Integration of Climate and Weather Information for Improving 15-Day-Ahead Accumulated Precipitation Forecasts

HUI WANG, A. SANKARASUBRAMANIAN, AND RANJI S. RANJITHAN

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

(Manuscript received 14 September 2011, in final form 2 August 2012)

ABSTRACT

Skillful medium-range weather forecasts are critical for water resources planning and management. This study aims to improve 15-day-ahead accumulated precipitation forecasts by combining biweekly weather and disaggregated climate forecasts. A combination scheme is developed to combine reforecasts from a numerical weather model and disaggregated climate forecasts from ECHAM4.5 for developing 15-day-ahead precipitation forecasts. Evaluation of the skill of the weather–climate information (WCI)-based biweekly forecasts under leave-five-out cross validation shows that WCI-based forecasts perform better than reforecasts in many grid points over the continental United States. Correlation between rank probability skill score (RPSS) and disaggregated ECHAM4.5 forecast errors reveals that the lower the error in the disaggregated forecasts, the better the performance of WCI forecasts. Weights analysis from the combination scheme also shows that the biweekly WCI forecasts perform better by assigning higher weights to the better-performing candidate forecasts (reforecasts or disaggregated ECHAM4.5 forecasts). Particularly, WCI forecasts perform better during the summer months during which reforecasts have limited skill. Even though the disaggregated climate forecasts do not perform well over many grid points, the primary reason WCI-based forecasts perform better than the reforecasts is due to the reduction in the overconfidence of the reforecasts. Since the disaggregated climate forecasts are better dispersed than the reforecasts, combining them with reforecasts results in reduced uncertainty in predicting the 15-day-ahead accumulated precipitation.

1. Introduction

Medium-range (10–15 days or submonthly time scales) weather forecasting has recently gained more attention owing to its practical importance with regard to water allocation and flood control in large basins. For instance, operation of reservoir systems critically depends on precipitation/streamflow over 2–4 weeks to develop water and energy management plans (Sankarasubramanian et al. 2009a,b). Chaotic characteristics of the atmosphere (Lorenz 1963) known as the ''butterfly effect'' limit weather predictability, resulting in limited skill beyond 10–15 days (Kalnay 2003). However, comparisons between current generation weather forecasting models and previous generation models show significant improvement in forecasting skill due to improved data collection and assimilation methods (Kalnay 2003; Gneiting and Raftery 2005). Efforts to improve the skill of weather forecasting models could be summarized broadly into three categories: 1) better process representation and parameterization of weather forecasting models, 2) improved estimates of initial conditions through better data assimilation, and 3) recalibration of predicted variables using model output statistics to reduce marginal bias and conditional bias in predictions. Over the past several decades, continued efforts to improve parameterization of numerical weather models has resulted in better weather forecasting skill (Kalnay 2003; Gneiting and Raftery 2005). Data assimilation based on Kalman filter techniques (Kalman 1960; Bishop et al. 2001; Hamill and Snyder 2002; Kalnay 2003) have been proposed to obtain better estimates of initial conditions. Ensemble forecasting techniques have also been pursued to account for uncertainties in initial conditions resulting in improved forecast reliability (Kalnay 2003).

Corresponding author address: A. Sankarasubramanian, 2501 Stinson Drive, Department of Civil, Construction, and Environmental Engineering, North Carolina State University, Campus Box 7908, Raleigh, NC 27695-7908.
E-mail: sankar_arumugam@ncsu.edu

DOI: 10.1175/JHM-D-11-0128.1

Various approaches have also been investigated for recalibration of weather forecasts that use the ensemble mean (Wilks and Hamill 2007) as well as the moments of the forecast probability distribution functions (PDFs; Wilks and Hamill 2007; Hamill et al. 2004).

Another approach that is gaining popularity for improving weather prediction is combining predictions from multiple weather models. Forecasts from single models suffer from spread deficiency since they tend to be overconfident under a given set of initial conditions (Whitaker and Loughe 1998; Whitaker et al. 2006). Raftery et al. (2005) applied Bayesian model averaging (BMA) to calibrate 48-h forecast ensembles of surface temperature and obtained better predictive PDFs from the raw ensembles. Johnson and Swinbank (2009) combined three individual models developed from the European Centre for Medium-Range Weather Forecasts (ECMWF), the Met Office of the United Kingdom, and the National Centers for Environmental Prediction (NCEP). They developed a global, 15-day multimodel ensemble for temperature, mean sea level pressure (MSLP), and 500-hPa height. Results from these studies show that a simple combination of ensembles from different forecast models could result in a significant improvement in the prediction skill. Several other studies (e.g., Krishnamurti et al. 1999; Rajagopalan et al. 2002; Barnston et al. 2003; Devineni et al. 2008; Devineni and Sankarasubramanian 2010a,b) also demonstrate that optimal combination of multiple climate models results in improved climate forecasts in comparison to the best model. Recently, Weigel et al. (2008) and Weigel and Bowler (2009) carried out a theoretical study based on synthetic forecasts to demonstrate that multimodel combination provides better forecasting skill than individual models and can even outperform the best single model provided the model is overconfident.

Though the current generation of numeric weather models has good skill in predicting precipitation up to a lead time of 3–5 days, the skill decreases substantially beyond one week (Kalnay 2003). In this study, we focus on improving medium-range (15-day ahead) precipitation forecasts by combining 15-day-ahead forecasts from a weather model and disaggregated precipitation forecasts over the same period developed from 1-month-ahead climate forecasts. Precipitation/streamflow forecasts obtained using disaggregation techniques (Valencia and Schaake 1973; Grygier and Stedinger 1990; Rajagopalan and Lall 1999) have been shown to provide better forecasts at subseasonal time scales. Precipitation forecasts updated every month from climate models have also been shown to have good skill in predicting 1-month-ahead precipitation over tropical regions (Sankarasubramanian et al. 2008). Thus, disaggregated forecasts from climate models could be utilized to improve the skill of 15-day accumulated precipitation forecasts. Further, most numerical weather models do not employ forecasted sea surface temperature (SST) conditions for obtaining 15-day-ahead precipitation forecasts. On the other hand, monthly climate forecasts updated every month use forecasted SSTs as boundary conditions for forcing the atmospheric general circulation models (GCMs; Li and Goddard 2005). Thus, any uncertainty in the initial conditions resulting from SSTs could possibly be reduced by utilizing the disaggregated climate forecasts to develop 15-day-ahead accumulated precipitation forecasts.

The main intent of this paper is to improve medium-range precipitation forecasts by combining 15-day-ahead precipitation forecasts from the reforecast dataset (Hamill et al. 2006) and disaggregated forecasts from 1-month-ahead precipitation forecasts from ECHAM4.5 GCM (Li and Goddard 2005). Figure 1 shows the schematic diagram of the proposed combination. First, we disaggregate 1-month-ahead precipitation forecasts from ECHAM4.5 to obtain 15-day-ahead precipitation forecasts (section 3a). Then, we combine the disaggregated precipitation forecasts from ECHAM4.5 with 15-day-ahead precipitation forecasts from reforecast model based on the combination scheme described in section 3b. Thus, the objectives of this study are 1) to develop a combination scheme that uses weather and climate forecasts to improve 15-day-ahead accumulated precipitation forecasts, 2) to evaluate their skill in developing improved 15-day-ahead forecasts, and 3) to investigate why the combination scheme results in improved biweekly forecasts.

This paper is organized as follows: Section 2 provides detailed information on weather and climate forecasts used in the study along with the motivation behind the proposed combination. The disaggregation approach and the combination algorithm are presented in detail in section 3. Section 4 discusses results by evaluating the 15-day-ahead weather–climate information-based precipitation forecasts and also investigates the basis behind the improved performance of the combination scheme. From here onward, we refer to the 15-day-ahead weather–climate information (WCI)-based precipitation forecasts as WCI precipitation forecasts. Finally, we conclude with the salient findings arising from the study.

2. Datasets

This study employs three datasets for developing WCI-based precipitation forecasts: (i) 15-day-ahead precipitation forecasts from the reforecasts database (Hamill et al. 2006), (ii) 1-month-ahead precipitation

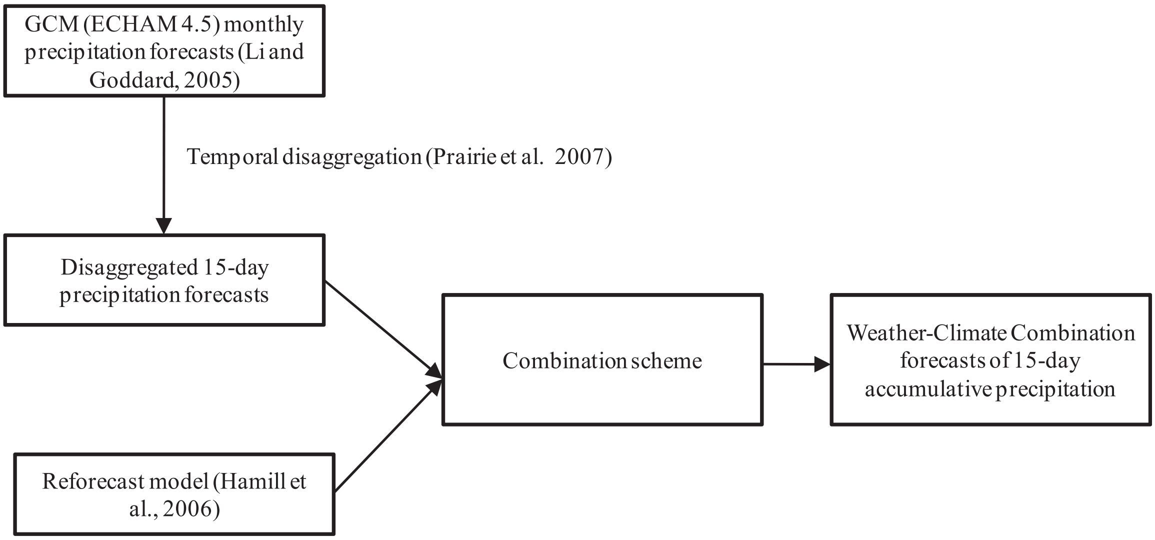

FIG. 1. Schematic diagram of the combination approach to improve medium-range weather forecasts.

