Decomposition of Sources of Errors in Monthly to Seasonal Streamflow Forecasts in a Rainfall–Runoff Regime

TUSHAR SINHA, A. SANKARASUBRAMANIAN, AND AMIRHOSSEIN MAZROOEI

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

(Manuscript received 25 August 2013, in final form 9 July 2014)

ABSTRACT

Despite considerable progress in developing real-time climate forecasts, most studies have evaluated the potential in seasonal streamflow forecasting based on ensemble streamflow prediction (ESP) methods, utilizing only climatological forcings while ignoring general circulation model (GCM)-based climate forecasts. The primary limitation in using GCM forecasts is their coarse resolution, which requires spatiotemporal downscaling to implement land surface models. Consequently, multiple sources of errors are introduced in developing real-time streamflow forecasts utilizing GCM forecasts. A set of error decomposition metrics is provided to address the following questions: 1) How are errors in monthly streamflow forecasts attributed to various sources such as temporal disaggregation, spatial downscaling, imprecise initial hydrologic conditions (IHCs), climatological forcings, and imprecise forecasts? and 2) How do these errors propagate with lead time over different seasons? A calibrated Variable Infiltration Capacity model is used over the Apalachicola River at Chattahoochee in the southeastern United States. The model is forced with a combination of daily precipitation forcings (temporally disaggregated observed precipitation, spatially downscaled and temporally disaggregated observed precipitation, ESP, ECHAM4.5 forecasts, and observed) and IHCs [simulated and climatological ensemble reverse ESP (RESP)] but with observed air temperature and wind speed at $^1/_8°$ resolution. Then, errors in forecasting monthly streamflow at up to a 3-month lead time are decomposed by comparing the forecasted streamflow to simulated streamflow under observed forcings. Results indicate that the errors due to temporal disaggregation are much higher than the spatial downscaling errors. During winter and early spring, the increasing order of errors at a 1-month lead time is spatial downscaling, model, temporal disaggregation, RESP, large-scale precipitation forecasts, and ESP.

1. Introduction

Over the last decade, considerable progress has been made in the ability to forecast seasonal streamflow through better understanding of climatic teleconnections (e.g., ENSO) as well as through the development of subgrid-scale land surface models (LSMs) that capture land–atmosphere interactions (Koster and Suarez 1995; Betts et al. 1997; Hamlet et al. 2002; Maurer et al. 2002; Mahanama et al. 2012). Nevertheless, the skill of climate forecasts varies significantly for different lead times, seasons, and geographic regions, restricting the widespread application of streamflow forecasts for water management (Stern and Easterling 1999; Pagano et al. 2001; Hartmann et al. 2002).

One of the key challenges in developing streamflow forecasts utilizing climate forecasts is the spatiotemporal scale mismatch between climate models and land surface models (Wood et al. 2002; Wood and Lettenmaier 2006; Luo and Wood 2008; Yuan et al. 2011). For instance, in order to utilize climate forecasts from general circulation models (GCMs), both spatial downscaling and temporal disaggregation (from monthly to daily) of climate forecasts are required to implement LSMs. Several strides have been made to bridge the mismatch in spatial scale by pursuing statistical downscaling (Wood et al. 2004; Maurer and Hidalgo 2008) or dynamic downscaling using regional climate models (RCMs) to obtain air temperature and precipitation at the LSMs scale (Wilby et al. 2000; Wood et al. 2002, 2004; Leung et al. 2003; 2004; Cocke et al. 2007; Yuan et al. 2012). Various studies show that dynamic downscaling approaches do not add significant skill in comparison to statistical approaches in developing downscaled products (Murphy

Corresponding author address: Tushar Sinha, 2501 Stinson Dr., Box 7908, Raleigh, NC 27695-7908.
E-mail: tsinha@ncsu.edu

DOI: 10.1175/JHM-D-13-0155.1

1999; Wilby et al. 2000; Caldwell 2010). Halmstad et al. (2013) suggested that not all the RCMs outperform their host GCMs in terms of correlation skill in predicting extreme precipitation. However, Di Luca et al. (2012) showed noticeable improvements in RCM prediction than their host GCMs because of the finer temporal scale of RCMs, but the authors also pointed out the need for further research to understand key aspects of climate predictions where RCMs add value in comparison to the GCM projections. In addition, the primary limitations in using RCMs to develop streamflow forecasts are the lack of available retrospective GCM forecast runs at a daily time scale as well as the enormous time required to set up RCMs over a given domain. Furthermore, to force the statistically downscaled climate forcings (available at 1/8°), which are typically available at a monthly time scale, temporal disaggregation is pursued to obtain hourly to daily time scale inputs for LSMs. Nevertheless, both spatial downscaling and temporal disaggregation techniques (e.g., Hayhoe et al. 2004; Prairie et al. 2007) introduce errors in the downscaled and disaggregated climate forcings, which are also translated in the developed streamflow forecasts.

In addition to climate input uncertainty, other sources of uncertainties, such as initial hydrologic conditions (IHC), parameter estimation, and the model itself, contribute to errors in streamflow forecasts from a given LSM. Such uncertainties can be reduced by 1) continuously updating IHCs (Sankarasubramanian et al. 2008; Sinha and Sankarasubramanian 2013), 2) combining parameter estimation and data assimilation in a Bayesian context (Duan et al. 2007; Ajami et al. 2007), and 3) using multimodel combinations (Li and Sankarasubramanian 2012; Devineni et al. 2008). The goals of this paper are to understand and quantify the errors that arise in the development of streamflow forecasts and to compare them with errors arising from large-scale climate forcings and IHCs for a basin dominated by the rainfall–runoff regime.

Numerous studies have used IHCs and climatological forcings to quantify the relative role of one over the other in snow-dominated regimes (e.g., Maurer et al. 2004; Mahanama et al. 2012; Shukla and Lettenmaier 2011). These studies mostly concur on the dominant role of IHCs over forecasting skills for lead times up to 1.5–3 months (Maurer and Lettenmaier 2003; Shukla and Lettenmaier 2011) in the snowmelt-dominated basins. It is natural to expect IHCs to be critical in the snow-dominated regimes (Maurer et al. 2004; Koster et al. 2010), whereas in the rainfall–runoff regimes, streamflow variability is primarily governed by climate forcings. Shukla and Lettenmaier (2011) reported that climatological forcings play a more important role than soil moisture in estimating streamflow beyond a 1-month lead time over the southeastern United States. Furthermore, very few land surface modeling studies have utilized retrospective climate forecasts to quantify the role of climate forecasts and IHCs (Luo and Wood 2008; Luo et al. 2007; Yuan et al. 2011). For example, Luo et al. (2007) reported that climate forecasts play a dominant role compared to IHCs in predicting summer flows for a lead time greater than 1 month. Li et al. (2009) found that IHCs are only dominant up to a 1-month lead time, beyond which climate forecasts are important for developing streamflow forecasts during summer months. Yuan et al. (2013) indicated that the forecasts from the Climate Forecast System, version 2 (CFSv2), add limited skill against climatology beyond a 1-month lead time.

The primary limitation in using real-time climate forecasts from GCMs is their coarse resolution, which requires spatial and temporal downscaling before they can be used to force land surface models. Consequently, multiple sources of errors are introduced in developing real-time streamflow forecasts utilizing climate information. This study explicitly quantifies and compares different sources of errors that arise from spatial downscaling, temporal disaggregation, imprecise IHCs, and GCM forecasts in the development of streamflow forecasts. For this purpose, we build on our previous study that considered monthly updated climate forecasts for developing monthly streamflow forecasts over the Apalachicola River at Chattahoochee, Florida (Sinha and Sankarasubramanian 2013). We systematically perform a retrospective analysis to decompose different sources of errors in monthly to seasonal streamflow forecasts as a function of lead time over the selected basin. Based on this analysis, we address the following science questions:

1) How do we attribute the errors in monthly streamflow forecasts to various sources such as (i) temporal disaggregation, (ii) spatial downscaling, (iii) imprecise IHCs, (iv) climatological forcings, (v) imprecise climate forecasts, and (vi) model error?
2) How do these errors propagate as a function of lead time under different seasons?

The manuscript is organized as follows. Section 2 details the experimental setup and decomposition of error sources under different streamflow forecasting schemes. Section 3 provides details on the results and analyses by application for the Apalachicola River at Chattahoochee. Section 4 presents the summary and findings from the study, and finally, section 5 describes the overall conclusions.

