Global Sea Surface Temperature Forecasts Using an Improved Multimodel Approach

MOHAMMAD ZAVED KAISER KHAN, RAJESHWAR MEHROTRA, AND ASHISH SHARMA

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

A. SANKARASUBRAMANIAN

Department of Civil, Construction and Environmental Engineering, North Carolina State University, Raleigh, North Carolina

(Manuscript received 12 August 2013, in final form 24 January 2014)

ABSTRACT

With the availability of hindcasts or real-time forecasts from a number of coupled climate models, multimodel ensemble forecasting systems have gained popularity in recent years. However, many models share similar physics or modeling processes, which may lead to similar (or strongly correlated) forecasts. Assigning equal weights to each model in space and time may result in a biased forecast with narrower confidence limits than is appropriate. Although methods for combining forecasts that take into consideration differences in model accuracy over space and time exist, they suffer from a lack of consideration of the intermodel dependence that may exist. This study proposes an approach that considers the dependence among models while combining multimodel ensemble forecast. The approach is evaluated by combining sea surface temperature (SST) forecasts from five climate models for the period 1960–2005. The variable of interest, the monthly global sea surface temperature anomalies (SSTA) at a 5° × 5° latitude–longitude grid, is predicted three months in advance using the proposed algorithm. Results indicate that the proposed approach offers consistent and significant improvements for all the seasons over the majority of grid points compared to the case in which the dependence among the models is ignored. Consequently, the proposed approach of combining multiple models, taking into account the interdependence that exists, provides an attractive strategy to develop improved SST forecasts.

1. Introduction

Skilful climate predictions provide useful scientific information to planners and operational agencies to plan and develop contingency measures and strategies to deal with the adverse conditions. In this regard, seasonal to interannual (long lead) climate forecasts are issued on a regular basis by various national and international agencies using both coupled ocean–atmosphere general circulation models (CGCMs) (Saha et al. 2006; Weisheimer et al. 2009; Palmer et al. 2004) and atmospheric general circulation models (AGCMs) (Devineni and Sankarasubramanian 2010a,b). Multiple sources of uncertainties including parameterization, process representation, and initial conditions control the development of seasonal to interannual climate forecasts. Efforts in reducing the errors and uncertainties in climate forecasts are focused on improving the representation of processes in individual CGCMs and AGCMs as well as by combining multiple models (Hagedorn et al. 2005). The use of a single best model often leads to poorer predictive performance as a result of the greater parameter and structural uncertainties in the model design (Chowdhury and Sharma 2011). Moreover, identifying a best or a poorest model from a set of models is not usually possible because of their individual strengths, preference in application, and varying skill depending on the variable of interest, location, season, forecast lead time, and so on (Hagedorn et al. 2005).

Developing multimodel forecasts has been pursued using various techniques ranging from simple pooling of the ensembles to optimizing weights to maximize the skill of multimodel forecasts (Rajagopalan et al. 2002; Robertson et al. 2004) or statistically estimating the weights conditioned on the dominant predictor conditions (Devineni and Sankarasubramanian 2010a). In addition, various statistical techniques such as simple regression

Corresponding author address: Ashish Sharma, School of Civil and Environmental Engineering, The University of New South Wales, High St., Kensington, NSW 2052, Australia.
E-mail: a.sharma@unsw.edu.au

DOI: 10.1175/JCLI-D-13-00486.1

(Krishnamurti et al. 1999), dynamic pairwise weighting based on logistic regression (Chowdhury and Sharma 2011), objective Bayesian analysis (Delsole 2007), and advanced statistical techniques such as the canonical variate method (Mason and Mimmack 2002) also have been applied. Simple multimodel ensembles (MMEs) are performed by combining the individual ensemble forecasts with equal weights (Hagedorn et al. 2005). However, these approaches do not consider any temporal variations of the component model skill. In more refined approaches, the participating single-model ensembles (SMEs) are weighted according to their prior performance (Rajagopalan et al. 2002). Here information is taken from all participating single models, including the less skilful one and the climatology. While linear combinations of multimodel responses are used in order to reduce the predictive uncertainty of climatic and hydrological variables, recent studies have focused on the technique of model combination weight wherein individual model weights are allowed to change with time, called a dynamic combination model (Chowdhury and Sharma 2009a,b, 2011; Devineni and Sankarasubramanian 2010a,b). It may be noted that the predictions from static combination of multiple models improve the skill compared to any single model prediction (Chowdhury and Sharma 2009a). The dynamic combination approach outperforms the static model combination approach in predicting SST and in long-range Niño-3.4 predictions (Chowdhury and Sharma 2009a, 2011).

Recently, Wasko et al. (2013) presented a copula-based approach for combining model rainfall predictions in a spatial context. All the above discussed techniques do not consider the spatiotemporal dependencies among the model skill. For instance, if a model performs well over a given region (e.g., tropical Pacific), then higher weights should be given for that model as opposed to another model that does not perform well in that region. Here we propose an approach to combine seasonal SST forecasts from multiple GCMs and provide an ensemble SST forecast for the entire globe taking into account the dependence among model forecasts, which varies with time. The proposed approach is an improvement over the multimodel combination methodology of Devineni and Sankarasubramanian (2010a) and Chowdhury and Sharma (2009a, 2011). The performance of the proposed approach is compared and evaluated against the case in which the dependence among the models is ignored.

2. Data description

Monthly SST forecasts from five available CGCMs of the Ensemble-Based Predictions of Climate Changes and Their Impacts (ENSEMBLES) project (Weisheimer et al. 2009) are used to develop the multimodel forecasts for SST for four seasons [February–April (FMA), May–July (MJJ), August–October (ASO), and November–January (NDJ)] over the entire globe. These include the models produced by European Centre for Medium-Range Weather Forecasts (ECMWF), the Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), the Leibniz Institute of Marine Sciences (IFM-GEOMAR), Météo-France, and the Met Office (UKMO). The data are obtained from the data library of the International Research Institute for Climate and Society (IRI), New York. These retrospective forecasts were issued on the first of February, May, August, and November based on the updated respective initial conditions and these forecasts extend up to seven months lead time except for November, during which the interest is in evaluating the skill in predicting ENSO conditions up to fourteen months. All the models have nine ensemble members and span over 46 years (1960–2005). The data are available at 2.5° × 2.5° spatial resolution and are interpolated onto the 5° × 5° grid to match the Kaplan observed monthly SST data (obtained from the IRI data library) (Kaplan et al. 1998; Reynolds and Smith 1994) for direct comparison.