forecasts from ECHAM4.5 (Li and Goddard 2005), and (iii) gridded observed daily precipitation for evaluating the developed forecasts (Higgins et al. 1996; http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.CPC/.REGIONAL/.US_Mexico/.daily/).

a. Weather forecasts

The National Oceanic and Atmospheric Administration (NOAA)'s reforecast model, which was run with 28 vertical sigma levels (Hamill et al. 2004) at a T62 resolution (about 200-km grid spacing), is employed in this study. NCEP–National Center for Atmospheric Research (NCAR) reanalysis data (Kistler et al. 2001) was used for initialization and seven bred pairs of initial conditions (Toth and Kalnay 1997) were used to develop 15 ensemble members. Fifteen-day-ahead retrospective precipitation forecasts are available every day from 1979 to date with forecasts updated every 12 h at about 2.5° resolution in latitude and longitude over the continental United States. The reforecast data were obtained from the Earth System Research Laboratory of NOAA (http://www.esrl.noaa.gov/psd/forecasts/reforecast/data.html). For this study, we consider the 15-day accumulated precipitation forecasts issued on the 1st and 15th days of the month for developing the WCI precipitation forecasts.

b. Climate forecasts

For climate forecasts, we consider 1-month-ahead retrospective precipitation forecasts from ECHAM4.5 (Roeckner et al. 1996) forced with constructed analog SSTs (Li and Goddard 2005). These retrospective forecasts are available for 7 months ahead with the forecasts being updated every month from January 1957 (http://iridl.ldeo.columbia.edu/SOURCES/.IRI/.FD/.ECHAM4p5/.Forecast/.ca_sst/.ensemble24/.MONTHLY/.prec/). We computed the ensemble mean of 1-month-ahead precipitation forecasts for the period 1957–2004 and then disaggregated them into two biweekly precipitation forecasts using the *k*-nearest-neighbor (*k*-nn) disaggregation algorithm of Prairie et al. (2007), which is discussed in detail in section 3a.

c. Observed precipitation

We obtained observed daily precipitation (Higgins et al. 1996), which is available at 1° resolution in latitude and longitude from the International Research Institute for Climate and Society (IRI) data library. Since the spatial resolution of reforecasts and the disaggregated climate model are at 2.5° in latitude and longitude, observed daily precipitation was also spatially averaged over a $2.5° \times 2.5°$ grid resulting in a total of 180 grid points over the continental United States. Further, daily time series of precipitation at 180 grid points were temporally aggregated to develop accumulated observed precipitation for the first 15 days and second 15 days of each month for evaluating the skill of the proposed scheme.

We utilize the data available from these three sources for evaluating the proposed scheme. The developed WCI-based forecasts are evaluated based on leave-five-out cross validation.

3. Methodology

In this section, we briefly describe the disaggregation methodology by Prairie et al. (2007) and the proposed scheme for combining weather and disaggregated climate forecasts.

a. Temporal disaggregation

Given that the 1-month-ahead precipitation forecasts from ECHAM4.5 are available over the entire month, we need to disaggregate them into two biweekly precipitation forecasts. Disaggregation is the process of dividing a given time series into its constituent parts. Both temporal disaggregation (e.g., from annual streamflow to monthly streamflow) and spatial disaggregation (e.g., from index gauge to upstream gauges) are useful in developing precipitation/streamflow scenarios related to water management. Temporal disaggregation of monthly precipitation forecasts to 15-day accumulated precipitation forecasts is of interest in this study. There are at least two special characteristics of disaggregation problems. First, the summability of the disaggregated precipitation should be maintained. This implies that the monthly precipitation is the sum of the disaggregated 15-day accumulated rainfall amounts. Second, the disaggregated precipitation series should preserve the statistical properties of the observed biweekly time series.

Various stochastic parametric models have been proposed for disaggregation. The linear stochastic framework was originally developed by Valencia and Schaake (1973), which was further modified by Grygier and Stedinger (1990) to obtain improved parameter estimates. Compared to parametric methods, nonparametric approaches have gained more attention for their simplicity and successful application in many hydrological problems (Lall and Sharma 1996). Tarboton et al. (1998) proposed a disaggregation approach using kernel-density estimators. A robust and simple approach based on resampling was proposed to achieve space–time disaggregation (Prairie et al. 2007) to avoid kernel density fitting. Recent studies on streamflow disaggregation based on other nonparametric approaches have also shown improvements over traditional parametric schemes (Sivakumar et al. 2004; Robertson et al. 2004; Lee et al. 2010). In this study, we apply the nonparametric disaggregation of Prairie et al. (2007) to disaggregate 1-month-ahead precipitation forecasts from ECHAM4.5 into biweekly precipitation forecasts. Daily observed precipitation is summed up from the 1st day to the 15th day and 16th day to the 30th day. We did not include the daily precipitation on the 31st day of relevant months (January, March, May, July, August, October, and December). For February, we considered the second forecasting period from the 16th to the 28th days of the month.

The disaggregation model takes the ensemble mean of the ECHAM4.5 monthly precipitation forecasts and then disaggregates that ensemble mean into two biweekly precipitation forecasts using the algorithm of Prairie et al. (2007). The disaggregation of climate forecasts into biweekly precipitation forecasts was performed in a leave-five-out cross-validation mode by leaving out the observed monthly precipitation and the ECHAM4.5 forecasts (i.e., ensemble mean) for the conditioning year and for another randomly chosen 4 years from the dataset available for the period 1957–2004. This resulted in a total of 43 years of observed monthly precipitation and ECHAM4.5 forecasts (i.e., ensemble mean) available as a training dataset for disaggregation. For the forecasted ensemble mean of precipitation forecasts from ECHAM4.5 in a given month, we identified "K" similar forecast conditions (i.e., K neighbors) by computing the Euclidean distance between the forecasted ensemble mean and the retained 43 years of ensemble means from the training dataset from that month. For the identified K similar years, we also obtained the observed monthly precipitation as well as the corresponding biweekly precipitation in those months from the training dataset. These biweekly observed precipitation values were resampled together to develop biweekly disaggregated forecasts having "N_2" ensembles, where "2" indicates the index for disaggregated forecasts. Thus, the identified neighbors/years on the forecasted precipitation space based on Euclidean distance were mapped to the corresponding observed biweekly precipitation in those years.

The number of ensemble members [$\mathrm{kw}(k) \times N_2$] that each identified neighbor/year—represented by the pair of observed biweekly precipitation values—was resampled in the disaggregated biweekly forecasts was specified by the kernel weighting function [Eq. (1)] suggested by Lall and Sharma (1996):

$$\mathrm{kw}(k) = \frac{k^{-1}}{\sum_{z=1}^{K} z^{-1}}, \tag{1}$$

where $\mathrm{kw}(k)$ represents the kernel weights of the kth neighbor, and k is the rank of the neighbor out of the total selected K neighbors. We evaluated the disaggregated series for its ability to reproduce the statistical properties, mean, standard deviation, and lag-1 correlation of the observed precipitation. Both the criteria—summability and preservation of statistical properties of the observed biweekly precipitation—were met in the disaggregated time series. We are not presenting these results here, since the performance of the proposed scheme in reproducing observed moments of the disaggregated series is well demonstrated in Prairie et al. (2007).

Leave-five-out cross-validated disaggregated forecasts showed reduced mean-square error in predicting the observed biweekly precipitation for $K = 10$ to 12 neighbors over all the 12 months (figure not shown). Thus, the temporal disaggregation from the ensemble mean of forecasted monthly precipitation to 15-day-ahead accumulated precipitation was carried out independently over each grid point by selecting $K = 10$ neighbors under leave-five-out cross validation. We selected $N_2 = 100$ ensemble members, since the estimated rank probability skill (RPS) of the disaggregated forecasts did not change by further increasing the ensemble size. Hence, we considered 100 ensemble members for developing the disaggregated forecasts. A total of $N_2 = 100$ ensembles of 15-day accumulated precipitation were produced to generate probabilistic biweekly precipitation forecasts for each month over the period 1957–2004. However, for this study, we only considered the disaggregated biweekly precipitation forecasts over the period 1979–2004 for developing WCI-based biweekly forecasts.

An initial investigation was carried out to examine the potential benefits in combining disaggregated climate forecasts and weather forecasts. One-month-ahead precipitation forecasts from ECHAM4.5 were disaggregated using the k-nn algorithm proposed by Prairie et al. (2007), and compared with the retrospective 15-day-ahead precipitation forecasts from the reforecast model. Since this was only an assessment, we did not consider cross validation for estimating mean-square skill score (MSSS) [Eq. (2)]. For a given grid point, mean-square errors (MSE) were calculated by averaging the squared error between the observed biweekly precipitation and the ensemble mean of the disaggregated forecast or the reforecast over 26 yr (1979–2004):

$$\mathrm{MSSS} = 1 - \frac{\mathrm{MSE}_c}{\mathrm{MSE}_w}. \tag{2}$$

The performance of both forecasts was compared using the MSSS in Eq. (2) with MSE_c and MSE_w denoting the MSE of the disaggregated climate forecasts and weather reforecasts, respectively.

