer. Interestingly, this seasonality appears remarkably consistent for different dimensions of **X**.

These results also show that predictability increases with the dimension of **X**. We confirm the statistical significance of these results by comparing these results with skill scores obtained by bootstrapping **X** with replacement to generate multiple realizations of **X*** and substituting this into Eq. (2). As expected from the synthetic example presented earlier, this analysis confirms that the global MSESS using the first 48 bootstrapped PCs is on average zero with a standard deviation of about 1.5%, such that skill scores reported in Fig. 7 represent genuine improvements in predictive skill as a consequence of introducing higher-order PCs into the model.

To develop a better understanding of how predictability increases as a function of variance accounted for in the global SSTA dataset, we compute the annually averaged MSESS (calculated by averaging over all seasons) for values of p ranging from 1 to 48. This is shown as the black line in Fig. 8 (left axis) and suggests a monotonic increase in the MSESS with the number of PCs included in **X**. This monotonic increase differs from the results of the synthetic example in which the MSESS plateaus after including the first two predictors, highlighting that even very high-order PCs contribute to observed global precipitation variability. The likelihood that a monotonic increase for all 48 PCs would occur by random chance is exceptionally low because of the mutual orthogonality of each of the PCs.

An interesting result is that the largest increase in the MSESS is associated with the first PC, followed by the second PC, and so on. This implies that there might be a direct link between the increase in globally averaged asymptotic MSESS and the variance of the global SSTA dataset contained in **X**. To test this assertion, we also plot the variance accounted for by successive PCs as a solid line in Fig. 8 (right axis). We adjusted the axes such that the variance accounted for by the first 48 PCs lines up with the asymptotic MSESS at this dimension.

The close alignment between the two curves in Fig. 8 is striking and suggests that the improvement in the MSESS, which is interpreted as the percentage reduction in variance resulting from the fitted model relative to a baseline climatology model, is directly proportional to the fraction variance accounted for by each individual PC relative to the global SSTA field. Considering that the asymptotic global MSESS is 13.6% when using 48 PCs, which together account for 92.8% of SSTA variance, a small extrapolation to the point where 100% of the SSTA variance is accounted for brings the asymptotic MSESS accounted for by the full global SSTA dataset to 14.7% (i.e., 13.6%/0.928). This forms the basis for our estimate for the global predictability of seasonal precipitation.

How robust is this relationship? A simple test is to use an alternative orthogonal representation of the global

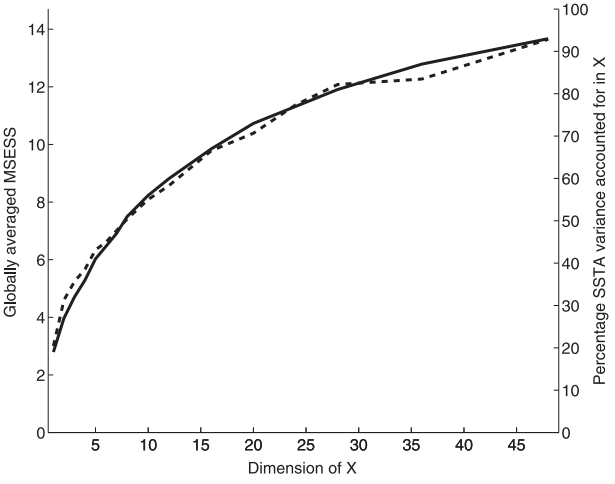

FIG. 8. Globally averaged MSESS (dotted line, left axis) calculated as the average over all seasons and cumulative percentage variance accounted for by successive PCs included in **X** (solid line, right axis).

SSTA dataset and ascertain whether the relationship between SSTA variance accounted for in the individual components and the globally averaged MSESS holds. We use independent component analysis (ICA) for this purpose (Lee 1998; Hyvarinen et al. 2001). The basis of ICA is that some orthogonal rotation of the principal component matrix **X** can be found that maximizes the statistical independence of the components, with this being a more stringent constraint than PCA, which only ensures that the covariance (correlation) matrix is diagonal. The most common applications of ICA are in signal processing and image processing, although ICA recently has been used to analyze the global SSTA field (Aires et al. 2000; Westra et al. 2009).

ICA was selected as this technique provides a different variance breakdown compared to PCA while still maintaining temporal orthogonality. The physical interpretability of the individual components is therefore not relevant here. We perform the ICA rotation of **X** for $p = 6$ because of instabilities associated with the ICA algorithm when the dimension becomes high. The results are presented in Fig. 9, with the variance accounted for using the PCA representation shown for reference. The scale of the vertical axes is the same as for Fig. 8 to maintain consistency.

As before the cumulative variance accounted for by ICs (solid line, triangles) increases with p; however, the variance maximizing property of PCA is now lost, with IC5 accounting for more variance of the SSTA data than any other component. The cumulative variance accounted for at $p = 6$ is the same for the ICA and PCA representations by construction, since the independent components (ICs) are simply a linear combination of the PCs.

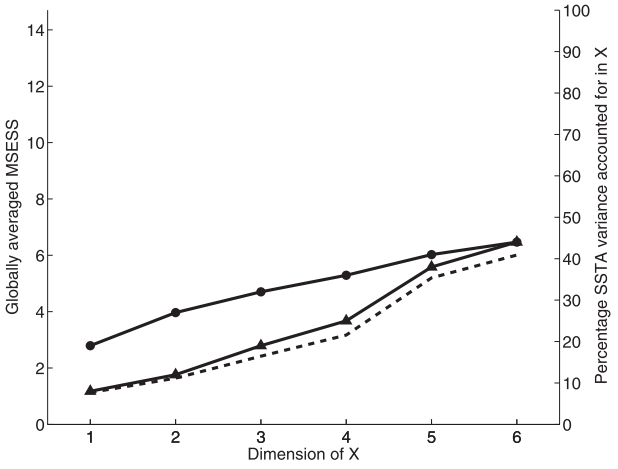

FIG. 9. Globally averaged MSESS (dotted line, left axis) calculated as the average over all seasons and cumulative percentage variance accounted for by successive ICs (solid line, triangles, right axis) and PCs (solid line, circles, right axis) included in **X**.