3. Methodology

As our aim is to develop seasonal forecasts, monthly SSTs are converted to seasonal SSTs. The individual model's systematic bias is removed by deducting the long-term average of the seasonal SST from individual values. Therefore, each model ensemble is represented as anomaly from the model's seasonal climatology. A nonparametric nearest neighbor approach is employed to identify K similar conditions to the observed monthly SST anomaly (SSTA) in October, January, April, and July (in the beginning of the seasonal forecast) for each year. Computed over K neighbors, the covariance matrix (discussed in detail in the next section) of the forecast is calculated every year, separately for each season, in order to ascertain the optimal weighting for each model. It is important to note that in Devineni and Sankarasubramanian (2010a) the weights for each month/season are estimated based on the mean squared deviation of each CGCM from the observation over K neighbors, whereas in our approach these are derived from a covariance matrix that accounts for the dependence among forecasts. As the covariance matrix changes with seasons and years, the model weights, derived from the covariance matrix, also change with time. More details on the algorithm are presented next.

a. Multimodel algorithm

Multimodel combinations of seasonal SSTA forecasts for the seasons NDJ, FMA, MJJ, and ASO for all grid points over the entire globe are developed following the procedure outlined below.

1) Read the SSTA data for the observed and model forecasts for a particular season (i.e., NDJ) over all global SST grid cells.
2) At each grid point, calculate the seasonal forecast errors of the five models over the period by subtracting the ensemble mean of the model forecast from the observed seasonal SSTA.
3) Let the observed SSTA of October 1960 (month before a particular season) be z. Search for K similar observations to z in the Octobers of the entire observed period (1961–2005; i.e., 45 yr, excluding the year under consideration) having minimum absolute difference (or Euclidean distance) between z and the observed monthly (October) SSTA of other years represented by x, according to the following:

$$d_i = |x_i - z|, \tag{1}$$

where the subscript i varies from 1 to 45. Also note down the corresponding seasonal forecast error. The distances d_i along with forecast error are sorted in ascending order and only the first K seasonal forecast errors are retained. Lall and Sharma (1996) proposed optimum value of K, $K = \sqrt{n}$ for a sample of size n.
4) From the forecast errors of the five models over the K neighbors, estimate the variance-covariance matrix for every year. Since we have five models, the size of the variance-covariance matrix is 5×5 with the diagonals representing the variance of the errors over the K neighbors.
5) Calculate weight associated with each model using the variance-covariance matrix [and Eq. (10)]. Compute multimodel forecast using weights and model forecasts.
6) Move on to the next grid point and go to step 2.
7) Move on to the next season and go to step 1.

b. Weight estimation using covariance matrix

At time t, the seasonal SSTA forecasts is expressed as Y_{t+h}; here h is referred to as the forecast horizon. Let $\hat{y}^c_{t+h,t}$ be the combined point forecast as a function of underlying forecasts $\hat{Y}_{t+h,t} = (\hat{y}_{1,t+h,t}, \hat{y}_{2,t+h,t}, \ldots, \hat{y}_{M,t+h,t})'$, with M denoting the number of models. It is considered that the loss function only depends on the forecast error $e_{t+h,t} = y_{t+h} - \hat{y}^c_{t+h,t}$. If the weights of models are $w_{t,h} = (w_{1,t+h,t}, \ldots, w_{M,t+h,t})'$, then a combined forecast $\hat{y}^c_{t+h,t}$ can be expressed as

$$\hat{Y}^c_{t+h,t} = w'\hat{Y}_{t+h,t}. \tag{2}$$

Timmermann (2006) suggests minimizing the mean squared error (MSE) loss by considering only the first two moments of the joint distribution of Y and $\hat{Y}$:

$$\begin{pmatrix} Y_{t+h} \\ \hat{Y}_{t+h,t} \end{pmatrix} \sim \begin{pmatrix} \mu_{yth} \\ \mu_{\hat{y}th} \end{pmatrix} \begin{pmatrix} \sigma^2_{yth} & \sigma'_{y\hat{y}th} \\ \sigma_{y\hat{y}th} & \Sigma_{\hat{y}\hat{y}th} \end{pmatrix}. \tag{3}$$

Here σ^2 is the variance of forecast error. We define the forecast errors associated with the M forecasts as $e = \iota Y - \hat{Y}$, where ι is an $M \times 1$ column vector of ones. Here the time subscript is dropped in order to avoid confusion. From Eq. (3) the covariance matrix of the forecast errors $\Sigma_e = E[ee']$ is given by Timmermann (2006):

$$\begin{aligned} \Sigma_e &= E[Y^2\iota\iota' + \hat{Y}\hat{Y}' - 2\hat{Y}\iota\hat{Y}'] \\ &= (\sigma_y^2 + \mu_y^2)\iota\iota' + \mu\mu' + \Sigma_{\hat{Y}\hat{Y}} \\ &\quad - \iota\sigma'_{Y\hat{Y}} - \sigma_{Y\hat{Y}}\iota' - \mu_y\iota\mu' - \mu_y\mu\iota'. \end{aligned} \tag{4}$$

For our analysis Σ_e has been estimated from the forecast errors of the models over the K nearest neighbors. Note that Σ_e can be quite unstable unless large datasets are available for estimation; also, the problem is particularly pronounced when the pairwise correlations are high (Clemen and Winkler 1986). To have unbiased combined forecasts, we consider minimizing the expected forecast error variance such that the individual model weights add up to one (Timmermann 2006):

$$\min(w'\Sigma_e w) \quad \text{such that} \quad w'\iota = 1 \tag{5}$$

provided that $\mu = \mu_y\iota$ so that

$$\mu_y^2\mu' + \mu\mu' - 2\mu_y\iota\mu' = 0. \tag{6}$$

We introduce a new variable called the Lagrange multiplier λ and the Lagrange function ζ is defined by

$$\zeta = w'\Sigma_e w - \lambda(w'\iota - 1). \tag{7}$$

Deriving the first-order necessary conditions by taking the partial derivatives for Eq. (7) and setting them equal to zero, we can get