From Eq. (2), we infer that for grid points with positive MSSS, disaggregated forecasts perform better than the reforecasts in terms of MSE. A comparison of forecast errors in January based on MSSS for the first 15 days (Fig. 2a) and the second 15 days (Fig. 2b) clearly showed that for various grid points, particularly over the upper Midwest, disaggregated ECHAM4.5 forecasts have lower MSE values than those of the reforecast model. This indicates that a proper combination of disaggregated climate forecasts and weather forecasts could result in improved medium-range forecasts. The next section discusses the methodology for obtaining the WCI-based forecasts over the continental United States.

b. Combination scheme

The combination algorithm (Fig. 3) presented here is based on the algorithm of Devineni and Sankarasubramanian (2010a) but was modified for combining biweekly forecasts. The essence of the combination scheme is to assign weights to each one of the candidate forecasts according to their skill in predicting the observed precipitation under similar forecast conditions. To start with, we have two biweekly precipitation forecasts in 180 grid points over the continental United States: a 15-member ensemble of biweekly forecasts from the NOAA reforecast model issued on the 1st and 16th days of the month and a 100-member ensemble of disaggregated biweekly forecasts from the ECHAM4.5 GCM. For these forecasts, we also have the corresponding observed biweekly precipitation over the 180 grid points. Using these two sources of information, we developed WCI-based biweekly forecasts for the period 1979–2004 in a leave-five-out cross-validation mode. Thus, developing WCI-based forecasts for each 15-day period had three major tasks: (i) obtaining the training dataset based on leave-five-out cross validation, (ii) estimating model weights based on the training dataset, and (iii) using the model weights to obtain probabilistic WCI-based forecasts.

1) Training dataset development

Given the 15-day forecasting period j, in a given year, we first developed a training dataset based on leave-five-out cross validation by dropping both the forecasts and the observation available for the conditioning year as well as another four randomly chosen years of observation and candidate biweekly forecasts. This resulted in a training dataset with 21 years of biweekly precipitation forecasts from two candidate models and the corresponding observed biweekly precipitation. Since most of the dynamical models exhibit significant bias in estimating their climatology, we converted the training dataset into anomaly values by subtracting both the candidate biweekly forecasts and the observed precipitation from their respective climatological values obtained from the retained 21 years. Similarly, we also converted the candidate ensemble forecasts available for the year "t" into anomaly values based on their

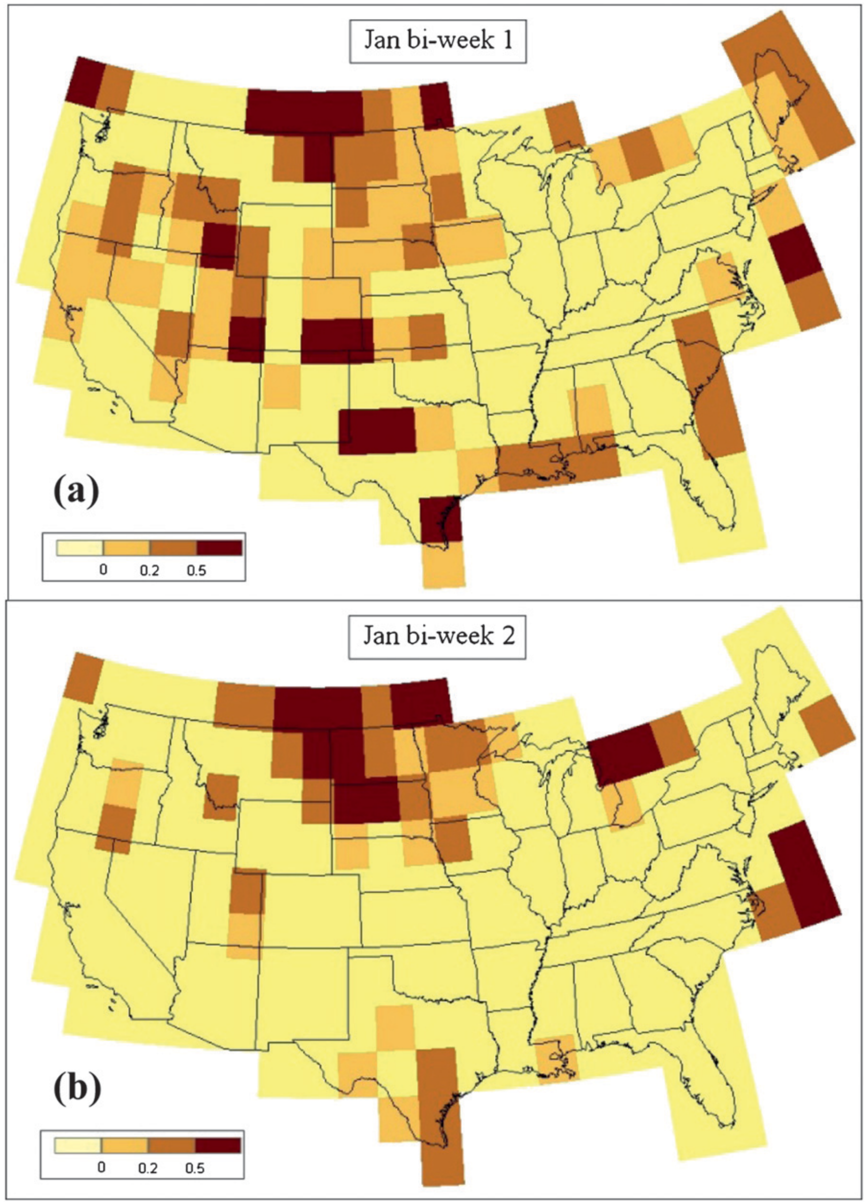

FIG. 2. MSSS between disaggregated ECHAM4.5 precipitation forecasts and the reforecasts for (a) the first 15 days and (b) the second 15 days in January. MSSS > 0 indicates MSE of disaggregated biweekly forecasts is lesser than the MSE of the reforecasts indicating potential for the proposed combination scheme.

respective climatological values estimated from the training dataset. Based on the 33rd and 67th percentiles of the observed precipitation anomalies from the training dataset, we converted the ensemble mean anomalies from each candidate forecast for year t into tercile probabilities, $\mathrm{PP}^m_{i,j,t}$, where $i = 1, 2, 3$ denotes the below-normal, normal, and above-normal categories, respectively; j denotes the biweekly period for which the forecast is developed; and m denotes the candidate forecasts (see Fig. 3 for further details). For details on how to calculate tercile probabilities for a given ensemble forecast, see Devineni et al. (2008) or Wilks (2006). Further, we also computed the ensemble mean of the candidate forecasts, $\mathrm{EP}^m_{j,t}$, for the year t based on the anomaly forecast ensembles.

To summarize the skill of the candidate forecasts in the training dataset, we first estimated the ensemble

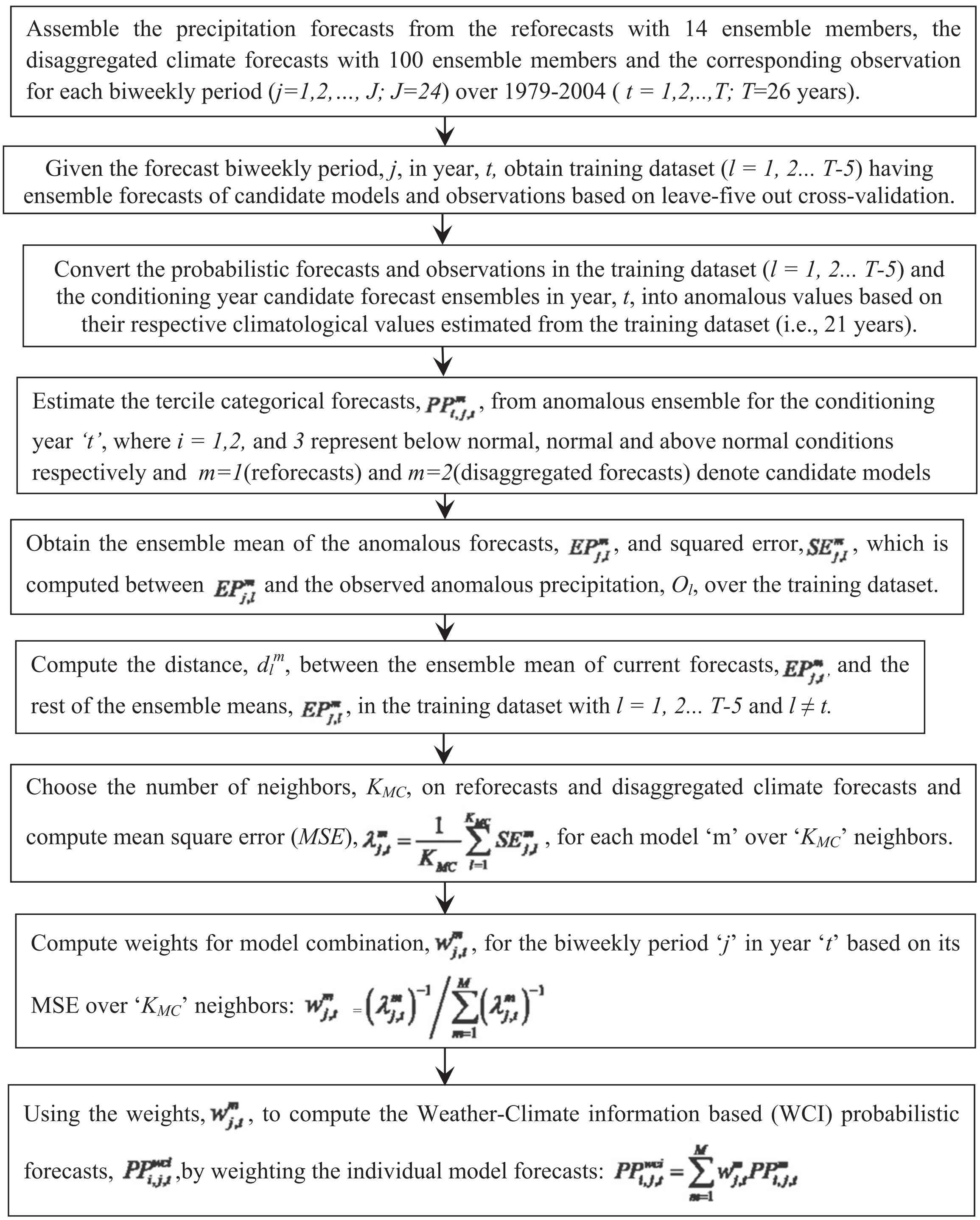

FIG. 3. Flowchart of the combination algorithm described in section 3b.

mean of the anomaly forecasts, $\mathrm{EP}^m_{j,t}$, for each model over the retained 21 years. Then, we obtained the skill of the forecasts in the training dataset based on squared errors, $\mathrm{SE}^m_{j,l}$, which is computed between the ensemble mean, $\mathrm{EP}^m_{j,l}$, and the observed precipitation anomalies, O_l, for each year "l" in the training dataset. These squared errors were used as the performance measure of the candidate forecasts for estimation of model weights in combination. It is important to note that both the tercile probabilities and the squared errors for the candidate forecasts were computed for each WCI forecast to be developed, since the biweekly climatological values of the forecasts and the observation were calculated based on the retained $(T - 5)$ data points under cross validation.