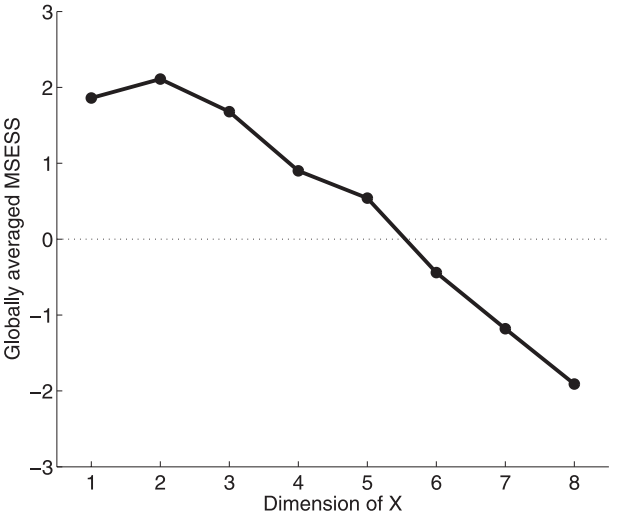

FIG. 10. Leave-5-out cross-validated MSESS plotted as a function of the dimension of **X** and calculated as the average over all seasons.

The global MSESS was calculated for different numbers of ICs and the results once again show a monotonic increase in MSESS with increasing dimension of **X**; however, the shape of the curve now clearly follows the ICA breakdown of variance among successive components. This adds weight to the assertion that the improvement in MSESS is a direct function of the variance in the total SSTA field accounted for by individual components.

b. Relation to actual predictive skill

To place the previous results in context, we estimate the cross-validated MSESS using leave-5-out cross validation to estimate the predictive skill for different dimensions of **X**. We use leave-5-out rather than leave-1-out cross validation in this case as this reduces any effects associated with the autocorrelation of **X**. We maintain a simple approach where, given a predictor matrix **X** of dimension p, we estimate the model parameters in Eq. (2) after withholding a portion of the sample, and then we apply the estimated parameters to the withheld portion to estimate $\hat{y}^{-k}$. Here, the superscript $-k$ indicates that the response $\hat{y}$ was estimated with the kth portion of the data removed, with this being repeated by sequentially withholding portions of data such that our final $\hat{y}^{-k}$ is of length n (see Hastie et al. 2001 for more details on cross validation). We note that in this case we do not use cross validation to select the optimal number of predictors at each location. Rather, we use the same approach that was used in section 4b where we gradually increase p from 1 to 48 and use this to estimate the cross-validated MSESS at all locations.

The results are provided in Fig. 10, and show a cross-validated MSESS for $p = 1$ of 1.86%, which is considerably below the unbiased estimate of 3.0% computed for $p = 1$ in section 4a and highlights the penalty that cross validation places on the skill score results. A marginal improvement (i.e., increase) can be observed for $p = 2$, with a cross-validated MSESS of 2.11%, with a deterioration (i.e., decrease) in predictive skills for higher dimensions of **X**. Although 2.11% appears to be a very low skill score, it should be remembered that it constitutes a global average and is much higher for certain seasons and regions, most notably for those regions indicated in Fig. 3 to be influenced by the ENSO phenomenon.

Although improved model formulations can be proposed that would likely result in greater predictive skill than our simple cross-validated linear regression model, this result does highlight the difficulty in engaging higher-order principal components in developing a predictive model. The reason for this is obvious; the marginal improvement in predictive skill for the addition of higher-order principal components as indicated in Fig. 8 is smaller than the penalty imposed by cross validation shown in Fig. 5. This result shows the challenge of using a 48-dimensional (or higher) predictor pool to estimate seasonal precipitation at any given location, and highlights why so much of the statistical seasonal forecasting literature concerns the derivation and/or identification of climate ''indices'' that efficiently represent the variance of the SSTA dataset most relevant to the region being analyzed (e.g., see Goddard et al. 2001). Nevertheless, it is clear that significant variability will be missed by using these reduced-dimension approaches, as all the relevant SSTA variability necessarily is only captured by engaging the full SSTA dataset.

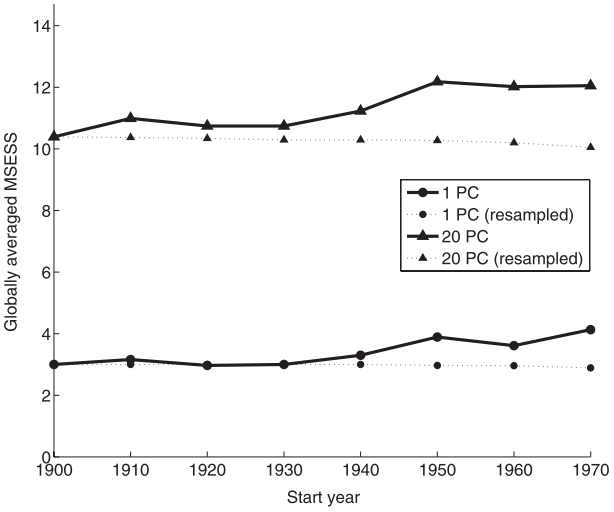

FIG. 11. Implications on globally averaged MSESS after removing earlier parts of the global SSTA and precipitation datasets from the analysis. Solid lines represent the MSESS considering only data after the start year indicated in the x axis. Dotted lines represent the MSESS after randomly removing years from the full 1900–2007 data such that the sample length corresponds to the sample length for the solid lines.

c. Implications of data length

The previous results were computed using precipitation data from 1900 to 2007. A potential limitation of using such an extended dataset is that, particularly for the earlier parts of the record, significant measurement and sampling errors may impact on the results. We therefore repeat the preceding analysis by only considering the global SSTA and precipitation data after a particular start date, which we gradually shift forward year by year from 1900 to 1970. We maintain the preprocessing step that no more than 15% of the sample may be missing at any location and season, such that the minimum value of n when the start year is 1970 is $38 \times 0.85 = 32$. The results for $p = 1$ and 20 are shown as solid lines in Fig. 11; we did not evaluate the case with $p = 48$ as p becomes greater than n.