2. Study area and experimental setup

To address the science questions discussed in the previous section, we primarily consider the calibrated

Variable Infiltration Capacity (VIC) land surface model that was employed in our previous study (Sinha and Sankarasubramanian 2013) focusing on the Apalachicola River at Chattahoochee, Florida, in the southeastern United States. The VIC model was implemented with observed 1/8° daily forcings to estimate the reference (i.e., simulated) monthly streamflow for the period 1981–2010. The VIC model was then forced with different precipitation forcings under updated IHCs, prior to the forecasting period, to estimate the relative root-mean-square errors (RMSEs) in monthly streamflow due to 1) temporal disaggregation, 2) spatial downscaling, 3) imprecise IHCs [reverse ensemble streamflow prediction (RESP)], 4) imprecise forcings [ensemble streamflow prediction (ESP)], and 5) ECHAM4.5 precipitation forecasts (candidate forecasts). Based on that, error propagation under 1–3-month lead time is analyzed over different months for the five different forecasting schemes considered in this study.

a. Study area and data description

The Apalachicola River at Chattahoochee, Florida, is composed of the Apalachicola, Chattahoochee, and Flint (ACF) Rivers and has a total drainage area of 44 032 km^2. The basin receives uniform precipitation throughout the year, resulting in significant runoff in all months. For additional details about the basin, see Sinha and Sankarasubramanian (2013). Daily time series of streamflow data during 1981–2010 were obtained from the U.S. Geological Survey (USGS) stream gauge station (02358000) available at the Apalachicola River at Chattahoochee. This site is included in the Hydro-Climatic Data Network (HCDN) database, and it is minimally affected by reservoir operations (Slack et al. 1993). Daily meteorological data including precipitation, maximum and minimum air temperature, and wind speed were obtained from Maurer et al. (2002) and are available at 1/8° spatial resolution over the conterminous United States for calibrating the VIC model under observed meteorological forcings.

Monthly updated precipitation forecasts from the ECHAM4.5 GCM, available at ~2.8° spatial resolution, were obtained from the International Research Institute of Climate and Society (IRI) data library (Li and Goddard 2005) from 1981–2010 for up to a 3-month lead time. A 3-month-lead forecast issued at the beginning of January denotes the forecasts for the month of March. These monthly ECHAM4.5 forecasts are available as 24 ensemble members, and we considered the ensemble mean of precipitation forecasts for statistical downscaling to 1/8° spatial resolution. Overall, seven adjacent ECHAM4.5 grid points that exhibited statistically significant correlation with spatially averaged precipitation over the ACF basin were selected for downscaling and disaggregation [see details in Sinha and Sankarasubramanian (2013)].

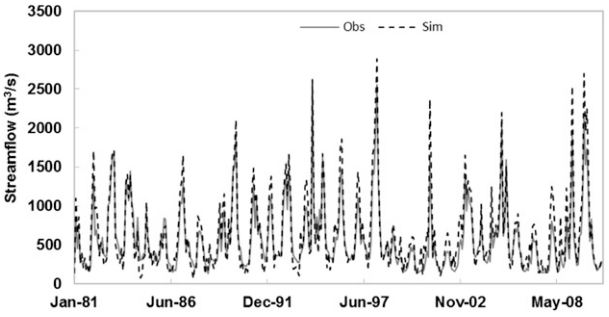

FIG. 1. Comparison of monthly USGS-observed and VIC model–simulated streamflow when forced under observed 1/8° meteorological forcings (Maurer et al. 2002) during the validation period of 1981–2010. The model was calibrated over 1957–80 (Sinha and Sankarasubramanian 2013).

b. VIC model

We used a semidistributed macroscale VIC land surface model (Liang et al. 1994, 1996; Cherkauer and Lettenmaier 2003) to analyze different sources of errors in monthly streamflow forecasting. The VIC model estimates energy and water balance fluxes on a gridcell by gridcell basis. A stand-alone routing model (Lohmann et al. 1998a,b) is implemented to route surface and subsurface flows to the basin outlet. Details on soil and vegetation parameters for the Apalachicola River at the Chattahoochee basin are available in Sinha and Sankarasubramanian (2013). The VIC model was calibrated during 1957–80, and the overall Nash–Sutcliffe efficiency during the analysis period of 1981–2010 is 0.81 on a monthly basis (Sinha and Sankarasubramanian 2013). Figure 1 shows a comparison of monthly time series of observed streamflow and the VIC model–simulated streamflow under observed forcings (Maurer et al. 2002) during the 1981–2010 period. Percentage bias corrections on the mean monthly flows were applied to all the VIC-simulated schemes during the period 1981–2010, when the comparisons of different forecasting schemes were made with respect to the USGS monthly observed flows.

c. Experimental design for error decomposition in streamflow forecasts

Figure 2 illustrates the overall experimental setup for analyzing different sources of errors arising in streamflow forecasts. A combination of different daily precipitation forcings and IHCs (either simulated or climatological ensemble ESP) were used to implement the VIC model at 1/8° scale during 1981–2010. IHCs were updated across all the schemes based on the last day of

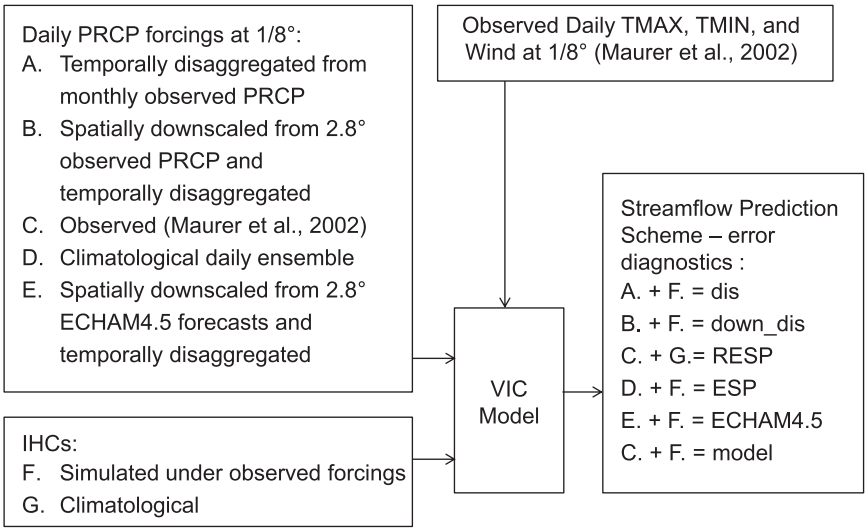

FIG. 2. Schematic showing the experimental design indicating different streamflow prediction schemes based on combinations of precipitation forcings (A–E) and IHCs (F and G) for implementing the VIC model. TMAX, TMIN, PRCP, and WIND represent max air temperature, min air temperature, precipitation, and wind speed, respectively; all other abbreviations are defined in the text.

the month prior to the forecasting period. For example, in order to forecast streamflow for January 1981, IHCs were updated based on 31 December 1980. Different daily precipitation forcings used in this study are as follows (Fig. 2):

1) temporally disaggregated daily precipitation time series from observed monthly data at 1/8° based on the Prairie et al. (2007) approach;
2) spatially downscaled and disaggregated daily precipitation that was obtained by downscaling observed monthly precipitation from 2.8° (obtained by upscaling all the relevant 1/8° grid points to 2.8° spatial scale) to 1/8° and then temporally disaggregating to daily scale using Prairie et al. (2007);
3) observed gridded precipitation (Maurer et al. 2002) at 1/8°;
4) climatological daily ensemble; and
5) spatially downscaled and temporally disaggregated ECHAM4.5 precipitation forecasts.

Other forcing variables such as observed daily maximum and minimum air temperature and wind speed were kept as the observed daily time series from Maurer et al. (2002) to obtain streamflow forecasts under different schemes. For each month, the performance of different streamflow forecasting schemes were compared with two reference streamflow values, the first one being the VIC model–simulated streamflow under observed forcings and the second one being the USGS-observed streamflow. The performance was evaluated based on the RMSE for the period 1981–2010 over 1–3-months lead time. Comparing the streamflow prediction scheme with the VIC-simulated streamflow Q_{it}^{sim} [where t denotes the month (1–12) and i denotes the year (1–30)] ignores the errors due to the model while comparison of the scheme with the observed streamflow Q_{it}^{obs} considers the error due to the VIC model.

d. Error decomposition metrics

Details of each forecasting scheme and its associated RMSE calculations, which are based on a given reference streamflow scheme Q_{it}^{ref} (where Q_{it}^{ref} denotes either Q_{it}^{sim} or Q_{it}^{obs}), are discussed next.