2) Model weights estimation

Given the ensemble mean of the reforecast $(m = 1)$ and the disaggregated climate forecasts $(m = 2)$, $\mathrm{EP}^m_{j,t}$, for the biweekly period "j" in year t, we identified similar forecast/neighboring conditions in the training dataset by computing the distance, d^m_l, between the current ensemble mean and the ensemble mean of the candidate forecasts, $\mathrm{EP}^m_{j,l}$, in the training dataset, where l $(l = 1, 2, \ldots, T - 5)$ denotes the years retained in the training dataset. For instance, to develop combined biweekly forecasts for the first two weeks of June in the year 1986 (j = 11th 15-day period in the year), we computed the distance d^m_l between the ensemble mean of the current forecast $\mathrm{EP}^m_{11,1986}$ and the rest of the forecasts $\mathrm{EP}^m_{11,l}$ ($l \neq$ 1986 and another four randomly left-out years) available for the 11th week in the training dataset. Based on the distance (d^m_l), we selected "K_{MC}" neighbors from both the reforecasts and the disaggregated climate forecasts and computed the mean-square error, $\lambda^m_{11,1986}$, by averaging the squared errors $(\mathrm{SE}^m_{j,l})$ over K_{MC} neighbors where the subscript "MC" denotes the identified neighbors under model combination. Figure 3 provides detailed steps and equations for estimating model weights for developing WCI-based precipitation forecasts. It is important to note that we also identify neighbors, K, under the disaggregation scheme (section 3a). Under disaggregation, to identify the K neighbors, we primarily look at similar conditions for years having similar monthly climate forecasts based on the monthly mean of the forecasts. On the other hand, under model combination, to identify the K_{MC} neighbors, we consider similar conditions for evaluating the performance of candidate forecasts. Based on the MSE value, the weight for an individual model "m" for the first two weeks of June in the year 1986 was calculated as

$$w^m_{11,1986} = \frac{(\lambda^m_{11,1986})^{-1}}{\sum_{m=1}^{M} (\lambda^m_{11,1986})^{-1}}, \tag{3}$$

where $w^m_{11,1986}$ represents the weight of model m for the first forecasting period in June for the year 1986. It is intuitive based on Eq. (3) that the lower the MSE, the higher its weights/influence in developing the combined biweekly forecasts.

3) Probabilistic WCI forecasts development

Given the weights, $w^m_{j,t}$, for the candidate models, a weighted average of the tercile forecasts (see Fig. 3) was calculated based on the tercile forecasts of reforecasts $(\mathrm{PP}^1_{i,j,t})$ and the disaggregated climate forecasts $(\mathrm{PP}^2_{i,j,t})$ to obtain WCI-based biweekly forecasts $(PP^{WCI}_{i,j,t})$ for the biweekly period j in year t. For a given value of K_{MC}, the procedure in Fig. 3 was repeated for each j over the period 1979–2004 for all the 108 grid points in the continental United States. To find the optimum K_{MC} value, we chose the K_{MC} neighbors that resulted in the best skill for the WCI-based forecasts for the selected biweekly forecast at a given grid point. The skill measure that we employed for selecting the optimal K_{MC} neighbors based on leave-five-out cross validation is the average RPS, which is discussed in detail in the next section using Eq. (4) (Wilks 2006). The optimum number of neighbors K_{MC} was found to be 10–15 depending on the biweekly forecast (not shown). Based on the optimum K_{MC} available for each grid point for the forecast biweekly period j, we developed an integrated biweekly precipitation forecasts derived from weather and climate information over the period 1979–2004. The next section presents a detailed evaluation of the skill of the WCI-based biweekly forecasts.

4. Results and analysis

RPS and rank probability skill score (RPSS) are used to evaluate the tercile forecasts (Wilks 2006) derived from the weather–climate information-based combination. RPS summarizes the sum of square of errors in the cumulative probabilities between the given categorical forecast and the observation. The cumulative probabilities of the observation, $\mathrm{PP}^{\mathrm{O}}_{i,j,t}$, is equal to zero below the observed category and is equal to one from the observed category. For instance, if the observation falls into the normal category in a given year, then cumulative probabilities for the observation could be represented as $\mathrm{PP}^{\mathrm{O}}_{1,j,t} = 0; \mathrm{PP}^{\mathrm{O}}_{2,j,t} = 1$ and $\mathrm{PP}^{\mathrm{O}}_{3,j,t} = 1$:

$$\mathrm{RPS}^m_{j,t} = \sum_{i=1}^{3} (\mathrm{PP}^m_{i,j,t} - \mathrm{PP}^{\mathrm{O}}_{i,j,t})^2. \tag{4}$$

Given the biweekly tercile forecasts from a model m $(\mathrm{PP}^m_{i,j,t})$, we can calculate the RPS using Eq. (4). Similar to MSSS [Eq. (1)], RPSS [Eq. (5)] compares the improvements in RPS of the WCI-based biweekly forecasts with the RPS of the reforecast model. The RPSS value represents how much of the RPS has been reduced (RPSS > 0) or increased (RPSS < 0) compared to the reference forecasts, which we consider as the reforecasts. When RPSS is positive, WCI forecasts outperform the reforecasts. On the other hand, when RPSS is negative, reforecasts outperforms WCI forecasts. It is important to note that RPS and RPSS are probabilistic performance measures. Thus, even a small improvement such as 5% in RPS is quite significant. Further details on the improvements in RPS and their significance testing could be found in Devineni et al. (2008) and Devineni and Sankarasubramanian (2010b). Previous studies have also shown that reforecasts are better than climatological odds (Hamill et al. 2004). Hence, we compared the improvements obtained from WCI forecasts over the reforecasts available for each 15-day period. Thus, we computed the average RPSS for each biweekly period over all the years:

$$\mathrm{RPSS}_{j,t}^{\mathrm{WCI}} = 1 - (\mathrm{RPS}_{j,t}^{\mathrm{WCI}}/\mathrm{RPS}_{j,t}^{1}). \tag{5}$$

Figure 4 shows the average RPSS for the two forecasting periods for 1 month in each of the four seasons. Grid points with light yellow color indicate where the reforecast model outperforms WCI-combined forecasts, whereas in the rest of the colored grids, the combination scheme reduces the error in probabilistic forecasts. The darker the color is, the higher the RPSS in the specific grid. For most of the grid points when RPSS > 0, combination approach reduces error by 10%–20% and for a few grids, the improvement is larger than 20% (RPSS $>$ 0.2). It is important to note that for a given grid point the skill is not the same for the two forecasting periods in a given month since the combination algorithm was carried out independently. Improvements in the skill of the WCI forecasts vary both spatially and temporally partly because of the individual models' performance in a given season and region. Improved skill in WCI biweekly forecasts is the highest over Texas and Utah in January. Similarly, WCI forecasts perform better over Arkansas, Ohio, and Pennsylvania in April and over the western United States in July and October.

We compare the RPS of the WCI forecasts and reforecasts by pooling the RPS for each month (Fig. 5) over the 180 grid points under both forecasting periods. The solid (dashed) boxplots represent RPS for the reforecast model (WCI-based forecasts). Comparing the median RPS of the forecasts from both models, WCI forecasts perform better in all the forecasting periods except in the second two weeks of January, first two weeks of February, and second two weeks of December. More importantly, the spread (interquartile range) of the RPS of the WCI-based forecasts is also lower than the spread of the RPS of reforecasts. Further, the improvements resulting from the WCI combination are substantial during the summer season compared to the other seasons. This is partly due to the poor performance of the reforecast model in the summer season (Clark et al. 2004; Hamill et al. 2008). The improved skill of WCI-based forecasts compared to the reforecast model reveals that there is utility in using disaggregated climate forecasts for improving the skill of medium-range weather forecasts. This is because the disaggregated climate forecasts capture the observed structure of the biweekly precipitation conditioned on the ensemble mean of the monthly climate forecasts.

Table 1 summarizes the total number of grids in which WCI forecasts outperform the reforecast model (i.e., RPSS > 0). RPSS_1 and RPSS_2 denote the RPSS for the first 15 days and the second 15 days, respectively. From Table 1, the performance of WCI-based forecasts is lower only during the second 15 days of January and November with 99 and 81 grid points exhibiting positive RPSS, respectively. The performance of WCI forecasts improves substantially during the spring and summer seasons. RPSS_1 is positive for 159 out of the 180 grid points and RPSS_2 is positive for 137 grid points in August. We also infer that the WCI-forecasts perform poorly over more than half of the selected 180 grid points during the second 15 days of January. This improved performance of WCI forecasts is mainly due to the poor performance of the reforecast model during the summer season. Further, the combination scheme results in improved biweekly forecasts particularly if the disaggregated ECHAM4.5 forecasts exhibit lower MSE than the reforecasts. For instance, the correlation between RPSS of WCI forecasts and the MSE of disaggregated ECHAM4.5 forecasts is -0.45 for the first 15 days in July. The negative correlation indicates that lower MSE values of the disaggregated forecasts are associated with higher values of RPSS. This is true for the other biweekly forecasting periods in summer season (results not shown).