The results show a slight gradual increase in the MSESS as sample size decreases, with maximum MSESS occurring when only post-1950 data are considered. The improvements are relatively small, however, with a 35% increase in MSESS for $p = 1$ and a 16% increase for $p = 20$. Applying the 16% increase in MSESS to the global precipitation predictability estimate of 14.7% derived in section 4a yields a revised predictability estimate of approximately 17.0%.

In section 3b we highlighted that the MSESS is a mildly biased estimator for the true MSESS, with this bias becoming significant as n becomes small. We therefore estimated the MSESS for different sample lengths by randomly withholding a fixed percentage of data from the analysis, with the percentage of data withheld being related to the available sample length corresponding to each start date. The results are presented as dashed lines in Fig. 11 and show a slight decrease in MSESS, with a maximum decrease when using the smallest sample lengths. This result is consistent with the results developed during the synthetic study, and therefore also indicates that the estimator of MSESS for the full sample size from 1900 to 2007 is likely to only slightly underestimate the true MSESS.

We have identified three possible reasons that might explain the increase in MSESS with decreasing sample length shown in Fig. 11. The first is that improved instrumentation and higher recording density in global SSTA and precipitation fields could result in a decrease in variance in the error term ε. The second and third possible reasons are that nonstationarity in the SSTA dataset due to low-frequency natural variability or anthropogenic climate change, respectively, might cause changes in the SSTA–precipitation relationship. The ultimate explanation is difficult to confirm and may comprise a combination of all three explanations. However, the results point to a relatively strong consistency in MSESS estimates for different sampling periods and do not show any evidence that anthropogenic climate change is degrading the relationship between SST anomalies and precipitation at the global scale. Furthermore, the improvement in asymptotic MSESS is less than 20% when using the shorter and presumably more reliable data, which in most cases will be outweighed by the statistical benefits of using longer sample sizes to train the parameters of a statistical forecasting model.

d. Implications of temporal and spatial scale

Thus far we have adopted a seasonal (3 month) time scale for this analysis, based on the premise that variance in the boundary conditions influence the probability of different weather outcomes, such that the maximum predictive skill is likely to occur when averaging across multiple synoptic systems (e.g., see Barnston 1994). This premise is tested here by repeating the analysis of section 4b using monthly rather than seasonal averages. We once again plotted the MSESS against the dimension of the predictor matrix p (figure not shown) and found that the MSESS increases incrementally with p in a similar way to the seasonal case provided in Fig. 8, except that now the maximum MSESS when p is 48 is only 8.9%. Given that for the monthly case the first 48 PCs account for 90.3% of the variance of the global SSTA field, we extrapolate to the case where 100% of the SSTA variance is accounted for and get an asymptotic MSESS of 9.9% (i.e., 8.9%/0.903). This represents about two-thirds of the MSESS when averaging over seasonal (3 month)

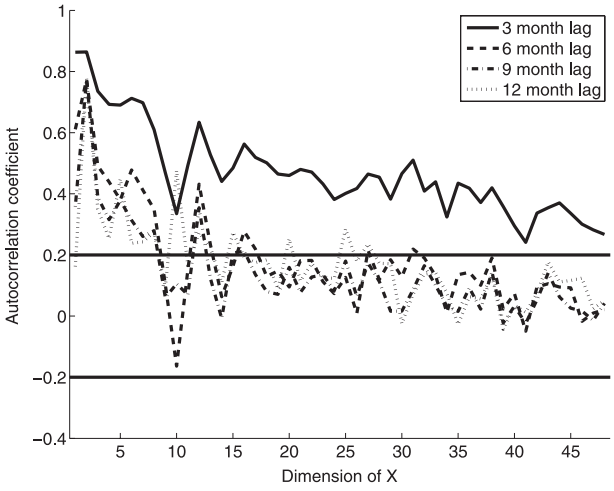

FIG. 12. Autocorrelation coefficient against principal component for lags ranging from 3 to 12 months.

blocks, suggesting that seasonal forecasting at the seasonal time scale results in an improved signal-to-noise ratio, which is probably mostly due to the reduction in the weather noise component of Eq. (2) due to the averaging out of multiple individual weather systems over the 3-month time frame.

We also examine the implications of spatial scale on the analysis. To this end, we use an alternative global gridded land surface precipitation dataset prepared by the Climate Research Unit (CRU) and described in Hulme and Osborne (1998), which has been derived using 11 880 separate time series across global land areas from 1900 to 1998 (http://www.cru.uea.ac.uk/~mikeh/datasets/global/). This dataset is available on both a 5° latitude by 5° longitude resolution and a 2.5° latitude by 3.75° longitude resolution grid. The finer-resolution data comprise 1520 separate grid boxes and therefore allow for the representation of additional detail in global precipitation variability. Although the station data for this dataset were also derived from the GHCN (version 2) database and therefore contain some overlap with the GHCN precipitation gridded dataset, these data are derived using a smaller number of total precipitation stations and are based in a Theissen polygon weighting to construct the individual gridded time series (see also discussion in New et al. 1999). We used the same filtering approach as was conducted for the GHCN data (refer to section 2); for the finer grid time scale we retained on average 1250 grid points.

We computed the MSESS for different numbers of PCs in an identical manner to the earlier analysis with the GHCN data and found the shape of the MSESS curve to be qualitatively similar to Fig. 8. The MSESS at 48 PCs was found to be 14.98% and 14.15% for the 5° by 5° and 2.5° by 3.75° gridded data, suggesting slightly improved performance for the lower-resolution data. We extrapolated to 100% variability accounted for in the SSTA field by dividing by percentage variance accounted for using the first 48 PCs and found an asymptotic MSESS of 16.1% and 15.4% for the larger and smaller grid sizes, respectively. We interpret this result using similar logic to the temporal resolution issue discussed in the previous section: climate variability at the seasonal time scale typically operates over large spatial areas, with the impact of averaging individual synoptic events over larger spatial scales being a slight increase in the signal-to-noise ratio.

e. Implications of lag

In the previous sections we estimated the asymptotic predictability of global precipitation that can be derived from the global SSTA dataset, assuming that the relationship between global SSTA and precipitation can be represented by the formulation in Eq. (2). The analysis used concurrent relationships between the SSTA and precipitation data, with the implicit assumption that precipitation variability is driven by the instantaneous variability in the boundary conditions and that any predictability over longer time horizons is derived from low-frequency variability of the boundary forcing. How valid is this assumption? And how is the SSTA–precipitation relationship expected to change if a lag was to be introduced?