$$\Sigma_e w = \frac{\lambda}{2}\iota. \tag{8}$$

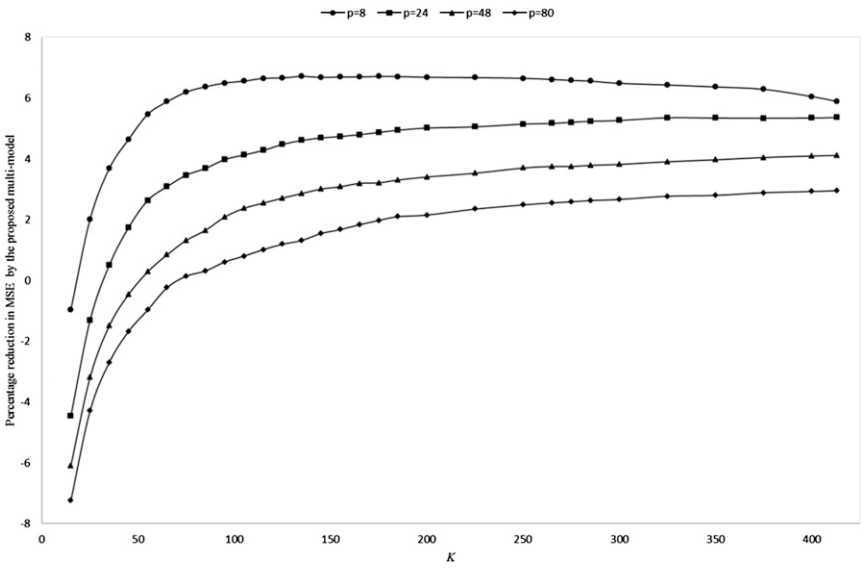

FIG. 1. Percentage reduction in MSE by the proposed multimodel averaged over all grid points vs number of nearest neighbors K considering different numbers of grid points p around a particular grid cell.

Assuming that Σ_e is invertible, after premultiplying by $\Sigma_e^{-1}\iota'$ and recalling that $\iota' w = 1$,

$$\frac{\lambda}{2} = (\iota'\Sigma_e^{-1}\iota)^{-1}. \tag{9}$$

Replacing the value of $\lambda/2$ in Eq. (8), the optimal weight w becomes

$$w = (\iota'\Sigma_e^{-1}\iota)^{-1}\Sigma_e^{-1}\iota. \tag{10}$$

The form of quadratic equations is solved using the quadprog package in R (see http://www.r-project.org/) in order to find the optimal weight for each model by considering the error covariance over K neighbors. It may be noted that in order to avoid the possibility of a negative weight being assigned to a model, we also imposed the condition that $w_i \geq 0$, where w_i ($i = 1, 2, 3, 4,$ and 5) represents the weight of each model. Once model weights are estimated over each grid point, the multimodel forecast for each season is calculated using Eq. (2) and then the MSE for each grid point over the 46 years is calculated.

It was observed that the covariance matrix estimated using the K nearest neighbors of 46 years of available data (45 data points for each season) was unstable on a few occasions. Hence, following the regionalization approach used in the statistical literature (Lowry and Glahn 1976) and the argument discussed in Hamill (2012), the data for a particular grid point were augmented by pooling the information from other nearby grid points that had relatively similar climatology. A sensitivity analysis was carried out by varying the number of nearby grid points and nearest neighbors. We considered 8, 24, 48, 80, and so on grid points p around the target grid cell and noted the percentage improvement in MSE (averaged over all the grid points) by the proposed multimodel versus the number of nearest neighbors K as shown in Fig. 1. The plots suggest that considering 8 nearby grid points offers the maximum improvement and, on average, 100 or more data points are needed to get a stable estimate of covariance matrix. Figure 2 shows the optimum K value for the both multimodels in four different seasons with 8 nearby grid points. The figure suggests that as few as 40 data points are needed to estimate the covariance matrix if correlations among the models are ignored, whereas it requires roughly 3 times more data points if they are considered. The suggested K value is found to be approximately 8 times more than the value proposed in Lall and Sharma (1996).

4. Results and discussion

Seasonal SST forecasts of the five models combined by the proposed multimodel combination approach described in section 3 are compared against the case in which dependence across the models is ignored (base model). Multimodel forecasts from both approaches are evaluated at each grid point over the entire globe for the four seasons (NDJ, FMA, MJJ, and ASO) during the

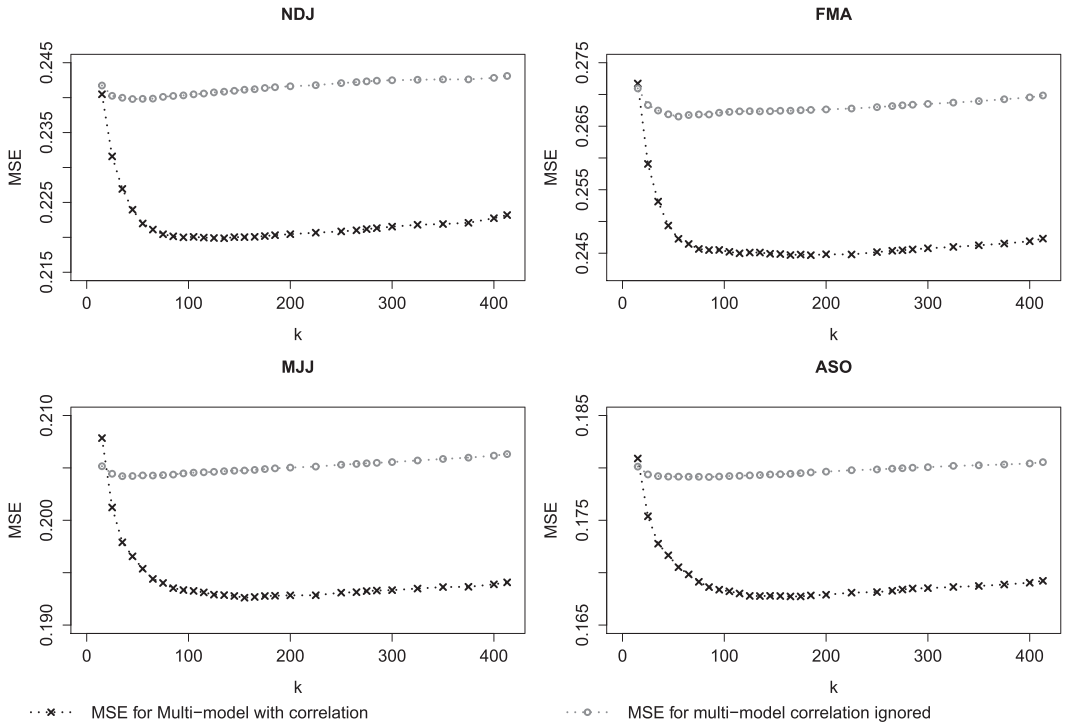

FIG. 2. Optimum K from 414-yr sample time series (46 yr $\times$ 9) in order to estimate variance-covariance matrix for different seasons.