1) ERRORS DUE TO TEMPORAL DISAGGREGATION

To determine errors in monthly streamflow forecasts due to the disaggregation scheme alone, daily observed precipitation at 1/8° from Maurer et al. (2002) was first aggregated to a monthly time step and then statistically disaggregated using the k–nearest neighbor (k-NN) approach to estimate daily disaggregated time series at the same grid point. For each month, the k-NNs were identified by estimating the distance between the predicted time series and the historical time series (Maurer et al. 2002) over the period 1981–2010 by leaving out the year from the historical series for which the predictions are to be made. For example, to identify the nearest neighbors for January 1982, monthly precipitation for the remaining 29 years during 1981–2010 was included in

the historical pool, excluding January 1982. The daily values corresponding to the identified monthly nearest neighbors were resampled using the Lall and Sharma kernel (Lall and Sharma 1996), which assigns higher weights to the nearest neighbors to obtain a single daily time series [see details in Sinha and Sankarasubramanian (2013)]. The *k*-NN disaggregation scheme preserves the monthly precipitation totals during the resampling process. This daily disaggregated precipitation time series along with daily updated IHCs, prior to the forecasting period, was used to implement the VIC model. The estimated streamflow under this scheme is the temporally disaggregated streamflow forecasting scheme, abbreviated as dis, for the reference streamflow. Thus, the RMSE due to disaggregation alone in predicting streamflow for a given month t under a specified lead time k could be written as

$$\mathrm{RMSE}_{tk}^{\mathrm{dis}} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}(Q_{it}^{\mathrm{ref}} - Q_{itk}^{\mathrm{dis}})^2\right]}, \qquad (1)$$

where Q_{it}^{ref} is the reference streamflow (either Q_{it}^{sim} or Q_{it}^{obs}), k is the lead time in months ranging from 1 to 3, n denotes number of years (30), and Q_{itk}^{dis} is the VIC-simulated streamflow when the model was forced with updated IHCs and disaggregated observed precipitation to estimate streamflow on a monthly basis. The *k*-NN disaggregation scheme was applied to all the forecasting schemes considered in this study, except for the ESP and the RESP schemes.

2) Errors due to spatial downscaling and disaggregation

Under this scheme, we quantify the error due to both spatial downscaling and disaggregation. For this purpose, we first upscaled observed precipitation at 1/8° grid points over the seven 2.8° grid points covering the study area. The selection of the seven 2.8° grid points is chosen since these grids are consistent with the extent and the spatial resolution of the ECHAM4.5 GCM precipitation forecasts. Thus, these seven 2.8° grid points were considered as predictors for developing the principal component regression (PCR). Since the precipitation forecasts over the selected seven grids are correlated, we used principal components of the predictors in developing the regression model. The observed 1/8° precipitation (Maurer et al. 2002) over the 251 grid points, which comprises of the study area, was considered as the predictand. For each 1/8° grid, a different PCR model was developed on a monthly basis under the leave-one-out cross-validation mode over 1981–2010. Using this downscaled 1/8° precipitation, we again disaggregated the precipitation into a daily precipitation time series using the same disaggregation method of Prairie et al. (2007), as described in the previous section. Thus, the developed precipitation forcings account for both downscaled and disaggregated errors. Based on this, we estimated the combined error due to both spatial downscaling and disaggregation with respect to a given reference streamflow scheme (Q_{it}^{sim} or Q_{it}^{obs}). Thus, streamflow prediction errors were estimated as follows (down_dis):

$$\mathrm{RMSE}_{tk}^{\mathrm{down\_dis}} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}(Q_{it}^{\mathrm{ref}} - Q_{itk}^{\mathrm{down\_dis}})^2\right]}. \qquad (2)$$

Based on this, the error due to spatial downscaling alone (down) is estimated as follows:

$$\mathrm{RMSE}_{tk}^{\mathrm{down}} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}(Q_{itk}^{\mathrm{down\_dis}} - Q_{itk}^{\mathrm{dis}})^2\right]}. \qquad (3)$$

3) Errors due to imprecise IHCs

Errors due to imprecise IHCs were estimated by implementing the VIC model with observed precipitation forcings along with the ensembles of IHCs. Observed precipitation forcings (Maurer et al. 2002) were used with climatological IHC's for a 3-month lead time during 1981–2010 using 29 years of updated IHCs, leaving out the year during which streamflow forecasts are developed. The average streamflow ensemble is estimated for each 3-month forecasting period and the RMSE under RESP is expressed as

$$\mathrm{RMSE}_{tk}^{\mathrm{RESP}} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}\left(Q_{it}^{\mathrm{ref}} - \overline{Q}_{itk}^{\mathrm{RESP}}\right)^2\right]}, \qquad (4)$$

where $\overline{Q}_{itk}^{\mathrm{RESP}}$ is the average VIC-simulated streamflow under the RESP with a lead time k.

4) Errors due to climatological forcings

Another streamflow prediction scheme is based on utilizing climatological precipitation ensembles. Climatological daily precipitation was estimated using the ESP approach (Day 1985; Franz et al. 2003). Under this approach, the daily precipitation, maximum and minimum air temperature, and wind speed time series for 3-months lead time during 1981–2010 were selected from the observed daily gridded data (Maurer et al. 2002) for 29 years, leaving out the year of the forecasting period to drive the VIC model. Finally, the mean ESP was estimated and the RMSE due to climatological forcings $\mathrm{RMSE}_{tk}^{\mathrm{ESP}}$ is expressed as

$$\mathrm{RMSE}_{tk}^{\mathrm{ESP}} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}\left(Q_{it}^{\mathrm{ref}} - \overline{Q}_{itk}^{\mathrm{ESP}}\right)^2\right]}, \quad (5)$$

where $\overline{Q}_{itk}^{\mathrm{ESP}}$ is average VIC-simulated streamflow using IHCs from 29 ensemble members in year i at lead time k.

5) Errors due to imprecise ECHAM4.5 climate forecasts

Given that the ECHAM4.5 precipitation forecasts are available at 2.8° spatial scale and at a monthly time step, both spatial downscaling and temporal disaggregation are required prior to implementing the VIC model. Therefore, the streamflow forecasts based on ECHAM4.5 precipitation is composed of errors due to three sources: 1) large-scale precipitation forecasts alone (i.e., errors due to the ECHAM4.5 forecasts, excluding errors due to spatial downscaling and temporal disaggregation), 2) spatial downscaling, and 3) temporal disaggregation. Monthly updated precipitation forecasts at 3-month-lead from the ECHAM4.5 GCM were first spatially downscaled from 2.8° to 1/8° scale using the PCR methods described earlier. Then, the downscaled monthly time series was disaggregated to obtain a daily time series at 1/8° using the k-NN approach [for additional details, see Sinha and Sankarasubramanian (2013)]. Finally, the ECHAM4.5 daily precipitation forecasts were used with updated IHCs to estimate the 3-month-lead streamflow forecasts (forecasts). This procedure was repeated for each month using the updated precipitation forecasts to obtain the 3-month-ahead streamflow forecasts over 1981–2010. The RMSE due to the imprecise ECHAM4.5 forecasts is estimated as

$$\mathrm{RMSE}_{tk}^{\mathrm{ECHAM4.5}} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}\left(Q_{it}^{\mathrm{ref}} - Q_{itk}^{\mathrm{ECHAM4.5}}\right)^2\right]}, \quad (6)$$

where $Q_{itk}^{\mathrm{ECHAM4.5}}$ is the VIC-simulated flow when the model was forced with the spatially downscaled and temporally disaggregated ECHAM4.5 monthly updated precipitation forecasts. Based on this, errors due to the imprecise large-scale forecasts alone $\mathrm{RMSE}_{tk}^{\mathrm{fcst}}$, which excludes the errors due to spatial downscaling and temporal disaggregation, are obtained by

$$\mathrm{RMSE}_{tk}^{\mathrm{fcst}} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}\left(Q_{itk}^{\mathrm{ECHAM4.5}} - Q_{itk}^{\mathrm{down\_dis}}\right)^2\right]}. \quad (7)$$

To compare the performance of different schemes across various months, we computed relative RMSE (R-RMSE) by taking the ratio of monthly RMSE of a forecasting scheme to that of the mean monthly reference streamflow in that month.

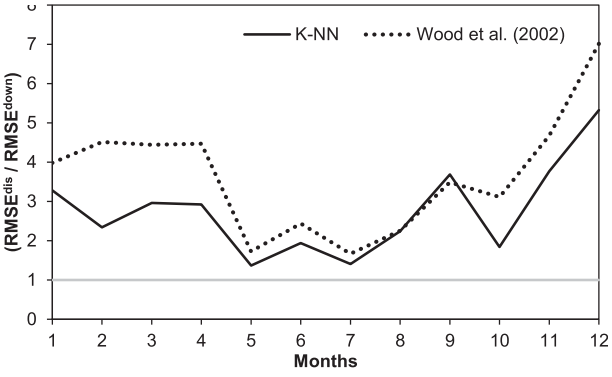

FIG. 3. Ratio of $\mathrm{RMSE}^{\mathrm{dis}}$ to $\mathrm{RMSE}^{\mathrm{down}}$ at a 1-month lead time under two different temporal disaggregation schemes: k-NN and Wood et al. (2002).