We have already suggested that the importance of the global SSTA dataset in providing atmospheric boundary forcing is derived largely from the fact that the oceans contribute to approximately 85% of water vapor in the atmosphere (Bigg et al. 2003), and that the recycling rate (defined as the proportion of water that precipitates out because of local evaporation compared horizontal transport) is less than 10% and 20% at spatial scales of 500 and 1000 km, respectively (Trenberth 1998). This implies that the majority of land surface precipitation would be ultimately derived from evaporation from the ocean surface. Furthermore, it has been estimated that residence times of water in the atmosphere are relatively short, with an *e*-folding residence time of atmospheric moisture calculated to be just over 8 days (Trenberth 1998). This suggests that the best results in terms of predictability should be derived from concurrent SSTA–precipitation relationships, with any predictability when using lagged relationships due almost exclusively to the persistence structure of the boundary forcing.

To test this hypothesis, we first examine the persistence structure of the SSTA field. To this end we evaluate the level of persistence in each of the individual principal component time series by calculating the lag-k autocorrelation coefficients with k ranging from 3 to 12 months. The results are provided in Fig. 12 for up to

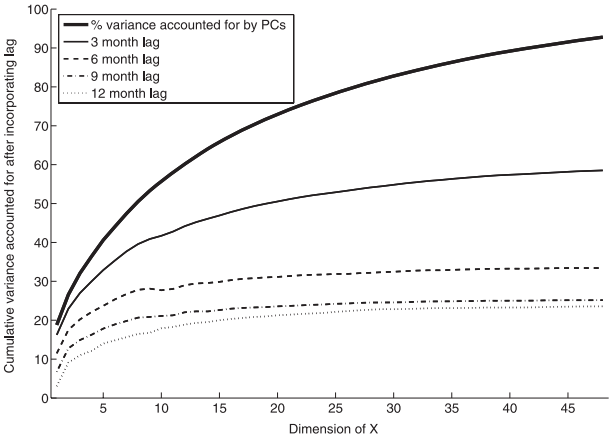

FIG. 13. Cumulative variance accounted for by individual PCs multiplied by the variance accounted for when introducing a lag.

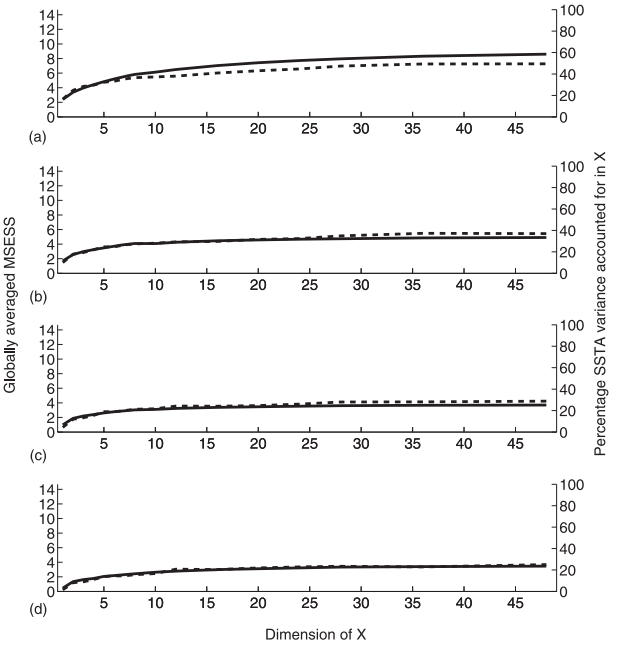

FIG. 14. Globally averaged MSESS obtained using a lagged relationship between SSTA and precipitation (dashed line, left axis), and the PCA variance accounted for in the original SSTA field multiplied by the square of the autocorrelation coefficient at the relevant lag (solid line, right axis). Lags are (a) 3, (b) 6, (c) 9, and (d) 12 months.

48 principal components, and they show that, as expected, the level of autocorrelation decreases with increased lag. The interesting result is that the level of autocorrelation generally also decreases with the order of the principal component. For example, the first principal component, which is representative of the ENSO phenomenon, shows the maximum autocorrelation at the 3 month lag, with the level of autocorrelation gradually decreasing with the order of the PC. A spike in the autocorrelation coefficient is observed for the second principal component, which we attribute to the significant trend exhibited by this component as shown in Fig. 2. This trend ensures high levels of persistence can be observed even at the 12 month lag. For the remaining PCs the level of autocorrelation appears to largely decrease with the order of the PC, with autocorrelation for 6 months or longer eventually falling below the 95% statistically significant level of 0.2. The implications of this result for seasonal forecasting are significant, since it shows that not only does PCA provide the most efficient (in a least squares sense) representation of a multivariate dataset, but it also represents the persistence structure of global SSTAs efficiently by the lower-order components.

The benefit of considering the lag-k autocorrelation coefficient is that its squared value can be interpreted as the variance of the principal component at time $t + k$ accounted for by the same principal component at time t. We can use this to calculate the cumulative variance accounted for by the full-dimensional PCA representation of the SSTA dataset at any given lag by multiplying the cumulative variance shown in Fig. 4 by the square of the autocorrelation coefficient as shown in Fig. 12. The cumulative variance curves for lags from 3 to 12 months are presented in Fig. 13, with the 0 lag (concurrent) curve reproduced from Fig. 4 for reference.

The interpretation of these curves is that they represent the cumulative variance of the SSTA field at time t, accounted for by the first p PCs at time $t + k$. These curves show that, as expected, the cumulative variance accounted for by the lagged PCs decreases with increasing lag. Continuing the calculations for autocorrelation up to the full 80-dimensional SSTA field shows that the 3, 6, 9, and 12 month lagged SSTAs represent approximately 60%, 35%, 26%, and 24% of the variability of the original SSTA field, respectively.