period 1960–2005. The results presented next are obtained in a leave-one-out cross-validation setting. A separate analysis was conducted to see if lead-1 autocorrelations were significant and we found them to be quite small (around 0.2–0.3) for most of the grid points. A comparison of MSE between multimodel forecast and climatology is used as a measure of relative skill between the two approaches. The detailed analysis is discussed next.

a. MSE

This study uses MSE as a relative measure of skill between the two multimodels. Figure 3 illustrates the improvement and nonimprovement in multimodel SST

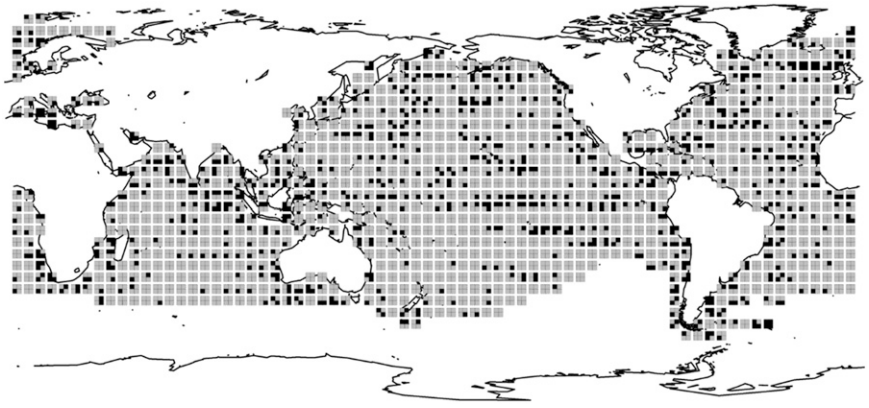

FIG. 3. Spatial distribution of the improvement or nonimprovement (difference in MSE) by the proposed multimodel in seasonal global SST forecast for the four seasons. Each 2 $\times$ 2 square represents a grid point in which each quadrant indicates a season. The top-left quadrant denotes NDJ, top-right quadrant denotes FMA, bottom-left quadrant denotes MJJ and bottom-right quadrant denotes ASO. Gray- and black-shaded boxes symbolize improvement and nonimprovement by the proposed multimodel, respectively.

TABLE 1. Percentage improvement in the number of grid points by the proposed multimodel approach.

Season	Improvement by the proposed multimodel	Avg MSE for the multimodel considering correlation	Avg MSE for the multimodel ignoring correlation
NDJ	81%	0.220	0.240
FMA	79%	0.245	0.267
MJJ	78%	0.193	0.204
ASO	78%	0.168	0.179

forecast expressed in terms of the difference in MSE across the time period (1960–2005) between the proposed approach and the base model for the four different seasons. In this figure, each grid cell is showing improvement or nonimprovement by appearing gray or black for every season. The plot shows that 78%–81% of grid cells around the entire globe exhibit significant improvement in terms of reduction of averaged MSE over time period (1960–2005) (Table 1). Maximum improvement is found in NDJ. On the other hand, Table 1 shows that MSE by the proposed multimodel averaged over all the grid points is the least for the season ASO. Moreover, Fig. 4 demonstrates the percentage improvement in terms of reduction of averaged MSE over the time period (1960–2005) by the proposed multimodel for the four different seasons. It may be noted that the forecasts have been bias corrected by centering them around their mean values. An alternate strategy could be to correct for the biases in the first-order as well as second-order moments. This issue was investigated and the results were found to be largely consistent with those reported here. However, as can be seen, the amount of improvement over grid cells varies from season to season. It would be useful to evaluate the performance of the proposed multimodel over different regions, such as Niño-3.4, the Indian Ocean dipole (IOD)–Indian Ocean west pole index (WPI) and Indian Ocean north pole index (NPI), the Indonesian index (II), the Tasman Sea index (TSI), the South Atlantic ocean dipole (SAOD) [northeast pole (NEP) and southwest pole (SWP)], and the Atlantic multidecadal oscillation (AMO) region (Table 2). Over the IOD–WPI region (10°N–10°S, 50°–70°E; Wajsowicz 2007), improvement is prevalent for all the seasons in the most of the grid cells. On the other hand, for the IOD–EPI region (0°–10°S, 90°–110°E; Wajsowicz 2007), substantial development over the number of grid cells can be identified for ASO and NDJ only; however, the percentage of MSE is increased in ASO. In case of the Niño-3.4 region (5°N–5°S, 170°–120°W; Wang et al. 2012), our model showed significant improvement for the seasons FMA, MJJ, and ASO, but not for NDJ. It appears that over this region all models perform equally well and therefore model weighting does not show significant improvements. Similarly, in