3. Results and analysis

Different streamflow prediction schemes outlined in the previous section were run for the period 1981–2010 and the RMSE for up to a 3-month lead time under both the reference streamflows (Q_{it}^{sim} or Q_{it}^{obs}) was computed for each month of forecast.

a. Assuming no model errors (reference streamflow Q_{it}^{sim})

In this section, we first summarize the performance under the reference streamflow being the VIC model–simulated streamflow, which assumes no model error.

1) Errors due to temporal disaggregation and spatial downscaling

Figure 3 shows the relative magnitude of errors due to temporal disaggregation (dis) and spatial downscaling (down) by estimating the ratio of the RMSE due to disaggregation [Eq. (1)] and the RMSE due to spatial downscaling [Eq. (3)] at a 1-month lead time ($k = 1$). Apart from the k-NN disaggregation scheme, we also included the VIC model–estimated streamflow when the model was forced with another commonly used disaggregation scheme for precipitation suggested by Wood et al. (2002). Under this scheme, for a given month, a year is selected randomly from the historical observations. The daily values of the selected year are rescaled based on the monthly ratio of precipitation in the forecasting month to the monthly observed precipitation [see Wood et al. (2002) for details]. Figure 3 indicates that the RMSE ratio of dis to down is greater than 1 in all the months. This implies that the errors due to temporal disaggregation are much larger than the errors due to spatial downscaling in all the months. Interestingly, the RMSE ratio of dis to down is smallest in the months of May–July, indicating that errors due to disaggregation and downscaling are similar in magnitude.

These months receive most of the precipitation in the form of convective storms. Such storms are associated with local processes, resulting in the high downscaling error for these months (Wang et al. 2013). Therefore, errors due to downscaling are expected to be relatively higher during late spring and early summer in comparison to the other seasons. On comparing the performance of the VIC model streamflow under both the disaggregation schemes, the *k*-NN method performs slightly better than the Wood et al. (2002) approach during most of the months, except September (Fig. 3). The most important point from Fig. 3 is that under both schemes, the error due to disaggregation is higher than the error arising from downscaling alone. Given the slight improved performance of the *k*-NN scheme, we considered the *k*-NN scheme for presenting further results on temporal disaggregation.

To further understand why the disaggregation error is higher than the downscaling error, we compared the flows obtained by forcing the VIC model with the disaggregated time series (Q_{itk}^{dis}) to the VIC-simulated flows (Q_{it}^{sim}; Fig. 4a) as well as to the flows obtained because of combined downscaling and disaggregation error ($Q_{itk}^{\text{down\_dis}}$; Fig. 4b) for July at a 1-month lead time ($k = 1$). Figure 4a clearly shows that the residual errors due to temporal disaggregation are relatively higher during the high flow years. Further, upon comparing $Q_{itk}^{\text{down\_dis}}$ with Q_{itk}^{dis}, we understand the predicted values are almost similar in all the years, except in the high flow year of July 1994 (as shown as a black circle in Fig. 4). This suggests that the major component of error in the downscaling and disaggregation primarily arises from the inability of disaggregation scheme in predicting the high flows. Thus, the error due to downscaling is much smaller when compared to the error due to the disaggregation scheme.

We also compared the daily time series of spatially averaged precipitation over the entire basin for the two high flow years, July 1994 and July 1982 (Fig. 5), which are marked as black circles in Fig. 4. Figure 5 suggests that the peak rainfall events are not well captured in the temporal disaggregation scheme. For instance, the peak observed rainfall event that occurred during 2–6 July 1994 is completely missed under the disaggregation scheme, resulting in high errors in daily streamflow (Figs. 5c,d). The primary reason for a higher disaggregation error is the dampening of peak rainfall events (on fewer days) that are typically spread over a larger number of days under the disaggregated precipitation. This analysis emphasizes the need to select *k*-NNs that minimize the error in disaggregating the intense precipitation events. We further discuss various approaches for reducing the disaggregation errors in the discussion.

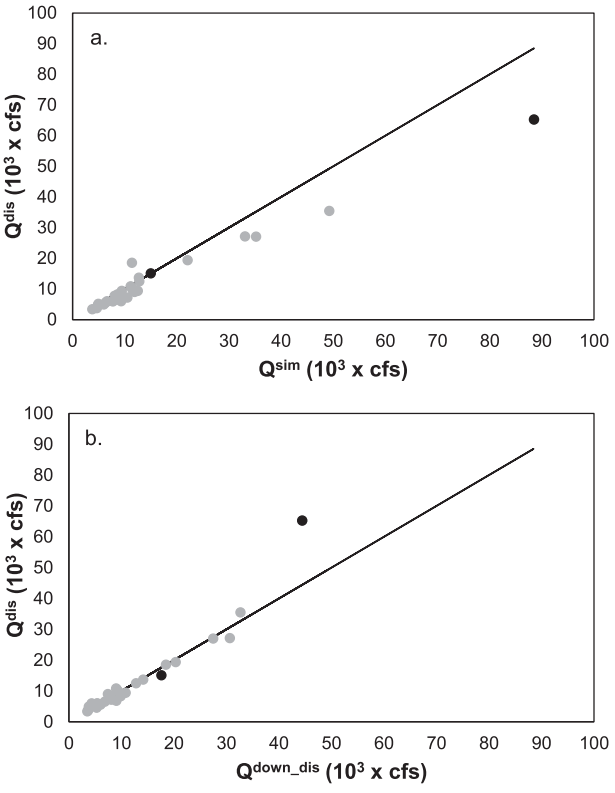

FIG. 4. Scatterplot of streamflow for July over the 1981–2010 period between (a) (Q_{itk}^{dis}) obtained using the *k*-NN approach and simulated streamflow (Q_{it}^{sim}) and (b) (Q_{itk}^{dis}) and spatially downscaled and temporally disaggregated flows ($Q_{itk}^{\text{down\_dis}}$).

2) ERRORS DUE TO FORCINGS, INITIAL CONDITIONS, AND CLIMATOLOGY

Figure 6 shows the R-RMSEs, which are estimated by computing the ratio of RMSE in different forecasting schemes [with lead times $k = 1$ (Fig. 6a) and $k = 3$ (Fig. 6b)] to the mean simulated monthly flows. The R-RMSEs are compared among the four different error diagnostics: 1) downscaling and disaggregation [down_dis; Eq. (2)], 2) climatological forcings [ESP; Eq. (5)], 3) imprecise IHCs [RESP; Eq. (4)], and 4) imprecise ECHAM4.5 forecasts (ECHAM4.5) and the large-scale precipitation forecasts [fcst; Eqs. (6) and (7), respectively]. The *x* axis in Fig. 6 indicates the month of streamflow forecasts but not the month when the forecasts were issued. At a 1-month lead time, errors in all the schemes, except those due to imprecise IHCs (RESP), are high in the month of July, including the scheme that uses observed precipitation but is obtained using downscaling and disaggregation (Fig. 6a). During July, the watershed receives mostly convective precipitation, which is highly variable, resulting in a larger RMSE. A comparison of errors among different schemes at a

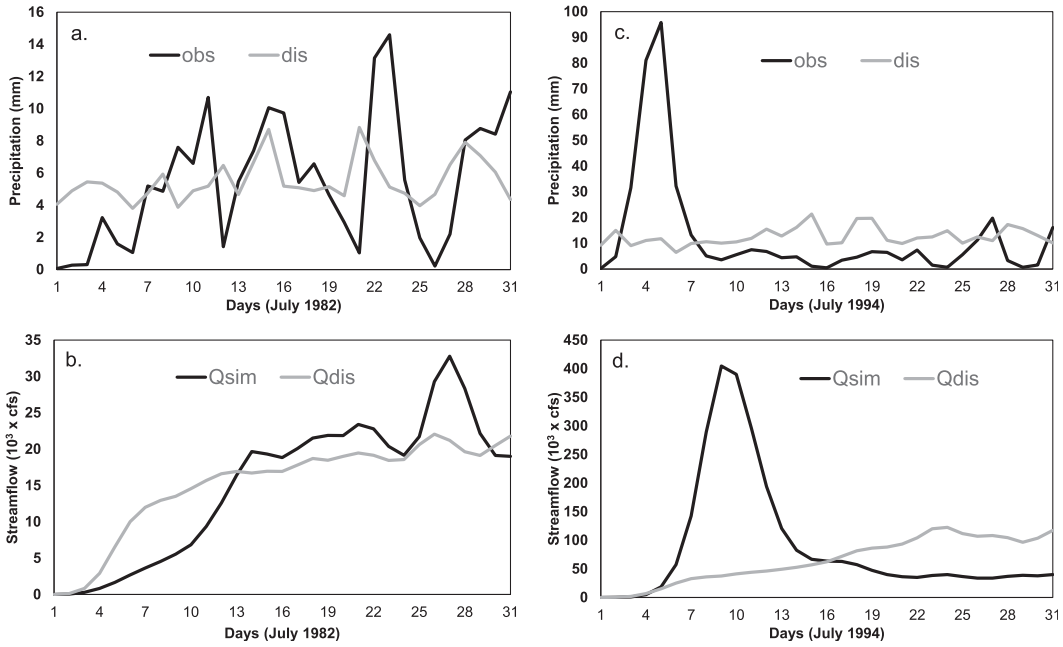

FIG. 5. Comparison of daily time series for July (a),(b) 1982 and (c),(d) 1994 between the observed precipitation (obs) and the temporally disaggregated precipitation (dis) and between the model-simulated streamflow under observed forcings (Q^{sim}) and streamflow obtained by forcing the VIC model with the temporally disaggregated precipitation (Q^{dis}).