If our hypothesis is that the predictable portion of the precipitation variance is due to the instantaneous state of the SSTA field and that seasonal predictability is derived from the low-frequency evolution of this field, then the curves presented in Fig. 13 should align with the MSESS values generated using lagged SSTA–precipitation relationships for each value of p. We test this for each lag, with results shown in Fig. 14. In all cases we display both the theoretical curves (solid line) and the MSESS curves (dotted line) using the same axis scales as was used in Fig. 8.

Although the relationship is not perfect, the proximity of the fit is once again remarkable. For lags of between 6 and 12 months, the MSESS is slightly higher than would be expected based on the preceding calculations, while for a lag of 3 months the MSESS indicates a somewhat more significant underestimation. We are unable to provide

definitive reasons for these divergences, with possible explanations including complex interactions between SSTA and other boundary conditions such as soil moisture, vegetation, and snow extent, or they might simply be due to sampling variability. Despite these minor departures, both the asymptotic MSESS for all 48 PCs and the breakdown of the improvement in MSESS by dimension p (in particular the ''plateauing'' of the MSESS with higher-order PCs for 6-month and longer lags) can only be explained by the conclusion that it is the instantaneous structure of the SSTA field that provides the lower-boundary forcing for global precipitation variability, and it is the persistence structure of the SSTA field that allows the global precipitation field to be forecast into the future.

The plateauing of the MSESS with higher-order PCs means that it is possible to find the asymptotic MSESS for each lag directly from the results in Fig. 14. This is achieved simply by reading the MSESS at the point where $p = 48$, as the improvements (increases) in the MSESS beyond this point are likely to be marginal. We therefore estimate the asymptotic MSESS at 3-, 6-, 9-, and 12-month lags to be 7.3%, 5.4%, 4.2%, and 3.7%, respectively.

f. Implications of a global trend

We finally turn to the issue of estimating the implications of the global warming trend. In section 2 we discussed the removal of the global trend by subtracting the global mean SSTA time series from the SSTA time series at each grid point, such that the final predictor data represent both anomaly (in the sense that the mean at each grid point is removed) and detrended (in the sense that the global mean time series is removed at each grid point) data. This preprocessing was conducted for statistical reasons to ensure orthogonality of both the principal components and the eigenvectors.

By regressing the trend time series against each of the SSTA grid points of the *original* SSTA dataset (i.e., before detrending), we calculate that this trend accounts for 6.1% of the variance of this original SSTA dataset. As discussed, this trend was removed from the SSTA field before performing the PCA operation, and therefore up to now the implications of the trend have been ignored. How is this likely to affect our estimate of global precipitation predictability?

In section 4a we showed that 100% of the global *detrended* SSTA variability accounted for 14.7% of the global precipitation variability. Thus, assuming that we had left the trend in the data, and also assuming that the MSESS is a direct function of the variance of the SSTA dataset, we would expect an increase in the MSESS of about 0.9% (i.e., $14.7\% \times 6.1\%/100\%$). To test whether this is the case, we use the linear regression formulation in Eq. (2) to regress precipitation at each grid point against the trend time series $\mathbf{x}_T$. The outcome of this regression is a globally averaged MSESS of 1.3%. Although this diverges slightly from the 0.9% we expected based on simple variance accounting, it is sufficiently close to provide additional support that the predictability of global precipitation can be described as a function of the variance of the SSTA dataset. Furthermore, these results suggest that the short-term variability in the SSTA field dominates the precipitation variability at the seasonal time scale, with the trend imparting some long-term persistence that would make statistical forecasting using the SSTA field at time scales greater than a year theoretically possible (although with limited skill).

This analysis provides qualified support to the conclusions of New et al. (2001) that the global warming trend contributes a significant, but thus far relatively small, component of the variance of the global precipitation dataset. However, we do not propose that the methodology and results provided here can be used to separate the variability due to ''natural'' climate variability and the variability due to anthropogenic greenhouse gas emissions, since the implications of global warming are likely to be felt in part through the modulation of natural climate modes (for example, see the discussion on the implications of climate change on ENSO variability in Meehl et al. 2007).

As discussed earlier, the results provided here support the conclusion that the relationship between global SSTA variance and global precipitation variance is likely to remain relatively stationary at the global scale into the foreseeable future, although such a conclusion may not be valid when considering precipitation variability at smaller regional scales.

5. Discussion and conclusions

In the preceding analysis, we presented a simple approach based on linear regression to estimate the upper bound of predictability of global precipitation at the seasonal time scale. We defined predictability as the unbiased estimate of maximum predictive skill, calculated through an appropriately specified statistic, which can be achieved if all the relevant predictors are included in the model. We proposed that the MSESS represents a sensible statistic to measure predictability, as it represents the percentage reduction in variance achieved by a given model compared to a climatology model estimated using the sample mean. We then presented a simple synthetic example to show that an estimate of predictability can be achieved using a dataset of finite length and demonstrated that the estimation of the globally averaged MSESS produces an approximately unbiased estimate for the true predictability.

TABLE 1. Summary of predictability estimates using a combination of alternative precipitation datasets, analysis periods, grid scales, temporal scales, and lags.

Precipitation dataset	Analysis period	Grid scale (lat × lon)	Temporal scale	Lag	Asymptotic MSESS
GHCN	1900–2007	5° × 5°	Seasonal	Concurrent	14.7%
CRU	1900–98	5° × 5°	Seasonal	Concurrent	15.4%
CRU	1900–98	2.5° × 3.75°	Seasonal	Concurrent	16.1%
GHCN	1900–2007	5° × 5°	Monthly	Concurrent	9.9%
GHCN	1950–2007	5° × 5°	Seasonal	Concurrent	17.0%
GHCN	1900–2007	5° × 5°	Seasonal	3 month	7.3%
GHCN	1900–2007	5° × 5°	Seasonal	6 month	5.4%
GHCN	1900–2007	5° × 5°	Seasonal	9 month	4.2%
GHCN	1900–2007	5° × 5°	Seasonal	12 month	3.7%

In the analysis, we used PCA to derive an orthogonal representation of the global SSTA dataset and use this as the model predictors. The globally averaged MSESS was estimated by fitting a linear model separately at each precipitation grid point using the same p-dimensional pool of predictors, calculating the MSESS at each grid point, and deriving the global MSESS by taking the weighted average accounting for the surface area of each grid point.