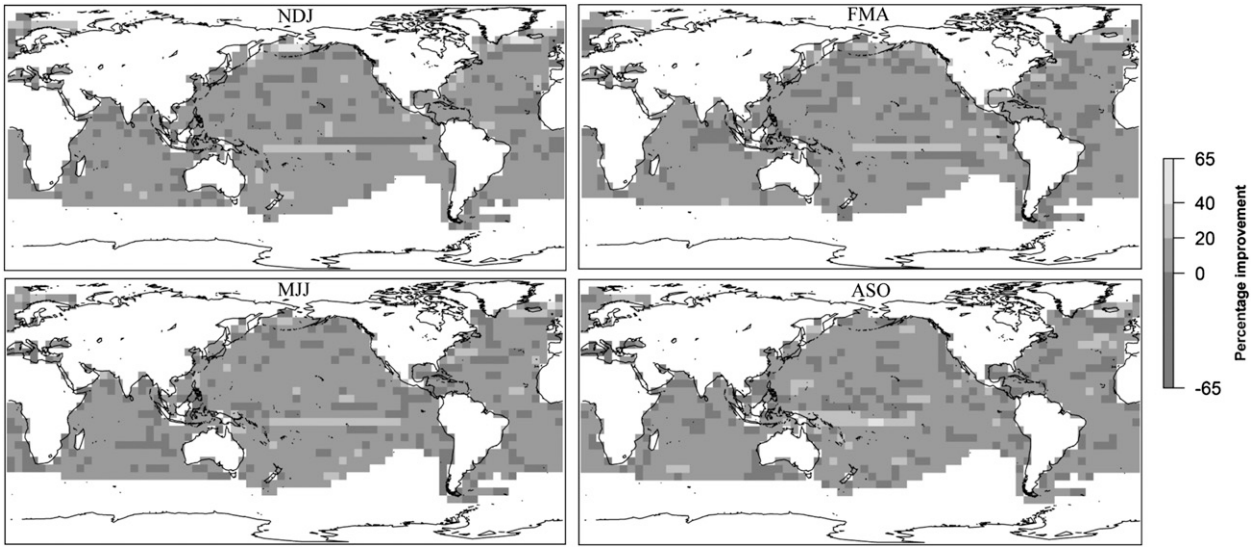

FIG. 4. Gradient maps showing the percentage of improvement by the proposed multimodel in respect to the multimodel where correlation is ignored in the seasonal global SST forecast for the four different seasons indicated at the top of the map.

TABLE 2. The performance evaluation of the proposed multimodel over different regions.

Region	Percentage of improved grid cells				Percentage reduction in MSE			
	NDJ	FMA	MJJ	ASO	NDJ	FMA	MJJ	ASO
IOD–WPI	94	94	94	81	6.5	5.9	5.2	1.9
IOD–EPI	67	33	50	83	4.3	1.1	2.9	−1.7
TSI	75	75	75	75	2.7	5.1	2.9	7.7
II	100	100	100	100	4.1	4.3	7.2	9.7
Niño-3.4	60	70	70	70	−1.6	6.7	4.7	8.0
AMO	71	75	81	71	4.3	3.2	4.7	5.6
SAOD–NEP	75	100	83	92	1.1	3.7	5.3	4.6
SAOD–SWP	100	83	67	94	10.0	6.4	4.3	7.7

case of TSI (30°–40°S, 150°–160°E; Wang et al. 2012), achievement is almost similar throughout the year. The II region (0°–10°S, 120°–130°E; Wang et al. 2012) shows improvement in all the seasons over all the grid points. The multimodel performed consistently better over the AMO region (0°–60°S, 80°W–0°; Gray et al. 2004) throughout all the four seasons. For the NEP of the SAOD region (0°–15°S, 10°–20°E; (Nnamchi et al. 2011), less improvement could be achieved in NDJ in contrast to the other three seasons. However, all the grid cells in the SWP of the SAOD (25°–40°S, 10°–40°W; Nnamchi et al. 2011) exhibit improvement in NDJ.

b. *Correlation among the models*

The use of model combinations for formulating improved climate forecasts has recently been extended for seasonal hydrologic predictions, with applications showing a range of situations where the advantages can be significant (Chowdhury and Sharma 2009b, 2011; Devineni and Sankarasubramanian 2010a; Robertson and Wang 2012; Wang et al. 2012). Here we further extend this multimodel combination approach to the cases when intermodel correlations are significant and illustrate the improvements that can be achieved if procedures for factoring in this dependence are utilized. Table 3 presents the pairwise spatial and temporal mean correlation averaged over all grid points for the five participating GCMs, those of the ECMWF, CMCC, IFM-GEOMAR, Météo-France, and the UKMO. The lower triangle of the top matrix, the upper triangle of the top matrix, the lower triangle of the bottom matrix, and the upper triangle of the bottom matrix indicate the correlations for the NDJ, FMA, MJJ, and ASO seasons, respectively. Correlation between models varies from 0.11 to 0.25 and is higher for NDJ and FMA seasons in comparison to the other two seasons.

c. *Weights for the models*

The weight associated with a model explains the contribution of that model in forecasting SSTA. As these

TABLE 3. Spatial and temporal mean correlation over all the grid points for the five participating GCMs—ECMWF, CMCC, IFM-GEOMAR, Météo-France, and the UKMO. The lower triangular matrix of the top-half of the table, the upper triangular matrix of the top-half of the table, the lower triangular matrix of the bottom-half of the table, and the upper triangular matrix of the bottom-half of the table represent the correlations among the CGCMs for the seasons NDJ, FMA, MJJ, and ASO, respectively.

		ECMWF	CMCC	IFM-GEOMAR	Météo-France	UKMO	
				FMA			
ECMWF	NDJ	1.00	0.18	0.13	0.21	0.19	FMA
CMCC		0.17	1.00	0.15	0.24	0.21	
IFM-GEOMAR		0.14	0.17	1.00	0.14	0.13	
Météo-France		0.18	0.23	0.15	1.00	0.25	
UKMO		0.17	0.22	0.14	0.22	1.00	
				NDJ			
				ASO			
ECMWF	MJJ	1.00	0.14	0.12	0.14	0.13	ASO
CMCC		0.15	1.00	0.12	0.15	0.13	
IFM-GEOMAR		0.13	0.14	1.00	0.12	0.11	
Météo-France		0.16	0.17	0.13	1.00	0.14	
UKMO		0.15	0.15	0.12	0.17	1.00	
				MJJ			