1-month lead time (Fig. 6a) indicates that the errors due to combined downscaling and disaggregation are lowest during winter and spring while these errors are higher during summer and fall. In summer and early fall, the region receives extreme (high) precipitation events partially because of tropical storms and hurricanes, which are difficult to predict through the temporal disaggregation schemes, resulting in higher errors. The errors due to the large-scale forecasts alone (fcst), which do not include the errors due to the spatial downscaling and temporal disaggregation, are relatively lower during winter and spring in comparison to other seasons. The large-scale precipitation forecasts derive their skill from ENSO, and the overall skill is relatively better during winter and spring over the study area (Sinha and Sankarasubramanian 2013). However, when we include the errors due to downscaling and disaggregation, that is, under operational forecasting, the ECHAM4.5 errors are relatively higher and are similar in magnitude to the ESP errors at a 1-month lead time. Despite the downscaling and disaggregation errors, the ECHAM4.5 forecast errors are relatively lower than the ESP errors during the winter and spring seasons. Errors under RESP are generally lower in comparison to the other schemes during late spring and summer while the RESP errors are larger during the winter season. This suggests that interannual variability in IHCs during the winter is higher, indicating the importance of updating the IHC. Furthermore, during late spring and summer, the soil moisture is relatively lower because of increased evapotranspiration, and it leads to low interannual variability in IHCs.

Looking beyond a 1-month lead time, the relative errors at both 2- and 3-month lead time are similar, and therefore, results corresponding to a 3-month lead time are presented here. At a 3-month lead time, relative errors are lowest under the RESP, followed by the downscaling and disaggregation errors (Fig. 6b). The errors due to climatological forcings are highest among all the other schemes considered in this study. As expected, all the errors are higher during July because of the inability to capture the convective precipitation in the summer. Comparing the relative errors of a 3-month lead time with a 1-month lead time, we see that the error magnitudes have increased under all the forecasting schemes. However, errors under imprecise IHCs have decreased during early winter (January), spring (April and May), and late fall (December) at a 3-month lead time, indicating the increased role of providing proper forcings over the forecasting period. The skill in the large-scale precipitation forecasts also goes down with the forecast lead time; however, the ECHAM4.5 forecasts still exhibit added value during all the months in comparison to climatological forcings at a 3-month lead time. With this understanding, we next evaluate the role of relative errors due to RESP, ESP, and precipitation forecasts in developing streamflow forecasts.

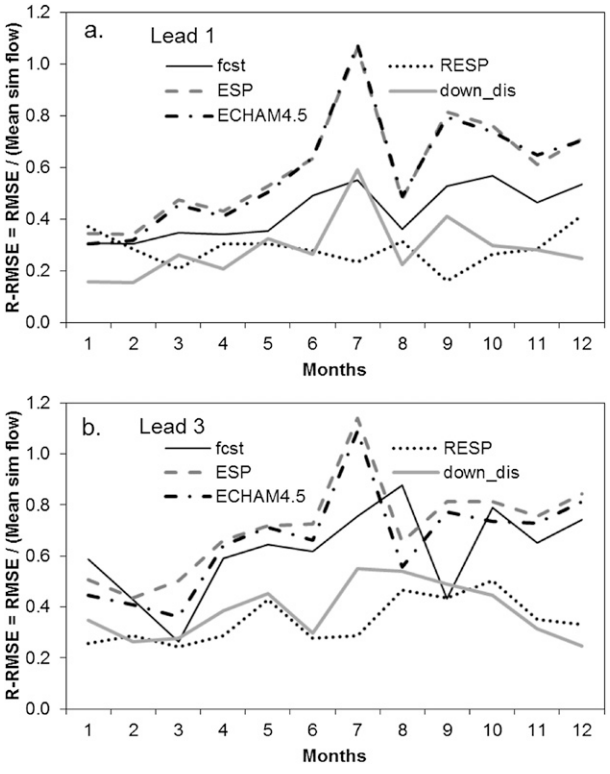

FIG. 6. R-RMSEs at (a) 1- and (b) 3-month lead times due to combined spatial downscaling and temporal disaggregation (down_dis), climatological forcings (ESP), imprecise IHCs (RESP), and ECHAM4.5 forecasts (ECHAM4.5) and the large-scale precipitation forecasts (fcst).

3) RELATIVE CONTRIBUTION OF ERRORS: RESP, ESP, AND fcst

In this section, we discuss the relative contribution of errors between imprecise IHC (RESP) and climatological forcing (ESP) errors as well as between the errors due to large-scale climate forecasts alone [Eq. (6)] and climatological forcing (ESP) errors. If the error ratio of imprecise IHCs to imprecise climatological forcings (ESP) is greater than 1, it implies that errors due to IHCs are high and IHCs play an important role. Similarly, if the ratio of imprecise fcst to imprecise ESP forcings is greater than 1, then the errors in the large-scale climate forecasts are higher, suggesting that climatological forcings perform better than the fcst. Figures 6 and 7 summarize the relative contribution of imprecise IHCs, forcings, and forecasts over different months as a function of lead time.

At a 1-month lead time, IHCs are dominant during early winter (January; Fig. 7a) in comparison to the climatological forcings. However, beyond a 1-month lead time, climatological forcings are dominant in all the months, indicated by the ratio being less than one. This is consistent with the findings of other studies such as Shukla and Lettenmaier (2011) and Li et al. (2009). On the other hand, errors due to large-scale climate forecasts alone (fcst) are lower than the errors due to climatological forcings at 1–3-months lead time during all the months, except January and August at 2–3-months lead time (Fig. 8). This clearly demonstrates that the large-scale precipitation forecasts have useful information in developing streamflow forecasts when compared to the climatological forcings (ESP). Our analyses also show that significant errors are added during the process of spatial downscaling and temporal disaggregation. Despite the errors arising from spatiotemporal downscaling and disaggregation, the ECHAM4.5 precipitation forecasts perform better than the ESP approach at 2–3-months lead time.

b. *Performance in predicting the observed streamflow* (Q_{it}^{obs})

In this section, we consider the performance of schemes when the observed streamflow is used as the reference flow. The advantage of considering the USGS-observed streamflow as a reference is that the forecasts can be compared to the actual basin response to the forcings, thereby providing an opportunity to compare the model error with other sources of errors. To compare different schemes with the observed streamflow, all the VIC model–simulated schemes were corrected for mean monthly bias over the 1981–2010 period of analysis. Percentage bias correction for each month was estimated by comparing the VIC model–simulated flows under observed forcings to the USGS-observed flows for the period 1981–2010. This bias correction was applied to all the VIC model–estimated streamflow schemes when the USGS-observed streamflow was considered as the reference flow. It is similar to Fig. 6, but the reference flow is the observed flow, and the errors due to downscaling and disaggregation are plotted as two separate sources of errors in addition to the model error.