To understand the importance of both candidate forecasts in reducing the uncertainty in predicting 15-day-ahead precipitation, we grouped the weights under each forecast model under two categories with RPSS > 0 and RPSS < 0 for both biweekly periods in few selected months [January (Fig. 6a), April (Fig. 6b), July (Fig. 6c), and October (Fig. 6d)] from various seasons. From Fig. 6, it is clear that the weights of the disaggregated forecasts are higher if the RPSS of the WCI-based forecasts is more than zero in all 4 months. On the other hand, in Fig. 6, the weights of the reforecasts are higher if the RPSS of the WCI-based forecasts is less than zero. Better performance of the disaggregated forecasts in July than the other months

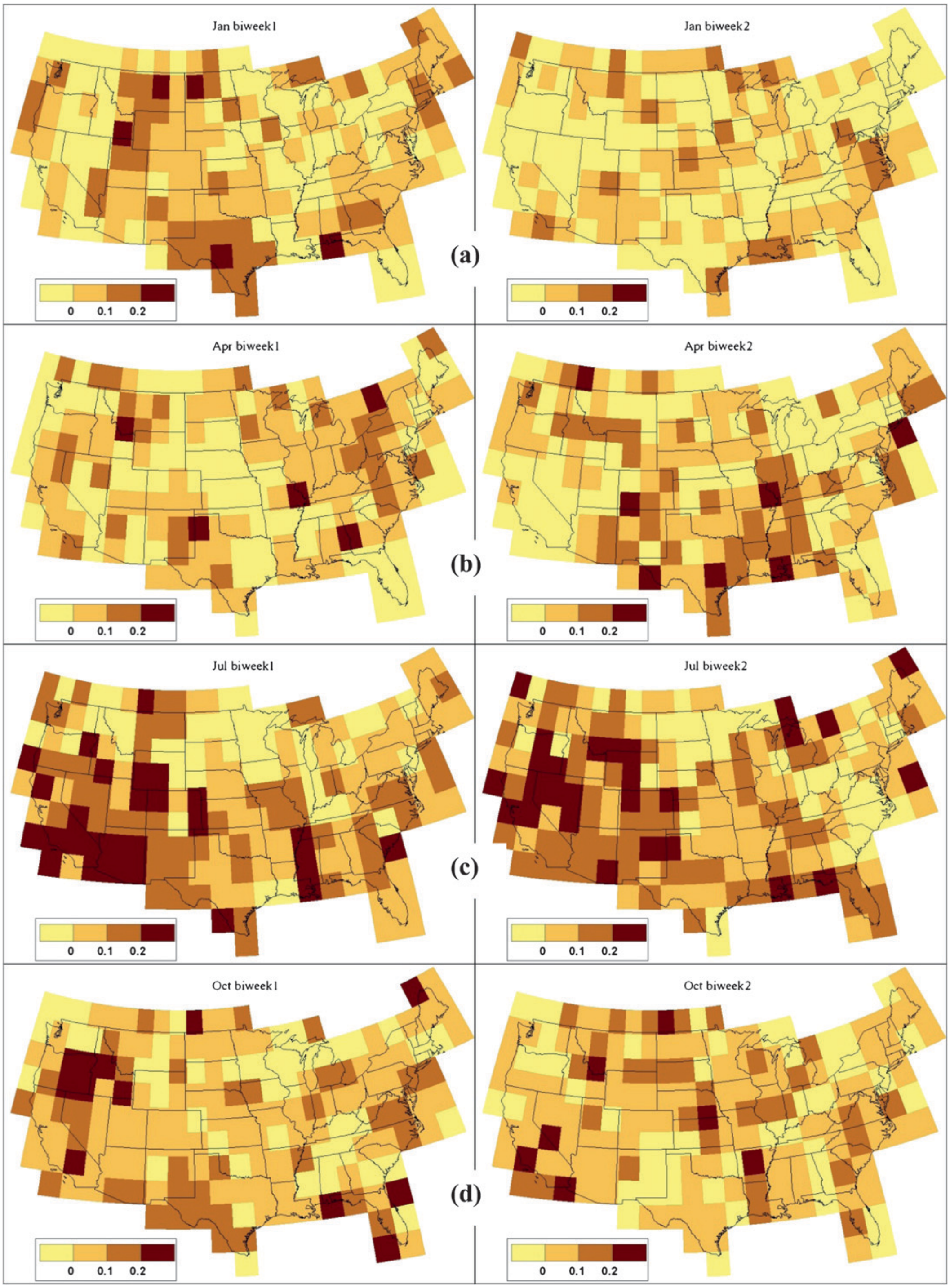

FIG. 4. RPSS of WCI-based forecasts over reforecasts with the left (right) column indicating the first (second) 15 days for the selected months over four seasons: (a) January, (b) April, (c) July, and (d) October.

is also clearly indicated in the boxplot (Fig. 6c), with the weight distribution of the disaggregated climate forecasts being positively skewed for RPSS > 0. Thus, the analysis of weights shows that the proposed combination scheme improves the skill of 15-day-ahead precipitation forecasts by appropriately evaluating the performance of the candidate forecasts under similar forecast/neighboring conditions.

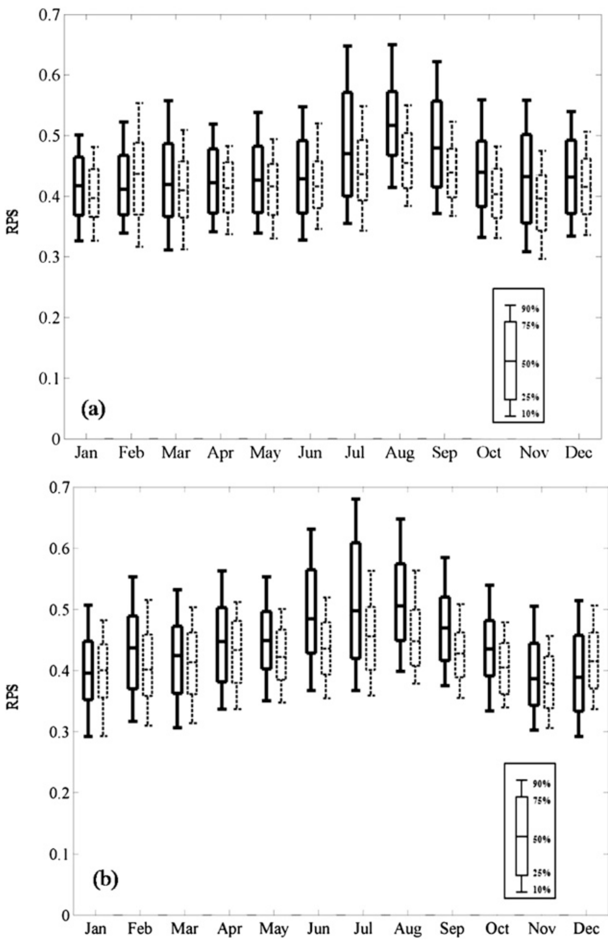

FIG. 5. Comparison of RPS between reforecast and WCI-based forecasts for all months. (a) Pooled RPS (180 × 26) for the first 15 days and (b) pooled RPS for the second 15 days. Solid (dashed) boxplots represent the RPS of reforecasts (WCI-based forecasts).

TABLE 1. Performance of the 15-day-ahead precipitation forecasts obtained from the combination of WCI expressed as the number of grid points exhibiting RPSS > 0 for the first 15 days ($RPSS_1$) and the second 15 days ($RPSS_2$).

	$RPSS_1 > 0$	$RPSS_2 > 0$
January	113	81
February	121	118
March	113	107
April	106	109
May	123	121
June	109	148
July	139	143
August	159	147
September	148	149
October	135	132
November	130	99
December	102	101

5. Discussion

Results presented in the previous section show that WCI-based combination scheme improves the 15-day-ahead precipitation forecasts over the continental United States. We also investigated the improved performance of WCI forecasts with various teleconnection indices such as Niño-3.4 (Kaplan et al. 1998) and Pacific–North American index (Barnston and Livezey 1987). Our analyses between the teleconnection indices and the improved skill of the WCI forecasts did not provide any clear association (figures not shown), thereby failing to provide any physical basis for the improved performance. However, Table 1 and Figs. 4 and 5 clearly show that more than 50% of the grid points exhibit improved skill in predicting 15-day-ahead precipitation using the proposed combination scheme with the exception being during the second biweekly forecast period in January. Further, the improvements are also particularly significant in forecasting warm-season precipitation (months of July–September), which is primarily due to the relatively lower skill of reforecasts during the summer season. The weights analysis (Fig. 6) also shows that the improved skill in WCI forecasts result from assigning higher weights for the disaggregated biweekly forecasts for conditions under which disaggregated forecasts have lower *MSE* values over K_{MC} similar conditions. This brings the discussion to two critical questions: (i) Given the three forecasts including the WCI-combined tercile forecasts, which model performs the best for a given grid point? and (ii) Given that there is no physical significance with any teleconnection/quasi-oscillatory climatic indices, why does the combination scheme improve the 15-day-ahead accumulated precipitation forecasts? Understanding these two issues is critical, since it can provide the basis for selecting the best forecasting model for a given region as well as the reasoning behind the improved skill resulting from the combination scheme.

With regard to the first question, we identified the best model at each grid point based on RPS (Figs. 7 and 8). Figure 7 shows the best model for the two forecasting periods under the selected months over the four seasons, whereas Fig. 8 summarizes the total number of grid points under which each model has the lowest RPS for each month. It is important to mention that disaggregated GCM forecasts are never the best-performing model and hence it is not shown in Figs. 7 and 8. From Figs. 7 and 8, WCI is the best-performing model and reforecasts from Hamil et al. (2006) perform better than climatology in predicting 15-day-ahead precipitation. Improvement from the combined forecasts is particularly significant during the summer and fall months (Figs.