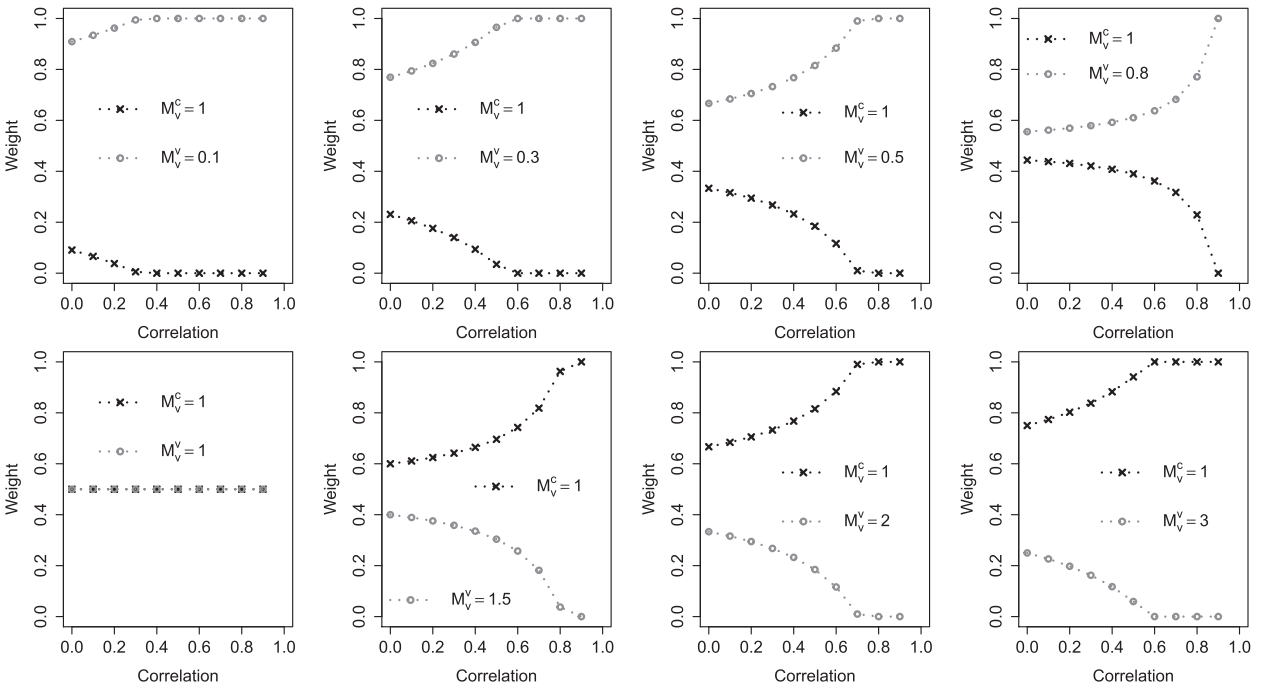

FIG. 5. Illustration of correlation-based weighing scheme in multimodel combination. The variance of one model (M_v^c) is constant (=1) and the variance of another model (M_v^v) varies from 0.1 to 3.0.

weights vary with time and space, their analysis and presentation is not straightforward. Figure 5 describes the correlation-based weighing scheme in the multimodel combination. In this illustration, two models are considered for multimodel combination. The variance of one model (M_v^c) is kept constant (=1) and the variance of another model (M_v^v) is varied from 0.1 to 3.0. This example indicates how the weight for the model with less variance (or the best model) increases as the correlation increases. While it is not straightforward to ascertain how this result will extend for more than two models, once can note the stability and added accuracy that is introduced by considering correlations, in contrast to the case where it is not taken into account (equivalent to the model weight for a correlation equal to zero in the figure).

Table 4 presents spatially and temporally averaged contribution of each model both for the proposed multimodel and the base multimodel in four different seasons. Considering the correlations among the models, ECMWF and UKMO are assigned higher weights, and CMCC and Météo-France are added as lower weights for the NDJ and FMA seasons. On the other hand, the IFM-GEOMAR and UKMO models contributed higher weights for the MJJ and ASO seasons. Interestingly, it can be observed form Table 5 that for all the four seasons, the models ECMWF and UKMO have the least error variances and these two models consistently get higher weights than the rest of the models if the correlations among the models are ignored. When considering correlations, for example, during the season NDJ, the correlation contributed higher weight to UKMO at the expense of CMCC and Météo-France (Table 3). It can also be noted that the UKMO model performs better in both the multimodel combinations for all the four seasons. Figure 6 shows the year-to-year distribution of model weights over the two randomly picked grid points for both cases (for NDJ). These grid points are picked up from each of the regions (e.g., Niño-3.4 and AMO).

TABLE 4. Weight in percentage of each CGCM for the multimodels averaged spatially and temporally in the four different seasons.

Season	ECMWF	CMCC	IFM-GEOMAR	Météo-France	UKMO
Multimodel considering correlation					
NDJ	26	10	24	9	31
FMA	27	8	21	10	33
MJJ	20	8	26	11	35
ASO	17	10	25	13	35
Multimodel ignoring correlation					
NDJ	24	16	18	19	23
FMA	24	16	18	19	23
MJJ	22	17	19	19	23
ASO	21	17	19	20	23

TABLE 5. Spatial and temporal mean error variance for each model.

Model	Season			
	NDJ	FMA	MJJ	ASO
ECMWF	0.23	0.24	0.21	0.20
CMCC	0.51	0.57	0.29	0.24
IFM-GEOMAR	0.35	0.37	0.24	0.21
Météo-France	0.35	0.41	0.24	0.20
UKMO	0.30	0.34	0.20	0.17

These plots confirm that the weights of the models for the both cases are time-space variant; however, this variation is significant in case of our proposed approach. It is also noteworthy that if correlations among the models are ignored, all of the participating CGCMs contribute in the combined forecast whereas the multimodel having correlation does not necessarily require the contribution of all the participating models. Finally, the contour plots of weights averaged over the entire time period across the globe for a season (e.g., NDJ) are shown in Fig. 7. This conveys that ECMWF is getting higher weight around the equator in the Pacific Ocean, the entire North American coasts of the North Atlantic Ocean, and the northern South American coasts of the South Atlantic Ocean. The UKMO model outperforms around the midlatitudes in the North and South Pacific Ocean, the midlatitudes in the Indian Ocean, and the midlatitudes in the North and South Atlantic Ocean. On the other hand, CMCC and Météo-France are contributing little.