At a 1-month lead time, the sources of error in forecasting the observed streamflow could be ranked (from low to high R-RMSE) as follows (Fig. 9a): spatial downscaling, model, disaggregation, large-scale precipitation forecasts (excluding disaggregation and downscaling errors), imprecise initial conditions (RESP), and the climatological forcings (ESP). However, there are exceptions with the RESP errors being greater than the ESP errors during early winter, late summer, and fall (Fig. 9a). During the winter and fall seasons, initialization of hydrologic models plays an important role. In general, during the fall season, all the schemes perform similarly since the model error is the predominant source of error. This primarily stems from the inability of the VIC model

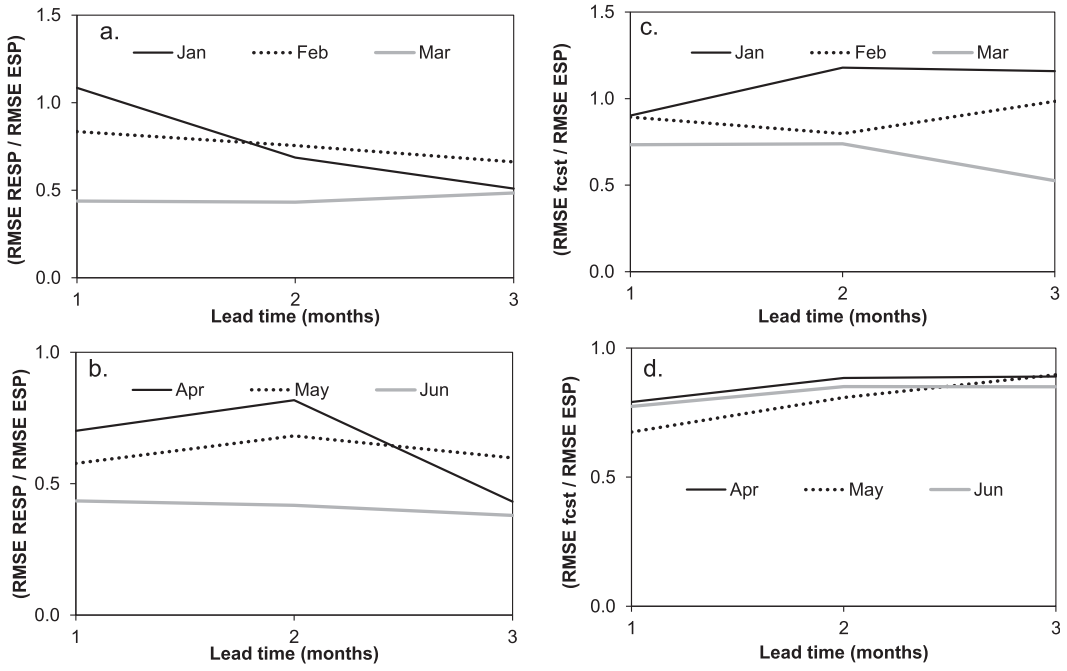

FIG. 7. Ratio of the (a),(b) RMSE under imprecise IHCs (RESP) to the RMSE under the climatological forcings (ESP) and (c),(d) RMSE under the imprecise large-scale precipitation forecasts (fcst) to the climatological forcings as a function of lead time during (top) winter and (bottom) spring.

in capturing the flows during the low flow season, when the Nash–Sutcliffe efficiency is lower than the other months (see Sinha and Sankarasubramanian 2013). Further, during the summer and fall seasons, the large-scale precipitation forecasts (fcst) have higher errors because of the poor skill of the ECHAM4.5 GCM forecasts in predicting precipitation during those seasons.

Looking at the performance of different schemes in predicting observed streamflow at a 3-month lead time (Fig. 9b), the major difference is that the errors due to the large-scale precipitation forecasts (fcst) are relatively higher than downscaling as well as disaggregation errors. Still, the forecast errors (fcst) are relatively lower than the climatological errors in all the seasons, except during fall. Comparing Figs. 6 and 9, the relative performance of most of the schemes remain similar, but the model error could be a significant source, particularly in seasons where the model performs poorly. For instance, in the fall season, the significant source of error is due to the inability of the model to capture the low flows. Thus, it is important to exercise caution on the ability to forecast, particularly in seasons when the model performs poorly.

4. Discussion

Results from Figs. 3 and 4 indicate that the errors due to temporal disaggregation are much higher than the errors due to spatial downscaling in all months. One of the possible reasons for this is that, under the temporal disaggregation scheme, the extreme precipitation events are averaged out and spread over several low-intensity events on multiple days, resulting in increased errors over all the months (Fig. 5). Only during the months of May–July, when convective precipitation is dominant, are the errors due to disaggregation and downscaling comparable. This basically stems from the well-known findings that large-scale climate forecasts have limited skill in predicting warm-season precipitation (Goddard et al. 2003; Barnston et al. 2003). The study considered two disaggregation methods (Prairie et al. 2007; Wood et al. 2002) for quantifying the errors in monthly streamflow forecasts arising from disaggregation. To improve this further, one could consider sophisticated optimization algorithms for disaggregation (e.g., Kumar et al. 2000) that preserve the properties of high-frequency weather events (e.g., 2-day wet/dry spells) in the disaggregated time series. One could also consider combining weather and climate information to develop 15-day-ahead precipitation forecasts, which has been shown to result in improved medium-range precipitation forecasts (Wang et al. 2013). We intend to explore this as part of the future study so that we can choose a strategy that overall reduces both downscaling and disaggregation errors.

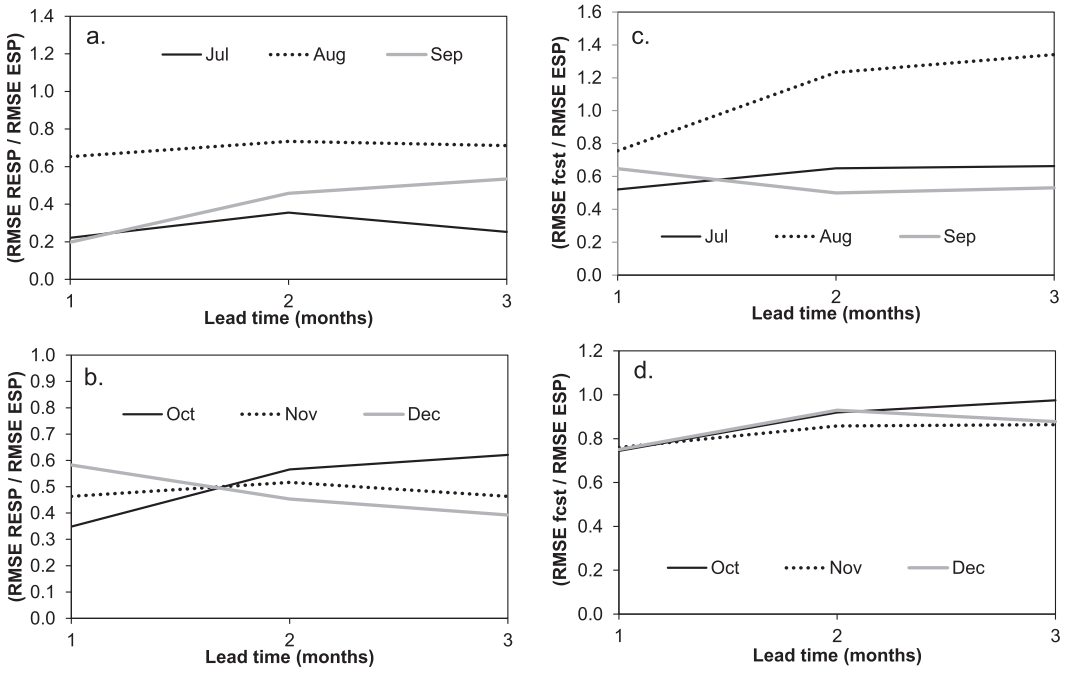

FIG. 8. As in Fig. 7, but for (top) summer and (bottom) fall.

Though the current study decomposed different sources of errors in developing monthly streamflow forecasts for a watershed from the Southeast, the findings, in general, should be applicable for basins under the rainfall–runoff-dominated regime. Extending the analyses at the regional scale is feasible only by considering the reference flow as the simulated streamflow, since comparison with observed streamflow would require routing over various large basins. We are currently working on quantifying the error decomposition in streamflow forecasts over the Sun Belt by considering multiple land surface models within the National Aeronautics and Space Administration's (NASA) Land Information System (LIS; Kumar et al. 2008). Given that the model error is a significant component, utilizing LIS's capability to handle multiple LSMs will provide opportunities for evaluation against multiple reference streamflows. It is also important to recognize that the study considered a single GCM: the ECHAM4.5 forecasts for developing monthly streamflow forecasts. Studies have shown that multimodel climate forecasts consistently perform better than the forecasts from a single GCM (Barnston et al. 2003; Devineni et al. 2008). Our recent study focusing on systematic uncertainty reduction strategies clearly demonstrated that climate model uncertainty should be reduced first, followed by hydrologic model uncertainty for developing improved seasonal streamflow forecast development (Singh and Sankarasubramanian 2014). Forcing multimodel climate forecasts with multiple LSMs could provide additional insights on reducing the uncertainties related to both climatic inputs and LSMs. Thus, a comprehensive understanding of the relative magnitudes of errors arising from climate models, statistical models facilitating scale mismatch in hydroclimatic attributes (e.g., downscaling and disaggregation), and LSMs under different climatic regimes (e.g., arid/humid) will provide useful information on reducing uncertainties related to streamflow forecast development.