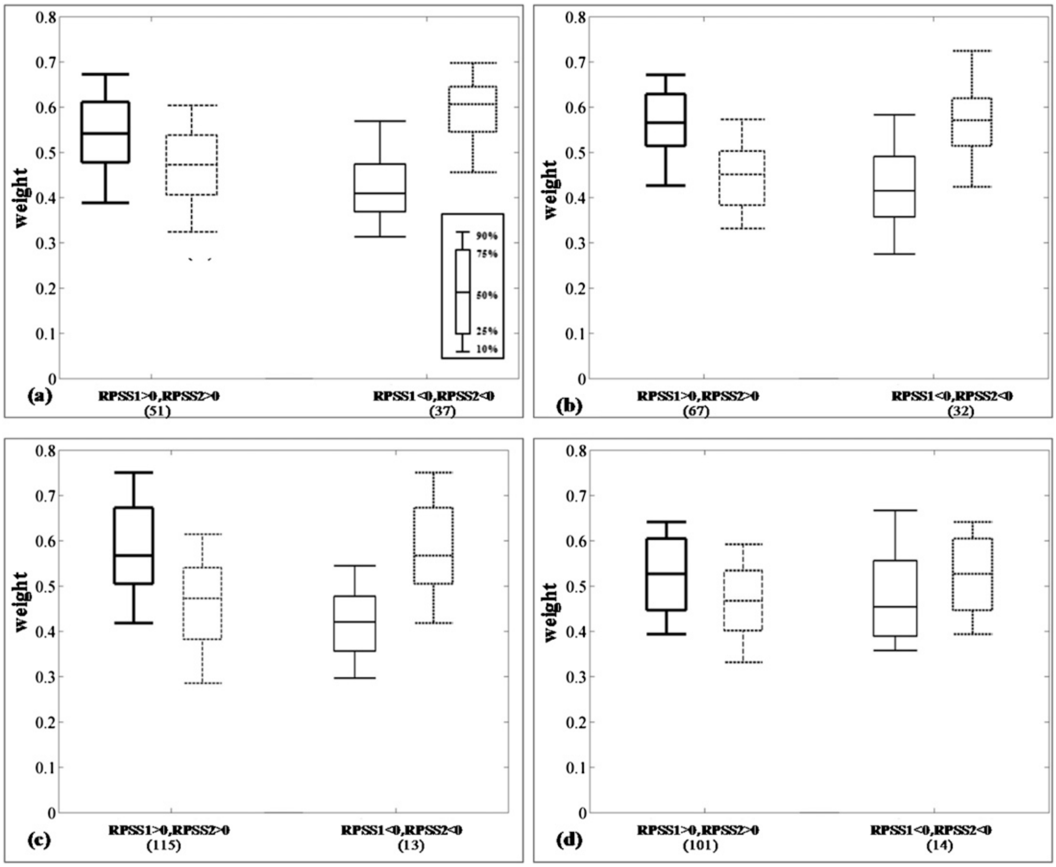

FIG. 6. Box plots of weights of the disaggregated GCM forecasts and reforecasts in the proposed combination scheme for different months: (a) January, (b) April, (c) July, and (d) October. Weights are grouped by pooling grid points solid (dotted) box plots represent the weights for the disaggregated climate forecasts (reforecasts). Numbers in parentheses under the RPSS indicate the number of grid points that have $RPSS_1 > 0$ and $RPSS_2 > 0$ and $RPSS_1 < 0$ and $RPSS_2 < 0$ for the WCI-based forecasts.

7 and 8). During the winter and spring months, improvement from the combination scheme is predominantly over the Sunbelt as well as over the upper Midwest and mid-Atlantic states. From Fig. 8, WCI-based forecasts perform better for many grid points during both the biweekly forecasting periods over the entire year. The only exception is during the second biweekly forecasting period in January under which the reforecasts performing better than WCI-based forecasts in a majority of grid points. Thus, it is clear that even though the disaggregated forecasts do not always perform the best, combining reforecasts and disaggregated ECHAM4.5 forecasts using the proposed combination scheme results in reduced RPS in many grid points over the continental United States.

To further understand why the combination of reforecasts and disaggregated forecasts perform better, we analyzed the probability distribution of these two candidate forecasts. Recently, Weigel et al. (2008) carried out a theoretical study to investigate under what situations multimodel forecasts can outperform even the best individual model. The study concluded that multimodel combination results in well-dispersed forecasts by reducing the overconfidence of the individual models. For instance, it is common for forecasts from numerical models to estimate high probability of occurrence of one particular tercile category (e.g., 90% probability of below-normal precipitation). This behavior is primarily due to the overconfidence of numerical models. Weigel et al. (2008) suggested methods to estimate the dispersion parameter, β, which ranges from zero to one, based on the standardized forecast ensembles having a climatology of zero and a standard deviation of one. When β is zero, the distribution of forecasts is well dispersed. For instance, a standardized climatological ensemble is a well-dispersed forecast but it does not have skill in predicting interannual variability in biweekly precipitation forecasts. On the other hand, when β is one, the distribution of candidate forecasts is overly dispersed.

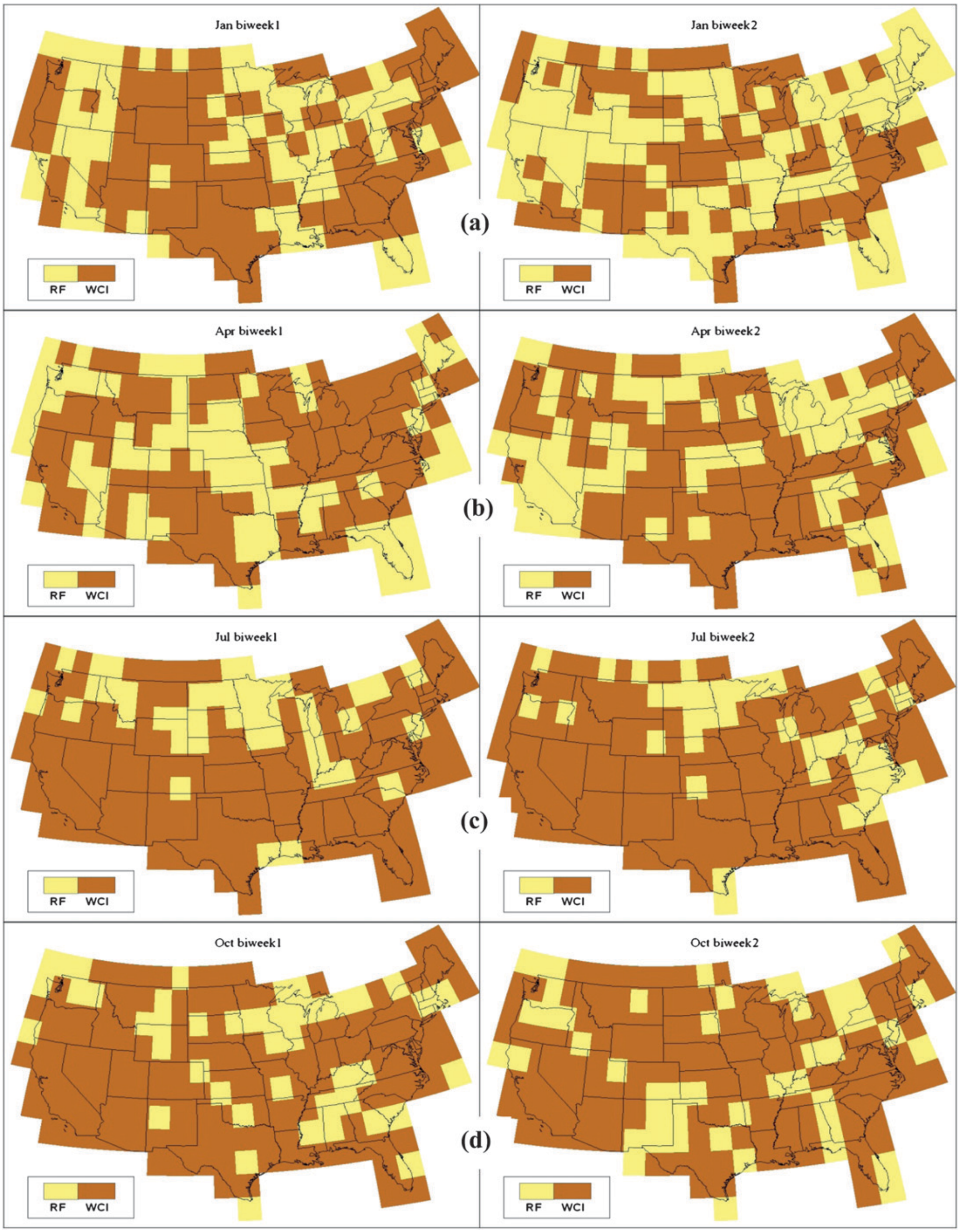

FIG. 7. Best model for 1 month under each season over the continental United States based on RPS. DF denotes disaggregated ECHAM4.5 forecasts, RF denotes the reforecasts, and WCI denotes weather–climate information-based biweekly forecasts.

The appendix provides details on estimating β for a given forecast. For additional details, see Weigel et al. (2008).

Using Eq. (A1), we estimated β for disaggregated ECHAM4.5 forecasts and reforecasts for four selected months over the 180 grid points. The estimated β values are grouped based on RPSS (Fig. 9). Figure 9 compares the estimated β between disaggregated GCM forecasts (solid box plots) and reforecasts (dotted box plots) based on RPSS_1 and RPSS_2 of the WCI-based forecasts.