5. Conclusions

Recent dynamic multimodel combination–based studies have the limitation of ignoring the correlations among the participating models. This study dynamically combined 3-month-ahead globally gridded sea surface temperature anomaly forecasts based on the degree of correlation between forecasts errors and the relative size of the individual models' forecast error variances. Five available CGCMs of the ENSEMBLES project are used to develop the multimodel forecast of the seasonal SSTA for the four seasons. Based on the observed monthly SSTA in October, January, April, and July computed over proposed K neighbors, the covariance matrix of the forecast errors is calculated in order to provide the weighting to each model every year. The proposed multimodel combination approach shows significant improvement in SSTA forecasts at a majority of the grid points regardless of seasons over the contemporary multimodel combination approach wherein correlations are ignored. The correlations among the models, and hence the model weights, vary from season

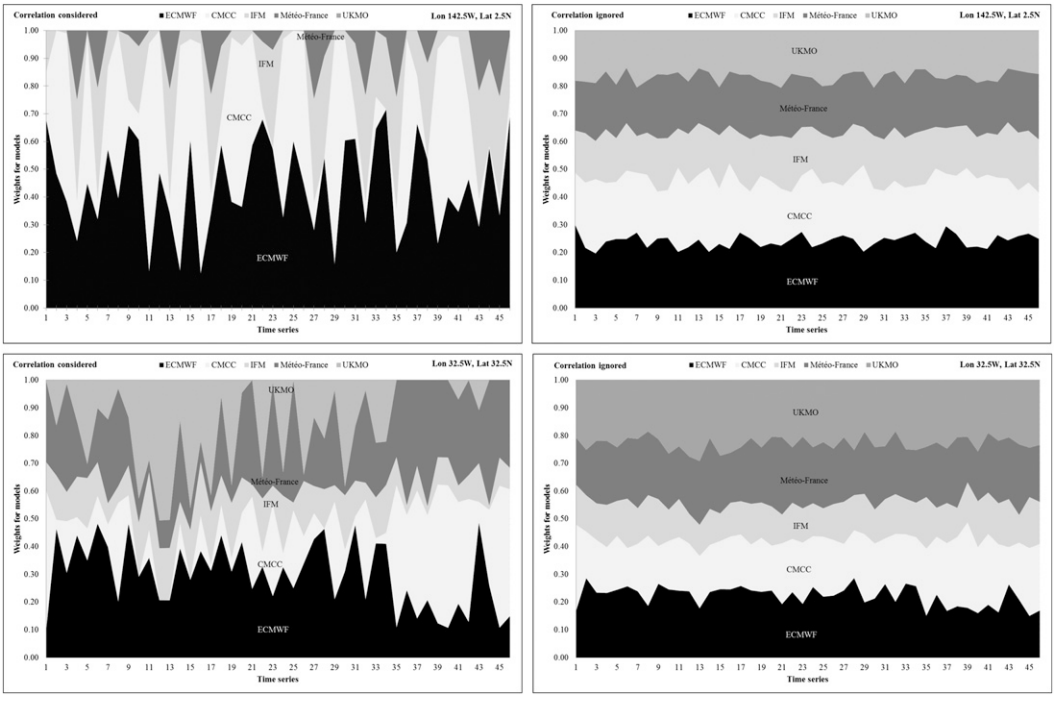

FIG. 6. The weights for the five GCMs at two selected grid cells for the both multimodel combination approaches (season: NDJ): (top) grid cell at 2.5°N, 142.5°W and (bottom) 32.5°N, 32.5°W.

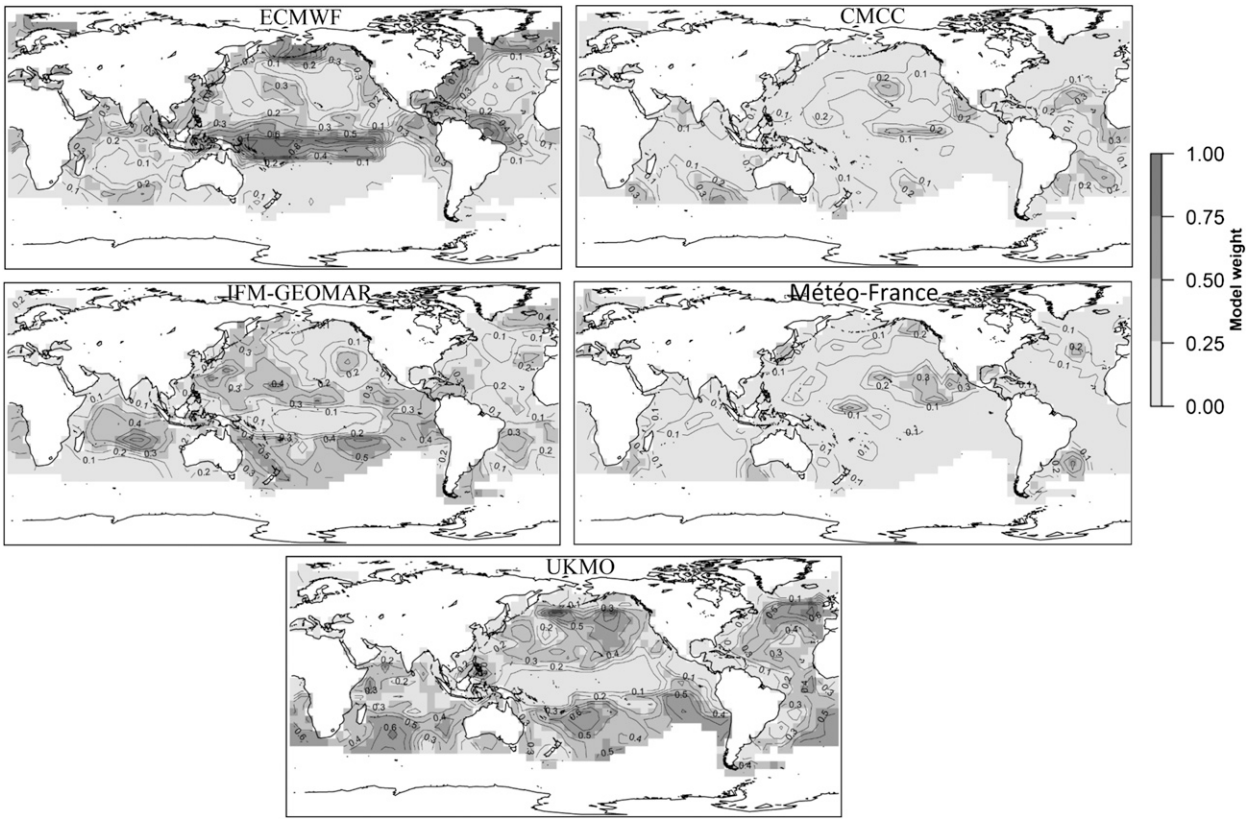

FIG. 7. The contour plots of weights averaged over the entire time period across the globe for a season (e.g., NDJ).

to season and from one location to another. In both multimodel approaches, the UKMO CGCM consistently contributes higher weights for all the four seasons.