5. Conclusions

Despite considerable progress in developing real-time climate forecasts, most studies have evaluated the potential in seasonal streamflow forecasting based on ensemble streamflow prediction (ESP) methods, which only utilize climatological forcings (i.e., no forecast) to develop streamflow forecasts, ignoring the GCM-based climate forecasts. The primary limitation in using real-time climate forecasts developed using GCMs is due to their coarse resolution, which requires spatial and temporal downscaling to force them on land surface models. Consequently, multiple sources of errors are introduced in developing real-time streamflow forecasts utilizing climate information. In this study, we provide a set of error decomposition metrics for addressing the following science questions: 1) How do we attribute the errors in monthly streamflow forecasts to (i) temporal

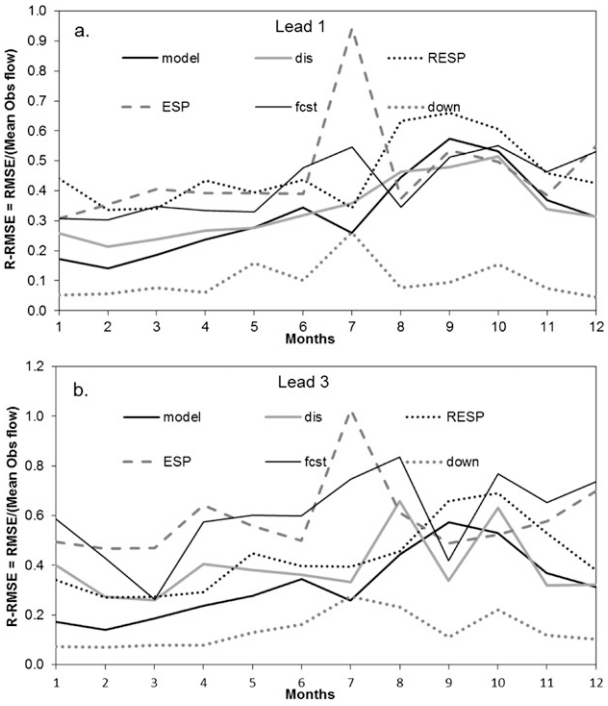

FIG. 9. R-RMSE at (a) 1- and (b) 3-month lead times for the model, temporal disaggregation (dis), spatial downscaling (down), imprecise IHCs (RESP), climatological forcings (ESP), and large-scale precipitation forecasts (fcst). Reference streamflow for comparison is the observed monthly flow at station 02358000.

disaggregation, (ii) spatial downscaling, (iii) imprecise initial conditions, (iv) climatological forcings, (v) imprecise forecasts, and (vi) model error? and 2) How do monthly streamflow forecast errors propagate as a function of different lead time over various seasons? In this section, we discuss these two questions.

a. *Errors due to different sources in monthly streamflow forecasts*

From Figs. 6a and 6b, we clearly understand that the role of IHCs is important up to a 1-month lead time and beyond, where forcings become dominant. This is consistent with the findings of Li et al. (2009) and Shukla and Lettenmaier (2011). The large-scale precipitation forecasts offer a great potential to improve monthly to seasonal streamflow forecasts over the ESP approach. However, the errors resulting from spatial downscaling and temporal disaggregation reduce the climate forecast skill, indicating the improvements in streamflow forecasts are limited to only winter and spring seasons at 1-month lead time. The errors due to temporal disaggregation are much higher than the errors due to spatial downscaling. This primarily arises because of the dampening of peak precipitation events, which are spread as smaller rainfall events over the entire month in the disaggregation scheme. Such distributed precipitation events result in increased storage in the basin, leading to increased evapotranspiration and reduced runoff from the basin. This necessitates the selection of disaggregation schemes that preserve the marginal frequency of extreme precipitation events in a given month. We intend to focus on developing such methods as part of our future effort.

b. *Error propagation of different forecasting scheme with lead time*

At 1–3-months lead time, the relative errors due to climatological forcings are highest among all the schemes, followed by errors in the ECHAM4.5 forecasts. At a 1-month lead time, the ECHAM4.5 precipitation-based streamflow forecasts, which include spatiotemporal downscaling errors, have relatively higher errors than the climatological forcings. Looking at the streamflow forecasts over 2–3-months lead time, the ECHAM4.5 forecasts show better performance than the climatological forcings during all the months, which clearly indicate the added value of monthly updated precipitation forecasts over the climatological forcings. At a 3-month lead time, the increasing order of errors during winter and early spring are RESP, downscaling and disaggregation, ECHAM4.5 forecasts, and climatological forcings. We understand the above findings are with regard to the ACF river basin, but we expect these findings to be similar for the basins experiencing a rainfall–runoff regime. We are currently working on decomposing these sources of errors considering various target basins over the U.S. Sun Belt.

Despite the advances in developing climate forecasts, there are several challenges that limit the application of large-scale climate forecasts in real-time streamflow forecasts. One of the primary challenges in streamflow forecasting is the scale mismatch between the GCM forecasts and the scale of land surface models at which streamflow forecasts are developed. To overcome this, we employ downscaling and disaggregation models to bring the large-scale information to the LSM scale. One could pursue dynamical downscaling for addressing the scale mismatch. Given the limited number of retrospective regional climate model runs, it takes substantial effort to pursue dynamical downscaling. It is clear that updating IHCs provides skillful streamflow forecasts up to a 1-month lead time in a rainfall–runoff regime. It is important to note that these updated IHCs constitute model errors, which could be reduced further by using data assimilation. Application of data assimilation could reduce uncertainty in IHCs and could result in improving IHC's role in developing skillful streamflow forecasts even beyond 1-month lead time. Given that the

basins are dominated with the rainfall–runoff regime, this could be useful as the variations in IHCs are higher as opposed to basins dominated by the snowmelt regime. We intend to consider this as part of the future study.

Acknowledgments. We thank NOAA for providing funding to conduct this research through Grant NA09OAR4310146. We are thankful for the comments of the three anonymous reviewers, particularly on the decomposition of error formulations, as their suggestions helped improve the overall quality of this manuscript.

REFERENCES

Ajami, N. K., Q. Duan, and S. Sorooshian, 2007: An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resour. Res.,* **43,** W01403, doi:10.1029/2005WR004745.

Barnston, A. G., 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, doi:10.1175/BAMS-84-12-1783.

Betts, R. A., P. M. Cox, S. E. Lee, and F. I. Woodward, 1997: Contrasting physiological and structural vegetation feedbacks in climate change simulations. *Nature,* **387,** 796–799, doi:10.1038/42924.

Caldwell, P., 2010: California wintertime precipitation bias in regional and global climate models. *J. Appl. Meteor. Climatol.,* **49,** 2147–2158, doi:10.1175/2010JAMC2388.1.

Cherkauer, K. A., and D. P. Lettenmaier, 2003: Simulation of spatial variability in snow and frozen soil. *J. Geophys. Res.,* **108,** 8858, doi:10.1029/2003JD003575.

Cocke, S., T. E. LaRow, and D. W. Shin, 2007: Seasonal rainfall prediction over the southeast United States using the Florida State University nested regional spectral model. *J. Geophys. Res.,* **112,** D04106, doi:10.1029/2006JD007535.

Day, G. N., 1985: Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plann. Manage.,* **111,** 157–170, doi:10.1061/(ASCE)0733-9496(1985)111:2(157).

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

Di Luca, A., R. de Elía, and R. Laprise, 2012: Potential for added value in precipitation simulated by high-resolution nested regional climate models and observations. *Climate Dyn.,* **38,** 1229–1247, doi:10.1007/s00382-011-1068-3.

Duan, Q., N. K. Ajami, X. Gao, and S. Sorooshian, 2007: Multimodel ensemble hydrologic prediction using Bayesian model averaging. *Adv. Water Resour.,* **30,** 1371–1386, doi:10.1016/j.advwatres.2006.11.014.

Franz, K. J., H. C. Hartmann, S. Sorooshian, and R. Bales, 2003: Verification of National Weather Service ensemble streamflow predictions for water supply forecasting in the Colorado River basin. *J. Hydrometeor.,* **4,** 1105–1118, doi:10.1175/1525-7541(2003)004<1105:VONWSE>2.0.CO;2.

Goddard, L., A. G. Barnston, and S. J. Mason, 2003: Evaluation of the IRI's "net assessment" seasonal climate forecasts: 1997–2001. *Bull. Amer. Meteor. Soc.,* **84,** 1761–1781, doi:10.1175/BAMS-84-12-1761.

Halmstad, A., M. R. Najafi, and H. Moradkhani, 2013: Analysis of precipitation extremes with the assessment of regional climate models over the Willamette River basin, USA. *Hydrol. Processes,* **27,** 2579–2590, doi:10.1002/hyp.9376.

Hamlet, A. F., D. Huppert, and D. P. Lettenmaier, 2002: Economic value of long-lead streamflow forecasts for Columbia River hydropower. *J. Water Resour. Plann. Manage.,* **128,** 91–101, doi:10.1061/(ASCE)0733-9496(2002)128:2(91).