Our estimates of the globally averaged MSESS are summarized in Table 1 and have been derived using two different precipitation datasets, a range of analysis periods, two spatial scales, and two temporal scales. The majority of the analysis was conducted using concurrent SSTA–precipitation relationships to reflect the relatively instantaneous (at the seasonal time scale) relationship between precipitation variability and variability in the external boundary forcing. The implication of lagging the SSTA–precipitation relationship was also tested by introducing lags of up to 12 months. The results of the analysis are summarized as follows:

1) There exists a direct relationship between the total variability accounted for by the principal components of SSTA and the globally averaged MSESS. This result was tested by using an alternative representation of the variability of the SSTA field using ICA and was found to be robust.
2) Using the first 48 principal components of SSTA field, which accounted for 92.8% of the total SSTA variance, the globally averaged MSESS is 13.6%, and a small extrapolation suggested that, if we could account for all the variance of the global SSTA field in our model, the asymptotic predictability would be 14.7%. This result only varied slightly using an alternative global gridded precipitation dataset.
3) The influence of record length was tested by incrementally removing the earlier parts of the record. The results showed that the MSESS increased by 35% for $p = 1$ and by 16% for $p = 20$ when considering only the post-1950 record, relative to the full 1900–2007 record. Although the relative contributions of decreasing measurement errors and nonstationarity due to low-frequency variability and/or anthropogenic climate change are difficult to separate, the results do not suggest any breakdown in the SSTA–precipitation relationship in the more recent record due to anthropogenic climate change at the global scale.
4) The hypothesis that the predictable component of precipitation is due to (approximately) instantaneous variability in the external boundary conditions, and that any seasonal predictability is due to the persistence structure of the boundary conditions, was confirmed for lags from 3 to 12 months.
5) Because of the relatively small percentage SSTA variance accounted for by the global warming trend compared to total seasonal SSTA variability, this trend accounts for only about 1.3% of total precipitation predictability. Caution is required when interpreting this as representing the fraction of variability attributed to anthropogenic global warming, since the impacts are unlikely to be expressed solely through the trend component. Furthermore, the use of global data obscures regional changes that may be much more significant. Nevertheless, the results highlight the importance of natural SST variability in accounting for temporal variability in precipitation data.

These results have numerous implications on the future of seasonal forecasting that we will describe below. First, however, we will review the assumptions on which the results were based.

a. Validity of model assumptions

As with any study, the validity of the conclusions is predicated on the validity of the underlying assumptions. In our case the principal assumptions were 1) that the global SSTA field is the dominant driver of long-term precipitation variability, 2) that precipitation variance can be partitioned between an unpredictable ''weather noise'' component and a more predictable externally

forced component, and 3) that the relationship between SSTA variability and global precipitation variability is linear and additive as suggested in Eq. (2). A pragmatic reason for these assumptions is that they made the analysis tractable, with the major alternative approach being the application of multiple AGCMs forced to historic SSTA variability but with differing initial conditions as proposed by Barnston et al. (2005). As we show below, however, these assumptions may not be unreasonable.

In the case of the first assumption, in addition to the sea surface temperatures field, external boundary conditions that are likely to be relevant for driving precipitation variability at the seasonal time scale include soil moisture, vegetation, snow cover, and sea ice. Nevertheless, SSTAs are generally regarded as the dominant boundary condition for precipitation predictability at the seasonal time scale (e.g., Charney and Shukla 1981; Palmer and Anderson 1994; Gershunov and Cayan 2003; Barnston et al. 2005). The relative dominance of the SST field for driving land surface precipitation is explained by the estimate that 85% of global evaporation occurs from the ocean surface and from the result that the level of recycling of moisture within a spatial region is small even for spatial areas of 1000 km or greater (Trenberth 1998; see also Shelton 2009), thereby indicating that water evaporated from the ocean is transported large distances before returning as precipitation. Furthermore, we propose that variability in the other proposed external boundary conditions such as soil moisture is largely driven by precipitation variability integrated over some previous time scale, such that much of the variability of these boundary conditions is likely to be highly correlated with variability in the SSTA dataset. Thus, the first assumption is likely to represent a reasonable reflection of reality.

The remaining assumptions, pertaining to the model structure in Eq. (2), are more difficult to justify physically. One approach to test the assumptions is to adopt the methodology proposed by Barnston et al. (2005) described above and examine the model outputs to estimate the proportion of precipitation variability that can be explained by variations in the boundary conditions. Such an approach would hopefully shed further light on relationship between the (potentially predictable) external forcing and the largely random internal atmospheric variability, including the question of whether the linear approximation is realistic. Limitations to this dynamical approach, however, include the likelihood that averaging across multiple AGCMs will not result in unbiased representations of precipitation and that the scale of precipitation processes are usually smaller than the scale of individual AGCM grids. Therefore, we suggest that this approach also would be limited by its own set of assumptions.

In the absence of the AGCM approach described above, our primary evidence that the assumptions underlying Eq. (2) are approximately valid are derived from the results presented in this paper. The first indication that the linearity assumption was approximately valid was presented in section 2, in which a range of skill scores including the MSESS, the Spearman rank correlation coefficient, and the linear error in probability space (LEPS) score were found to provide almost identical results for precipitation predictability using an index of ENSO as the predictor. This would not be expected to occur if the ENSO–precipitation relationship exhibited significant nonlinearity. A second, perhaps more convincing, line of evidence is that if the SSTA–precipitation relationship was highly nonlinear, then the relationship between variance accounted for by the PCs of the SSTA field and the MSESS would be expected to be much more complicated than that found in Fig. 8. For example, the PCs that represent variability in the tropics (e.g., PC1) might account for a disproportionate amount of the variance in global precipitation compared to PCs that more evenly account for variance across all latitudes (e.g., PC2). We emphasize that we do not suggest that the SSTA–precipitation relationship is linear; rather, we conclude that the relationship appears to be well approximated by a locally linear (i.e., linear within the bounds of variability implied by the historical record) relationship when averaged over the global scale.