Acknowledgments. We acknowledge the support from the Australian Research Council. The computations have been carried out using the freely available R statistical computing platform (http://www.r-project.org/).

REFERENCES

Chowdhury, S., and A. Sharma, 2009a: Long-range Niño-3.4 predictions using pairwise dynamic combinations of multiple models. *J. Climate,* **22,** 793–805, doi:10.1175/2008JCLI2210.1.

——, and ——, 2009b: Multisite seasonal forecast of arid river flows using a dynamic model combination approach. *Water Resour. Res.,* **45,** W10428, doi:10.1029/2008WR007510.

——, and ——, 2011: Global sea surface temperature forecasts using a pairwise dynamic combination approach. *J. Climate,* **24,** 1869–1877, doi:10.1175/2010JCLI3632.1.

Clemen, R. T., and R. L. Winkler, 1986: Combining economic forecasts. *J. Bus. Econ. Stat.,* **4,** 39–46. [Available online at http://www.jstor.org/stable/1391385.]

Delsole, T., 2007: A Bayesian framework for multimodel regression. *J. Climate,* **20,** 2810–2826, doi:10.1175/JCLI4179.1.

Devineni, N., and A. Sankarasubramanian, 2010a: Improved categorical winter precipitation forecasts through multimodel combinations of coupled GCMs. *Geophys. Res. Lett.,* **37,** L24704, doi:10.1029/2010GL044989.

——, and ——, 2010b: Improving the prediction of winter precipitation and temperature over the continental United States: Role of the ENSO state in developing multimodel combinations. *Mon. Wea. Rev.,* **138,** 2447–2468, doi:10.1175/2009MWR3112.1.

Gray, S. T., L. J. Graumlich, J. L. Betancourt, and G. T. Pederson, 2004: A tree-ring based reconstruction of the Atlantic multidecadal oscillation since 1567 A.D. *Geophys. Res. Lett.,* **31,** L12205, doi:10.1029/2004GL019932.

Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus,* **57A,** 219–233, doi:10.1111/j.1600-0870.2005.00103.x.

Hamill, T. M., 2012: Verification of TIGGE multimodel and ECMWF reforecast-calibrated probabilistic precipitation forecasts over the contiguous United States. *Mon. Wea. Rev.,* **140,** 2232–2252, doi:10.1175/MWR-D-11-00220.1.

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,** 18 567–18 589, doi:10.1029/97JC01736.

Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E Williford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. *Science,* **285,** 1548–1550, doi:10.1126/science.285.5433.1548.

Lall, U., and A. Sharma, 1996: A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resour. Res.,* **32,** 679–693, doi:10.1029/95WR02966.

Lowry, D. A., and H. R. Glahn, 1976: An operational model for forecasting probability of precipitation—PEATMOS PoP. *Mon. Wea. Rev.,* **104,** 221–232, doi:10.1175/1520-0493(1976)104<0221:AOMFFP>2.0.CO;2.

Mason, S. J., and G. M. Mimmack, 2002: Comparison of some statistical methods of probabilistic forecasting of ENSO. *J. Climate,* **15,** 8–29, doi:10.1175/1520-0442(2002)015<0008:COSSMO>2.0.CO;2.

Nnamchi, H. C., J. Li, and R. N. C. Anyadike, 2011: Does a dipole mode really exist in the South Atlantic Ocean? *J. Geophys. Res.,* **116,** D15104, doi:10.1029/2010JD015579.

Palmer, T. N., and Coauthors, 2004: Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER). *Bull. Amer. Meteor. Soc.,* **85,** 853–872, doi:10.1175/BAMS-85-6-853.

Rajagopalan, B., U. Lall, and S. E. Zebiak, 2002: Categorical climate forecasts through regularization and optimal combination of multiple GCM ensembles. *Mon. Wea. Rev.,* **130,** 1792–1811, doi:10.1175/1520-0493(2002)130<1792:CCFTRA>2.0.CO;2.

Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate,* **7,** 929–948, doi:10.1175/1520-0442(1994)007<0929:IGSSTA>2.0.CO;2.

Robertson, A. W., U. Lall, S. E. Zebiak, and L. Goddard, 2004: Improved combination of multiple atmospheric GCM ensembles for seasonal prediction. *Mon. Wea. Rev.,* **132,** 2732–2744, doi:10.1175/MWR2818.1.

Robertson, D. E., and Q. J. Wang, 2012: A Bayesian approach to predictor selection for seasonal streamflow forecasting. *J. Hydrometeor.,* **13,** 155–171, doi:10.1175/JHM-D-10-05009.1.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate,* **19,** 3483–3517, doi:10.1175/JCLI3812.1.

Timmermann, A., 2006: Forecast combinations. *Handbook of Economic Forecasting,* Vol, 1, G. Elliott et al., Eds., Elsevier, 135–196.

Wajsowicz, R. C., 2007: Seasonal-to-interannual forecasting of tropical Indian Ocean sea surface temperature anomalies: Potential predictability and barriers. *J. Climate,* **20,** 3320–3343, doi:10.1175/JCLI4162.1.

Wang, Q. J., A. Schepen, and D. E. Robertson, 2012: Merging seasonal rainfall forecasts from multiple statistical models through Bayesian model averaging. *J. Climate,* **25,** 5524–5537, doi:10.1175/JCLI-D-11-00386.1.

Wasko, C., A. Sharma, and P. Rasmussen, 2013: Improved spatial prediction: A combinatorial approach. *Water Resour. Res.,* **49,** 3927–3935, doi:10.1002/wrcr.20290.

Weisheimer, A., and Coauthors, 2009: ENSEMBLES: A new multimodel ensemble for seasonal-to-annual predictions—Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys. Res. Lett.,* **36,** L21711, doi:10.1029/2009GL040896.