Hartmann, H. C., T. C. Pagano, S. Sorooshian, and R. Bales, 2002: Confidence builders: Evaluating seasonal climate forecasts from user perspectives. *Bull. Amer. Meteor. Soc.,* **83,** 683–698, doi:10.1175/1520-0477(2002)083<0683:CBESCF>2.3.CO;2.

Hayhoe, K., and Coauthors, 2004: Emissions pathways, climate change, and impacts on California. *Proc. Natl. Acad. Sci. USA,* **101,** 12 422–12 427, doi:10.1073/pnas.0404500101.

Koster, R. D., and M. J. Suarez, 1995: Relative contributions of land and ocean processes to precipitation variability. *J. Geophys. Res.,* **100,** 13 775–13 790, doi:10.1029/95JD00176.

——, S. P. P. Mahanama, B. Livneh, D. P. Lettenmaier, and R. H. Reichle, 2010: Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nat. Geosci.,* **3,** 613–616, doi:10.1038/ngeo944.

Kumar, N., U. Lall, and M. R. Petersen, 2000: Multisite disaggregation of monthly to daily streamflow. *Water Resour. Res.,* **36,** 1823–1833, doi:10.1029/2000WR900049.

Kumar, S., C. D. Peters-Lidard, J. L. Eastman, and W. K. Tao, 2008: An integrated high-resolution hydrometeorological modeling testbed using LIS and WRF. *Environ. Model. Software,* **23,** 169–181, doi:10.1016/j.envsoft.2007.05.012.

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.

Leung, L. R., Y. Qian, X. D. Bian, and A. Hunt, 2003: Hydroclimate of the western United States based on observations and regional climate simulation of 1981–2000. Part II: Mesoscale ENSO anomalies. *J. Climate,* **16,** 1912–1928, doi:10.1175/1520-0442(2003)016<1912:HOTWUS>2.0.CO;2.

——, ——, ——, W. M. Washington, J. G. Han, and J. O. Roads, 2004: Mid-century ensemble regional climate change scenarios for the western United States. *Climatic Change,* **62,** 75–113, doi:10.1023/B:CLIM.0000013692.50640.55.

Li, H., L. Luo, E. F. Wood, and J. Schaake, 2009: The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *J. Geophys. Res.,* **114,** D04114, doi:10.1029/2008JD010969.

Li, S., and L. Goddard, 2005: Retrospective forecasts with the ECHAM4.5 AGCM. IRI Tech. Rep. 05-02, Earth Institute, Columbia University, Palisades, NY, 16 pp.

Li, W., and A. Sankarasubramanian, 2012: Reducing hydrologic model uncertainty in monthly streamflow predictions using multimodel combination. *Water Resour. Res.,* **48,** W12516, doi:10.1029/2011WR011380.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.,* **99,** 14 415–14 428, doi:10.1029/94JD00483.

——, ——, and ——, 1996: One-dimensional statistical dynamic representation of sub-grid spatial variability of precipitation in the two layer Variable Infiltration Capacity model. *J. Geophys. Res.,* **101,** 21 403–21 422, doi:10.1029/96JD01448.

Lohmann, D., E. Raschke, B. Nijssen, and D. P. Lettenmaier, 1998a: Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model. *Hydrol. Sci. J.,* **43,** 131–142, doi:10.1080/02626669809492107.

——, ——, ——, and ——, 1998b: Regional scale hydrology: II. Application of the VIC-2L model to the Weser River, Germany. *Hydrol. Sci. J.,* **43,** 143–158, doi:10.1080/02626669809492108.

Luo, L., and E. F. Wood, 2008: Use of Bayesian merging techniques in a multimodel seasonal hydrologic ensemble prediction system for the eastern United States. *J. Hydrometeor.,* **9,** 866–884, doi:10.1175/2008JHM980.1.

——, ——, and M. Pan, 2007: Bayesian merging of multiple climate model forecasts for seasonal hydrological predictions. *J. Geophys. Res.,* **112,** D10102, doi:10.1029/2006JD007655.

Mahanama, S., B. Livneh, R. Koster, D. P. Lettenmaier, and R. Reichle, 2012: Soil moisture, snow, and seasonal streamflow forecasts in the United States. *J. Hydrometeor.,* **13,** 189–203, doi:10.1175/JHM-D-11-046.1.

Maurer, E. P., and D. P. Lettenmaier, 2003: Predictability of seasonal runoff in the Mississippi River basin. *J. Geophys. Res.,* **108,** 8607, doi:10.1029/2002JD002555.

——, and H. G. Hidalgo, 2008: Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.,* **12,** 551–563, doi:10.5194/hess-12-551-2008.

——, A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous Unites States. *J. Climate,* **15,** 3237–3251, doi:10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.

——, D. P. Lettenmaier, and N. J. Mantua, 2004: Variability and potential sources of predictability of North American runoff. *Water Resour. Res.,* **40,** W09306, doi:10.1029/2003WR002789.

Murphy, J., 1999: An evaluation of statistical and dynamical techniques for downscaling local climate. *J. Climate,* **12,** 2256–2284, doi:10.1175/1520-0442(1999)012<2256:AEOSAD>2.0.CO;2.

Pagano, T. C., H. C. Hartmann, and S. Sorooshian, 2001: Using climate forecasts for water management: Arizona and the 1997–1998 El Niño. *J. Amer. Water Resour. Assoc.,* **37,** 1139–1153, doi:10.1111/j.1752-1688.2001.tb03628.x.

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.

Sankarasubramanian, A., A. Sharma, U. Lall, and S. Espinueva, 2008: Role of retrospective forecasts of GCMs forced with persisted SST anomalies in operational streamflow forecasts development. *J. Hydrometeor.,* **9,** 212–227, doi:10.1175/2007JHM842.1.

Shukla, S., and D. P. Lettenmaier, 2011: Seasonal hydrologic prediction in the United States: Understanding the role of initial hydrologic conditions and seasonal climate forecast skill. *Hydrol. Earth Syst. Sci.,* **15,** 3529–3538, doi:10.5194/hess-15-3529-2011.

Singh, H., and A. Sankarasubramanian, 2014: Systematic uncertainty reduction strategies for developing streamflow forecasts utilizing multiple climate models and hydrologic models. *Water Resour. Res.,* **50,** 1288–1307, doi:10.1002/2013WR013855.

Sinha, T., and A. Sankarasubramanian, 2013: Role of climate forecasts and initial conditions in developing streamflow and soil moisture forecasts in a rainfall–runoff regime. *Hydrol. Earth Syst. Sci.,* **17,** 721–733, doi:10.5194/hess-17-721-2013.

Slack, J. R., A. Lumb, and J. M. Landwehr, 1993: Hydro-Climatic Data Network (HCDN) streamflow data set, 1874–1988. USGS Water-Resources Rep. 93-4076. [Available online at http://pubs.usgs.gov/wri/wri934076/.]

Stern, P. C., and W. E. Easterling, 1999: *Making Climate Forecasts Matter*. National Academies Press, 192 pp.

Wang, H., A. Sankarasubramanian, and R. S. Ranjithan, 2013: Integration of climate and weather information for improving 15-day-ahead accumulated precipitation forecasts. *J. Hydrometeor.,* **14,** 186–202, doi:10.1175/JHM-D-11-0128.1.

Wilby, R. L., L. E. Hay, W. J. Gutowski, R. W. Arritt, E. S. Takle, Z. Pan, G. H. Leavesley, and M. P. Clark, 2000: Hydrological responses to dynamically and statistically downscaled climate. *Geophys. Res. Lett.,* **27,** 1199–1202, doi:10.1029/1999GL006078.

Wood, A. W., and D. P. Lettenmaier, 2006: A test bed for new seasonal hydrologic forecasting approaches in the western United States. *Bull. Amer. Meteor. Soc.,* **87,** 1699–1712, doi:10.1175/BAMS-87-12-1699.

——, E. P. Maurer, A. Kumar, and D. P. Lettenmaier, 2002: Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.,* **107,** 4429, doi:10.1029/2001JD000659.

——, L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change,* **62,** 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.

Yuan, X., E. F. Wood, L. Luo, and M. Pan, 2011: A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophys. Res. Lett.,* **38,** L13402, doi:10.1029/2011GL047792.

——, X.-Z. Liang, and E. Wood, 2012: WRF ensemble downscaling seasonal forecasts of China winter precipitation during 1982–2008. *Climate Dyn.,* **39,** 2041–2058, doi:10.1007/s00382-011-1241-8.

——, E. F. Wood, J. K. Roundy, and M. Pan, 2013: CFSv2-based seasonal hydroclimatic forecasts over the conterminous United States. *J. Climate,* **26,** 4828–4847, doi:10.1175/JCLI-D-12-00683.1.