Finally, we emphasize that, even if the underlying assumptions of Eq. (2) are not correct (e.g., see Hoerling et al. 1997) and the asymptotic predictability presented in this paper significantly underestimates true predictability, the additional predictability may not be accessible in a practical sense. For example, the cross-validation results described in section 4b highlight that the actual predictive skill that can be derived from regression modeling is substantially lower than our estimate of the asymptotic predictive skill because of the difficulty in correctly estimating model parameters from a finite training sample. Although our statistical formulation is very simple, the example highlights the necessity of making simplifying assumptions in order to generate a statistically robust result. This issue may not be as valid for dynamical approaches as it is for statistical forecasting; however, the difficulty in generating dynamical seasonal forecasts that outperform statistical forecasts suggests that our asymptotic estimate is likely to represent a practical upper bound for the foreseeable future.

b. Implications and future directions

How do our results align with our present understanding of climate variability? As discussed in the introduction, much of the research into climate variability

has focused on the ENSO phenomenon. Our results both confirm the importance of ENSO in seasonal forecasting and diminish it: confirm, because we show that the ENSO mode is the single most significant mode in the global SSTA field that also has high levels of seasonal persistence and therefore would logically be considered to be the most important climate mode for many regions around the world; diminish, because more than 80% of the total SSTA variance is not attributable to ENSO, suggesting the presence of significant non-ENSO-related predictability, which has yet to be fully exploited.

Beyond ENSO, much of the research into statistical forecasting has involved identifying stable climate modes, which usually are represented by climate ''indices.'' The advantage of using indices is that they often account for a large proportion of the SSTA variance (either regionally or globally) and exhibit a persistence structure that allows for the generation of forecasts with lead times of one or more seasons. This approach also can be justified with reference to the cross-validated MSESS presented in Fig. 10 and discussed in section 4b. In that case the best cross-validated forecast performance was obtained using only the first two principal components, with the contribution of higher-order principal components to the predictive skill being outweighed by the penalty that cross validation places on model complexity. In consequence, when one wishes to develop operational seasonal forecasts, the focus must be on the one or several climate modes that account for the greatest amount of variance in the response dataset rather than on the variability of the full SSTA field, which can only be accounted for by using a high-dimensional predictor dataset.

How might the limitations of short data length be overcome to develop improved seasonal prediction models? From a statistical perspective the use of larger spatial fields can to some degree substitute for short temporal records, which justifies the widespread use of multivariate statistical techniques such as canonical correlation analysis (CCA; e.g., Barnston and Ropelewski 1992; Bretherton et al. 1992). Furthermore, although in section 4c we show some small improvement in predictability by using only the more recent and presumably higher quality data, this must be weighed against the benefits of using a longer train dataset to develop the statistical model (e.g., see discussion in Wilks 2008).

From a dynamical perspective, by generating a large ensemble with differing atmospheric initial conditions, GCMs provide a tool that ultimately may overcome data limitations associated with empirical techniques. Currently, statistical and dynamical approaches exhibit comparable forecast skill (Coelho et al. 2006; Wilks 2008), and the question of whether such dynamical approaches will be capable of exceeding empirically based methods remains a subject of active debate (e.g., Anderson et al. 1999; Van den Dool 2007). Recently, there also has been research on the development of hybrid approaches that combine the advantages of empirical and dynamical modeling methods. Such approaches include a Bayesian forecast assimilation procedure that combines empirical and dynamical approaches and has been shown to outperform empirical or coupled multimodel predictions in isolation when applied to South American rainfall (Coelho et al. 2006) and the development of statistical correction procedures to account for biases in atmospheric general circulation model outputs (Tippett et al. 2005). Such hybrid approaches hold considerable promise in developing the next generation of seasonal forecasts.

Regardless of approach, our estimate of 14.7% predictability at the global scale suggests that, on average, 85.3% of the variability in seasonal precipitation field is due to random variations or weather noise (Barnston et al. 2005), which is not predictable beyond the deterministic weather predictability barrier. Specific regions and seasons may exhibit predictability significantly higher than this, and the value of even small amounts of predictive skill to decision makers is often high (e.g., Hamlet et al. 2002). The ultimate conclusion, however, is that estimates of future seasonal precipitation throughout most of the world will be dominated by a term that can best be described as random.

Acknowledgments. We wish to thank the Australian Research Council for funding this research. We also wish to acknowledge the contributions of the anonymous reviewers of this manuscript.

REFERENCES

Aires, F., A. Chedin, and J. P. Nadal, 2000: Independent component analysis of multivariate time series: Application to the tropical SST variability. *J. Geophys. Res.,* **105** (D13), 17 437–17 455.

Anderson, J., H. Van den Dool, A. Barnston, W. Chen, W. Stern, and J. Ploshay, 1999: Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bull. Amer. Meteor. Soc.,* **80,** 1349–1361.

Barnston, A. G., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere. *J. Climate,* **7,** 1513–1564.

——, and C. F. Ropelewski, 1992: Prediction of ENSO episodes using canonical correlation analysis. *J. Climate,* **5,** 1316–1345.

——, A. Kumar, L. Goddard, and M. P. Hoerling, 2005: Improving seasonal prediction practices through attribution of climate variability. *Bull. Amer. Meteor. Soc.,* **86,** 59–72.

Bigg, G. R., T. D. Jickells, P. S. Liss, and T. J. Osborne, 2003: The role of oceans in climate. *Int. J. Climatol.,* **23,** 1127–1159.

Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate,* **5,** 541–560.

Casey, K. S., and P. Cornillon, 2001: Global and regional sea surface temperature trends. *J. Climate,* **14,** 3801–3818.

Charney, J. G., and J. Shukla, 1981: Predictability of monsoons. *Monsoon Dynamics. Proceedings of the Joint IUTAM/IUGG International Symposium on Monsoon Dynamics,* J. Lighthill and R. P. Pierce, Eds., Cambridge University Press, 99–110.

Coelho, C. A. S., D. B. Stephenson, M. Balmaseda, F. J. Doblas-Reyes, and G. J. van Oldenborgh, 2006: Towards an integrated seasonal forecasting system for South America. *J. Climate,* **19,** 3704–3720.