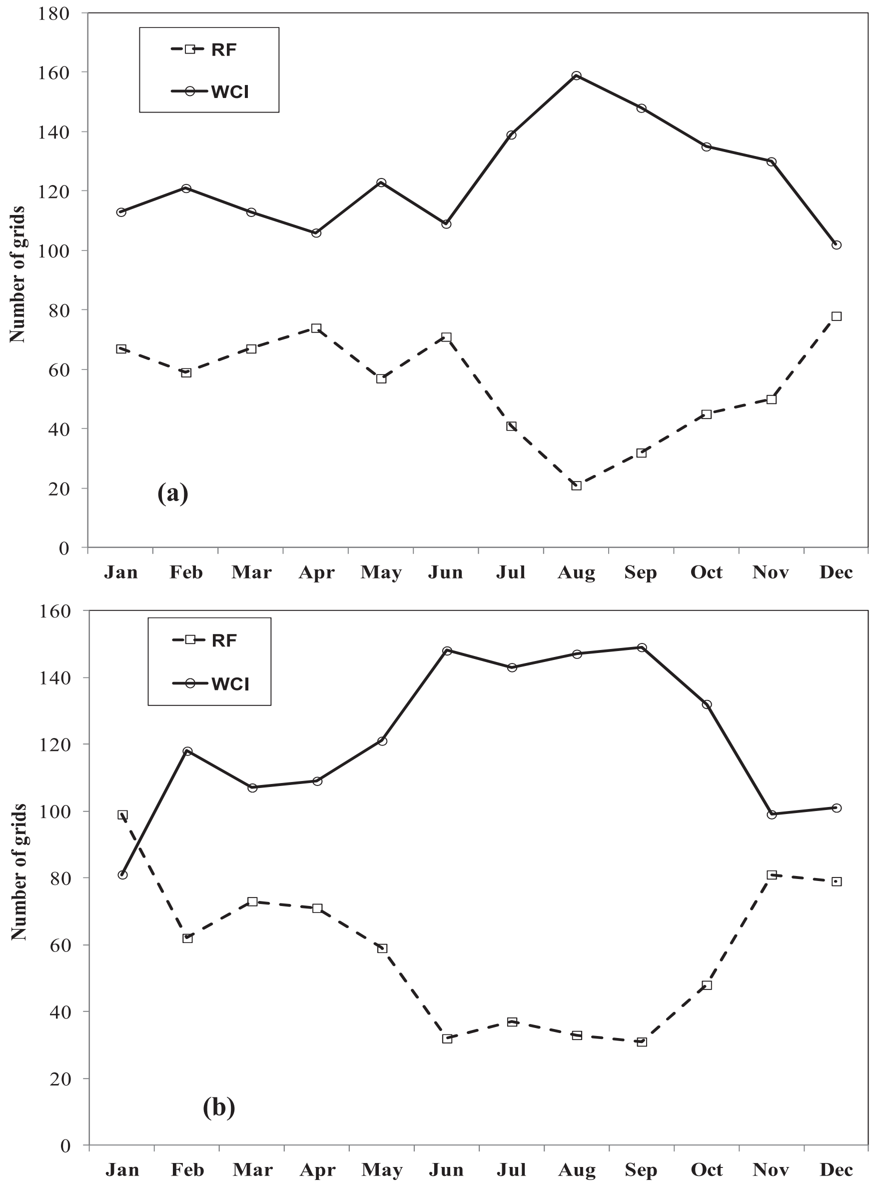

FIG. 8. Number of grid points under which reforecasts and WCI-based forecasts have the lowest RPS during (a) the first 15 days and (b) the second 15 days for each month over the continental United States.

It is clear that β values of disaggregated GCM forecasts are smaller than that of reforecasts under WCI's RPSS > 0. Given that reforecasts are primarily obtained from the numerical model of NCEP's Global Forecast System, it exhibits overconfidence in the reforecast ensemble with all ensemble members predicting similar outcomes under certain initial/boundary conditions. Thus, by combining better-dispersed disaggregated GCM forecasts with the overly dispersed reforecasts, the algorithm in Fig. 3 produces biweekly precipitation forecasts that are more skillful than the reforecasts. Based on Fig. 9, we argue that the basic reason why WCI-based biweekly forecasts perform better is due to their better dispersive characteristics. Even though the skill of disaggregated ECHAM4.5 is worse than that of the reforecast model, combining disaggregated climate forecasts with reforecasts result in biweekly forecasts that are more skillful than reforecasts. Given that the reforecasts are available only from 1979, we employed leave-five-out cross validation for evaluating the WCI forecasts. As the length of reforecasts increases, WCI-based forecasts could be evaluated based on split-sample

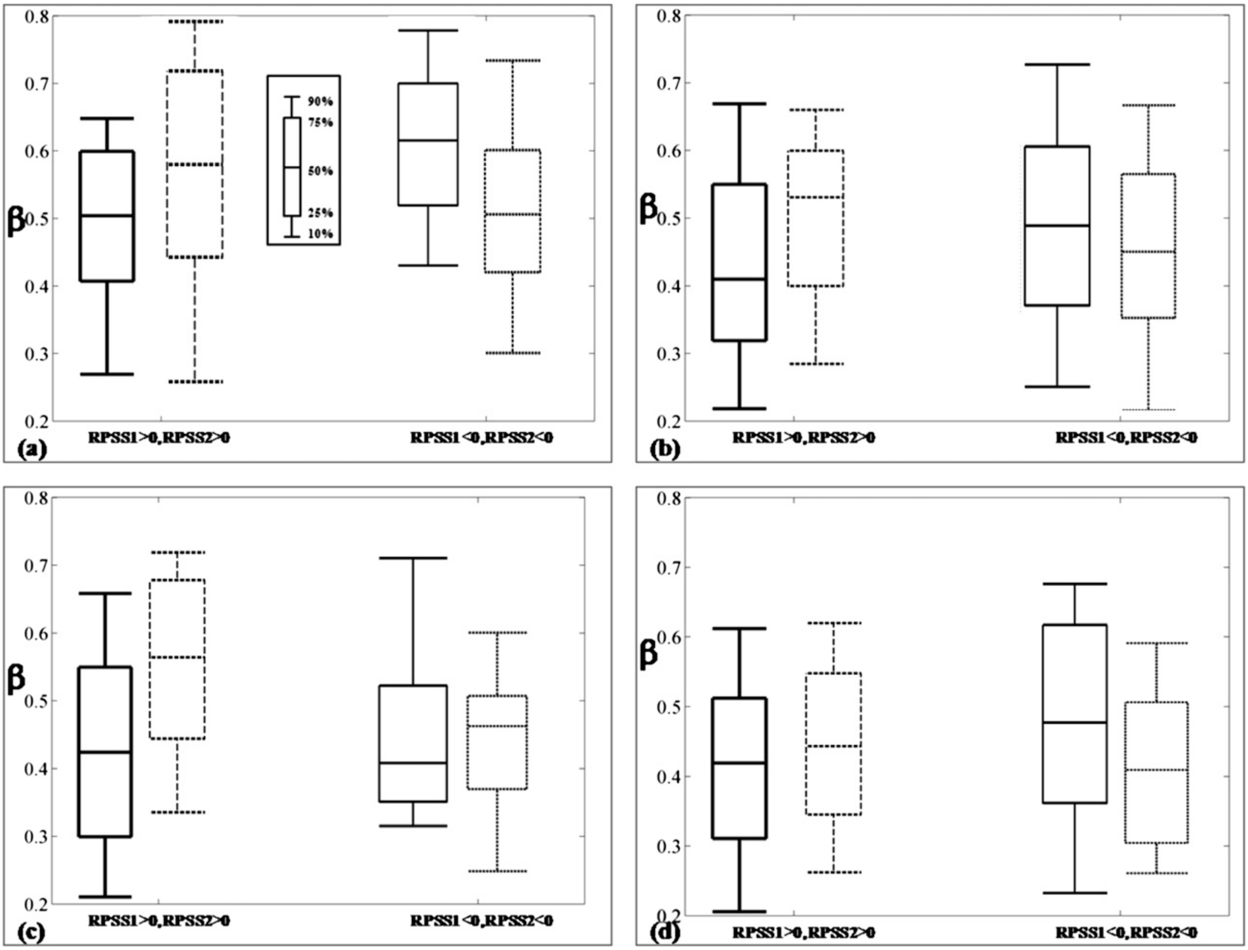

FIG. 9. Box plots of dispersion parameter, β, of the disaggregated GCM forecasts (solid box) and reforecast (dotted box) in the proposed combination scheme for different months: (a) January, (b) April, (c) July, and (d) October. Dispersion parameter β s are grouped in such a way with $RPSS_1$ and $RPSS_2$ of WCI being greater than and lesser than 0.

validation. However, implementation of the combination scheme in real-time forecast development faces one critical issue, which is the availability of climate forecasts at the first day of each month. Typically, climate forecasts are updated by the middle of the month. One way to overcome this is to consider the 2-month-ahead precipitation forecasts from ECHAM4.5 for disaggregation. Our future study will evaluate the combination scheme in developing real-time experimental biweekly precipitation forecasts as well as in utilizing them for developing 15-day-ahead real-time streamflow forecasts.

6. Summary and conclusions

A combination scheme is proposed to combine reforecasts from a numerical weather model and disaggregated climate forecasts from ECHAM4.5 for developing 15-day-ahead precipitation forecasts. The proposed scheme was evaluated by developing biweekly precipitation forecasts for each month in 180 grid points in the continental United States. Evaluation of this weather–climate information (WCI)-based biweekly forecasts under leave-five-out cross validation shows that the WCI-based forecasts perform better than the reforecasts in many grid points in the continental United States. Correlation between RPSS of the WCI-based forecasts and MSE of the disaggregated forecasts reveals that the lower the error in the disaggregated forecasts the better the performance of WCI forecasts. Weights analysis also shows that the combination scheme performs better by assigning higher weights to the better-performing model. Particularly, WCI-based biweekly forecasts perform better in the summer months during which the reforecasts have limited skill. Even though the disaggregated climate forecasts did not perform well, the primary reason why WCI-based forecasts perform better than reforecasts is due to their better dispersive characteristics. Analysis of the dispersion parameter, β, of reforecasts and disaggregated forecasts show that the disaggregated forecasts are better dispersed than the reforecasts particularly when the RPSS of the WCI-based forecasts are greater than zero. Thus, combining the disaggregated climate forecasts with the reforecasts result in improved 15-day-ahead precipitation forecasts. Our future research will evaluate the utility of these medium-range precipitation forecasts in developing 15-day-ahead streamflow forecasts for various selected river basins across the country.