Dai, A., and T. M. L. Wigley, 2000: Global patterns of ENSO-induced precipitation. *Geophys. Res. Lett.,* **27,** 1283–1286.

Diaz, H. F., M. P. Hoerling, and J. K. Eischeid, 2001: ENSO variability, teleconnections and climate change. *Int. J. Climatol.,* **21,** 1845–1862.

Drosdowsky, W., and L. E. Chambers, 2001: Near-global sea surface temperature anomalies as predictors of Australian seasonal rainfall. *J. Climate,* **14,** 1677–1687.

Efron, B., and R. Tibshirani, 1993: *An Introduction to the Bootstrap*. Chapman & Hall, 436 pp.

Gershunov, A., and D. R. Cayan, 2003: Heavy daily precipitation frequency over the contiguous United States: Sources of climate variability and seasonal predictability. *J. Climate,* **16,** 2752–2765.

Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol.,* **21,** 1111–1152.

Hamlet, A. F., D. Huppert, and D. P. Lettenmaier, 2002: Economic value of long-lead streamflow forecasts for Columbia River hydropower. *J. Water Resour. Plann. Manage.,* **128,** 91–101.

Hastie, T., R. Tibshirani, and J. Friedman, 2001: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 533 pp.

Hoerling, M. P., and A. Kumar, 2003: The perfect ocean for drought. *Science,* **299,** 691–699.

——, ——, and M. Zhong, 1997: El Niño, La Niña and the nonlinearity of their teleconnections. *J. Climate,* **10,** 1769–1786.

Horel, J. D., and J. M. Wallace, 1981: Planetary-scale atmospheric phenomena associated with the Southern Oscillation. *Mon. Wea. Rev.,* **109,** 813–829.

Hulme, M., and T. J. Osborne, 1998: Precipitation sensitivity to global warming: Comparison of observations with HadCM2 simulations. *Geophys. Res. Lett.,* **25,** 3379–3382.

Hyvarinen, A., J. Karhunen, and E. Oja, 2001: *Independent Component Analysis*. John Wiley and Sons, 481 pp.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.

Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res.,* **103** (C9), 18 567–18 589.

Lee, T. W., 1998: *Independent Component Analysis: Theory and Applications*. Kluwer Academic, 248 pp.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.,* **20,** 130–141.

Madden, R. A., 1989: On predicting probability distributions of time-averaged meteorological data. *J. Climate,* **2,** 922–928.

Meehl, G. A., and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 747–846.

New, M., M. Hulme, and P. Jones, 1999: Representing twentieth-century space–time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *J. Climate,* **12,** 829–856.

——, M. Todd, M. Hulme, and P. Jones, 2001: Precipitation measurements and trends in the twentieth century. *Int. J. Climatol.,* **21,** 1899–1922.

Nicholls, N., 1989: Sea surface temperatures and Australian winter rainfall. *J. Climate,* **2,** 965–973.

Palmer, T. N., and D. L. T. Anderson, 1994: The prospects for seasonal forecasting. *Quart. J. Roy. Meteor. Soc.,* **120,** 755–793.

Peterson, T. C., and R. S. Vose, 1997: An overview of the Global Historical Climatology Network temperature data base. *Bull. Amer. Meteor. Soc.,* **78,** 2837–2849.

——, ——, R. Schmoyer, and V. Razuvaev, 1997: Global Historical Climatology Network (GHCN) quality control of monthly temperature data. *Int. J. Climatol.,* **18,** 1169–1180.

Potts, J. M., C. K. Folland, I. T. Jolliffe, and D. Sexton, 1996: Revised ''LEPS'' scores for assessing climate model simulations and long-range forecasts. *J. Climate,* **9,** 34–53.

Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography*. Elsevier, 425 pp.

Quan, X., M. Hoerling, J. Whitaker, G. Bates, and T. Xu, 2006: Diagnosing sources of U.S. seasonal forecast skill. *J. Climate,* **19,** 3279–3293.

Rajeevan, M., D. S. Pai, R. A. Kumar, and B. Lal, 2007: New statistical models for long-range forecasting of southwest monsoon rainfall over India. *Climate Dyn.,* **28,** 813–828.

Richman, M. B., 1986: Rotation of principal components. *J. Climatol.,* **6,** 293–335.

Ruiz, E. J., I. Cordery, and A. Sharma, 2005: Integrating ocean subsurface temperatures in statistical ENSO forecasts. *J. Climate,* **18,** 3571–3586.

Shelton, M. L., 2009: *Hydroclimatology: Perspectives and Applications*. Cambridge University Press, 426 pp.

Stuart, A., and J. K. Ord, 1987: *Kendall's Advanced Theory of Statistics*. Oxford University Press, 604 pp.

Tippett, M. K., L. Goddard, and A. G. Barnston, 2005: Statistical-dynamical seasonal forecasts of central southwest Asia winter precipitation. *J. Climate,* **18,** 1831–1843.

Trenberth, K. E., 1997: The definition of El Niño. *Bull. Amer. Meteor. Soc.,* **78,** 2771–2777.

——, 1998: Atmospheric moisture residence times and cycling: Implications for rainfall rates and climate change. *Climatic Change,* **39,** 667–694.

Van den Dool, H., 2007: *Empirical Methods in Short-Term Climate Prediction*. Oxford University Press, 215 pp.

Walker, G. T., 1923: World Weather I. *Mem. Indian Meteor. Dep.,* **24,** 75–131.

Ward, M. N., and C. K. Folland, 1991: Prediction of seasonal rainfall in the north Nordeste of Brazil using eigenvectors of sea-surface temperature. *Int. J. Climatol.,* **11,** 711–743.

Westra, S., C. Brown, U. Lall, I. Koch, and A. Sharma, 2009: Interpreting variability in global SST data using independent component analysis and principal component analysis. *Int. J. Climatol.,* **30,** 333–346, doi:10.1002/joc.1888.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. Elsevier, 627 pp.

——, 2008: Improved statistical seasonal forecasts using extended training data. *Int. J. Climatol.,* **28,** 1589–1598.