Acknowledgments. The work is supported by Award 0756269 from the Environmental Sustainability Program of the National Science Foundation (NSF). Any opinions, findings, and conclusion expressed in this materials are those of the authors and do not necessarily reflect the views of NSF. The authors would also like to thank the two anonymous reviewers whose comments led to substantial improvements in the manuscript.

APPENDIX

Estimation of the Dispersion Parameter (β)

Weigel et al. (2008) demonstrated that multimodel forecasts can outperform the best individual model only if the individual models exhibit overconfidence, which is primarily due to the overly dispersed characteristics of the ensembles that represent the conditional probability distribution functions. We employ Weigel et al.'s (2008) approach to estimate the dispersion parameter, β, based on the ensembles of the reforecasts and the disaggregated ECHAM4.5 forecasts.

In this study, the number of ensemble members (N_m with m denoting the models) represented by reforecasts and disaggregated ECHAM4.5 forecasts are 15 and 100, respectively. To begin with, both the candidate forecasts ensemble members are standardized for each biweekly period, j, based on their respective climatological mean (μ_j^m) and standard deviation (σ_j^m). Based on the standardized ensemble forecasts for each model m, we calculated the ensemble mean, $\mathrm{ep}_{j,t}^m$, and the ensemble standard deviation, $\sigma_{j,t}^m$, for each year t. Given the ensemble mean, $\mathrm{ep}_{j,t}^m$, and the ensemble variance, $\widehat{(\sigma_{j,t}^m)^2}$, of the standardized forecasts from both models, we estimated $\widehat{\beta_j^m}$ for each model m over the forecasting biweekly period, j, using Eq. (A1):

$$\widehat{\beta_j^m} = \sqrt{\widehat{\mathrm{var}}(\widehat{\mathrm{ep}_{j,t}^m}) - \frac{1}{N_m}\left(\widehat{\sigma_{j,t}^{m^2}}\right) - (\widehat{\alpha_j^m})^2}, \quad \text{(A1)}$$

where $\widehat{\mathrm{var}}(\widehat{\mathrm{ep}_{j,t}^m})$ denotes the variance of the ensemble mean ($\mathrm{ep}_{j,t}^m$) over the period 1979–2004, α_j^m denotes the correlation between the ensemble mean ($\mathrm{ep}_{j,t}^m$) of the standardized forecasts and the observed precipitation over the period 1979–2004, $\overline{(\sigma_j^m)^2}$ represents the average of the ensemble variance $[\widehat{(\sigma_{j,t}^m)^2}]$ over the period 1979–2004, and N_m denotes the number of ensemble members in each candidate forecasts. For additional details, see Weigel et al. (2008).

REFERENCES

Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.,* **115,** 1083–1126.

——, S. J. Mason, L. Goddard, D. G. DeWitt, and S. E. Zebiak, 2003: Multimodel ensembling in seasonal climate forecasting at IRI. *Bull. Amer. Meteor. Soc.,* **84,** 1783–1796.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.,* **129,** 420–436.

Clark, M. P., S. Gangopadhyay, L. Hay, B. Rajagopalan, and R. Wilby, 2004: The Schaake shuffle: A method for reconstructing space–time variability in forecasted precipitation and temperature fields. *J. Hydrometeor.,* **5,** 243–262.

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: Improved prediction of winter precipitation and temperature over the continental United States: Role of ENSO state in developing multimodel combinations. *Mon. Wea. Rev.,* **138,** 2447–2468.

——, ——, and S. Ghosh, 2008: Multimodel ensembles of streamflow forecasts: Role of predictor state in developing optimal combination. *Water Resour. Res.,* **44,** W09404, doi:10.1029/2006WR005855.

Gneiting, T., and A. Raftery, 2005: Weather forecasting with ensemble methods. *Science,* **310,** 248–249, doi:10.1126/science.1115255.

Grygier, J. C., and J. R. Stedinger, 1990: SPIGOT, a synthetic streamflow generation software package, technical description, version 2.5. Cornell University School of Civil and Environmental Engineering.

Hamill, T. M., and C. Snyder, 2002: Using improved background-error covariances from an ensemble Kalman filter for adaptive observations. *Mon. Wea. Rev.,* **130,** 1552–1572.

——, J. S. Whitaker, and X. Wei, 2004: Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.,* **132,** 1434–1447.

——, ——, and S. L. Mullen, 2006: Reforecasts, an important dataset for improving weather predictions. *Bull. Amer. Meteor. Soc.,* **87,** 33–46.

——, R. Hagedorn, and J. S. Whitaker, 2008: Probabilistic forecast calibration using ECMWF and GFS ensemble reforecast. Part II: Precipitation. *Mon. Wea. Rev.,* **136,** 2620–2632.

Higgins, R. W., J. E. Janowiak, and Y.-P. Yao, 1996: A gridded hourly precipitation data base for the United States (1963-1993). NCEP/Climate Prediction Center ATLAS 1, 47 pp.

Johnson, C., and R. Swinbank, 2009: Medium-range multimodel ensemble combination and calibration. *Quart. J. Roy. Meteor. Soc.,* **135,** 777–794.

Kalman, R. E., 1960: A new approach to linear filter and prediction problems. *Trans. ASME,* **82D,** 35–45.

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

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

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

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

Lall, U., and A. Sharma, 1996: A nearest neighbor bootstrap for resampling hydrological time series. *Water Resour. Res.,* **32,** 679–693.

Lee, T., J. D. Salas, and J. Prairie, 2010: An enhanced nonparametric streamflow disaggregation model with genetic algorithm. *Water Resour. Res.,* **46,** W08545, doi:10.1029/2009WR007761.

Li, S., and L Goddard, 2005: Retrospective forecasts with ECHAM4.5 AGCM. IRI Tech. Rep. 05–02, 16 pp.

Lorenz, E. N., 1963: Deterministic non-periodic flows. *J. Atmos. Sci.,* **20,** 130–141.

Prairie, J., B. Rajagopalan, U. Lall, and T. Fulp, 2007: A stochastic nonparametric technique for space-time disaggregation of streamflows. *Water Resour. Res.,* **43,** W03432, doi:10.1029/2005WR004721.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski, 2005: Using Bayesian averaging to calibrate forecast ensembles. *Mon. Wea. Rev.,* **133,** 1155–1174.

Rajagopalan, B., and U. Lall, 1999: A *k*-nearest-neighbor simulator for daily precipitation and other weather variables. *Water Resour. Res.,* **35,** 3089–3101, doi:10.1029/1999WR900028.

——, ——, and S. E. Zebiak, 2002: Categorical climate forecasts through regularization and optimal combination of multiple GCM ensembles. *Mon. Wea. Rev.,* **130,** 1792–1811.

Robertson, A. W., S. Kirshner, and P. Smyth, 2004: Downscaling of daily rainfall occurrence over northeast Brazil using a hidden Markov model. *J. Climate,* **17,** 4407–4424.

Roeckner, E., and Coauthors, 1996: The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate. Max-Planck Institute for Meteorology Tech. Rep. 218, 90 pp.

Sankarasubramanian, A., U. Lall, and N. Devineni, 2008: Role of retrospective forecasts of GCM forced with persisted SST anomalies in operational streamflow forecasts development. *J. Hydrometeor.,* **9,** 212–227.

——, ——, ——, and S. Espunevea, 2009a: The role of monthly updated climate forecasts in improving intraseasonal water allocation. *J. Appl. Meteor. Climatol.,* **48,** 1464–1482.

——, ——, F. D. Souza Filho, and A. Sharma, 2009b: Improved water allocation utilizing probabilistic climate forecasts: Short-term water contracts in a risk management framework. *Water Resour. Res.,* **45,** W11409, doi:10.1029/2009WR007821.

Sivakumar, B., W. W. Wallender, C. E. Puente, and M. N. Islam, 2004: Streamflow disaggregation: A nonlinear deterministic approach. *Nonlinear Processes Geophys.,* **11,** 383–392.

Tarboton, D., A. Sharma, and U. Lall, 1998: Disaggregation procedures for stochastic hydrology based on nonparametric density estimation. *Water Resour. Res.,* **34,** 107–119, doi:10.1029/97WR02429.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.,* **125,** 3297–3319.

Valencia, D., and J. C. Schaake Jr., 1973: Disaggregation processes in stochastic hydrology. *Water Resour. Res.,* **9,** 580–585.

Weigel, A. P., and N. E. Bowler, 2009: Comment on 'Can multimodel combination really enhance the prediction of probabilistic ensemble forecasts?' *Quart. J. Roy. Meteor. Soc.,* **135,** 535–539.

——, M. A. Linier, and C. Appenzeller, 2008: Can multi-model combination really enhance the prediction of probabilistic ensemble forecasts? *Quart. J. Roy. Meteor. Soc.,* **134,** 241–260.

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

——, X. Wei, and F. Vitart, 2006: Improving week-2 forecasts with multimodel reforecast ensembles. *Mon. Wea. Rev.,* **134,** 2279–2284.

Wilks, S. D., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Elsevier, 627 pp.

——, and M. T. Hamill, 2007: Comparison of ensemble-MOS methods using GFS reforecasts. *Mon. Wea. Rev.,* **135,** 2379–2390.