AMERICAN
METEOROLOGICAL
SOCIETY

Earth Interactions

The DOI for this manuscript is doi: 10.1175/2013EI000519.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Li, H., 2013: Seasonal hydrological forecasts for watersheds over the Southeastern United States for boreal summer and fall seasons. Earth Interact. doi:10.1175/2013EI000519.1, in press.

Seasonal hydrological forecasts for watersheds over the Southeastern United States for boreal summer and fall seasons

Satish Bastola[1, #] , Vasubandhu Misra[1,2,3] and Haiqin Li[1]

[1]Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida

[2] Department of Earth, Ocean and Atmospheric Sciences, Florida State University, Tallahassee, Florida

[3] Florida Climate Institute, Florida State University, Tallahassee, Florida

Submitted to Earth Interactions

18

Corresponding author; Email: sbastola@fsu.edu

Abstract

We evaluate the skill of a suite of seasonal hydrological prediction experiments over 28 watersheds throughout the Southeastern United States (SEUS), including Florida, Georgia, Alabama, South Carolina, and North Carolina. The seasonal climate retrospective forecasts (the Florida Climate Institute–Florida State University Seasonal Hindcast at 50 km [FISH50]) is initialized in June and integrated through November of each year from 1982 through 2001. Each seasonal climate forecast has six ensemble members. An earlier study showed that FISH50 represents state-of-the-art seasonal climate prediction skill for the summer and fall seasons, especially in the subtropical and higher latitudes. The retrospective prediction of streamflow is based on multiple calibrated rainfall-runoff models. The hydrological models are forced with rainfall from FISH50, (quantile-based) bias-corrected FISH50 rainfall (FISH50_BC), and resampled historical rainfall observations based on matching observed analogues of forecasted quartile seasonal rainfall anomalies (FISH50_Resamp).

The results show that direct use of output from the climate model (FISH50) results in huge biases in predicted streamflow, which is significantly reduced with bias correction (FISH50_BC) or by FISH50_Resamp. On a discouraging note, we find that the deterministic skill of retrospective streamflow prediction as measured by the normalized root mean square error is poor compared to the climatological forecast irrespective of how FISH50 (e.g., FISH50_BC, FISH50_Resamp) is used to force the hydrological

models. However, our analysis of probabilistic skill from the same suite of retrospective prediction experiments reveals that over the majority of the 28 watersheds in the SEUS, significantly higher probabilistic skill than climatological forecast of streamflow can be harvested for the wet/dry seasonal anomalies (i.e., extreme quartiles) using FISH50_Resamp as the forcing. We contend that given the nature of the relatively low climate predictability over the SEUS, high deterministic hydrological prediction skills will be elusive. Therefore, probabilistic hydrological prediction for the SEUS watersheds is very appealing, especially with the current capability of generating a comparatively huge ensemble of seasonal hydrological predictions for each watershed and for each season, which offers a robust estimate of associated forecast uncertainty.

1. Introduction

The Southeastern United States (SEUS) has the fastest growing population in the country (Seager et al. 2009) and has a multibillion dollar agricultural industry, which is largely rain fed (Hansen et al. 1998). Therefore, the SEUS requires a reliable prediction of streamflow in its watersheds. Several studies have already indicated the benefits of skillful streamflow prediction for water supply, agriculture, and hydropower generation (e.g., Broad et al. 2007; Yao and Georgakakos 2001). The SEUS is the wettest region in the continental United States during the boreal summer season (Chan and Misra 2010). Chan and Misra (2010) further indicate that the SEUS also exhibits the largest seasonal variance in the summer. However, most global climate models exhibit a poor seasonal prediction skill, especially for rainfall during the summer and fall seasons over the SEUS (Stefanova et al. 2012). In the winter and spring seasons, large-scale variations of global SST (e.g., El Niño and the Southern Oscillation) have significant influence on the rainfall variability in the SEUS (Ropelewski and Halpert 1986, 1987). There is an absence of such teleconnection in the summer and fall seasons (especially at seasonal to interannual time scales), which leads to corresponding reduction of the seasonal climate prediction skill from climate models (Stefanova et al. 2012a).

Seasonal climate predictability is, however, rapidly evolving, with state-of-the-art climate models that have increased in complexity and resolution over time (Saha et al. 2006, 2010; Gent et al. 2010, 2011), with updated data-assimilated products available for initializing the various components of the climate system (Derber and Rosati 1989; Rosati et al. 1997; Balmaseda et al. 2008; Cazes-Boezio et al. 2008; Saha et al. 2010; Xue et al. 2011; Paolino et al. 2012), with newer ways of initializing the climate models (Zhang et al. 2007; Yang et al. 2009; Balmaseda and Anderson 2009), and with increased participation from several modeling groups (Palmer et al. 2004; Kirtman and Min 2009) to generate multimodel estimates of seasonal climate. More recently, two multi-institutional real-time projects in seasonal prediction have been instituted: the U.S. National Multimodel Ensemble (NMME; http://www.cpc.ncep.noaa.gov/products/ctb/MMEWhitePaperCPO_revised.pdf) and the European Seasonal to Interannual Prediction (EUROSIP; http://cosmos.enes.org/uploads/media/TStockdale.pdf) project that is anticipated to further improve seasonal climate prediction skill and more importantly produce robust estimates of seasonal forecast uncertainty.

Much of hydrological forecasting is based on empirical methods, like linear regression, which use initial conditions and information on future climate conditions as predictors (Rosenberg et al. 2011; Pagano et al. 2009). However, since the introduction of extended streamflow prediction (Day 1985), the field of surface hydrological forecasting has developed relatively fast with ensemble streamflow prediction (ESP; Franz et al. 2003; McEnery et al. 2005; Wood and Schaake 2008; Bohn et al. 2010). ESP makes probabilistic forecasts by running the calibrated hydrological models with a number of realizations of meteorological forcing resampled from historical observations based on matching analogues of the forecasted climate state. Calibrated hydrological models are first run with historical observed climate data until the time of forecast to obtain initial conditions for the hydrological model. Then the weather sequence resampled from observed historical record corresponding to the closest analogue of the forecasted climate state is used to force the hydrological models. Generations of daily precipitation by resampling from historical record are widely used in hydrological, ecological, and agriculture applications (Lobell et al. 2006; Wood and Lettenmaier 2006). The physical and conceptual basis of hydrological models allows ESP to overcome some of the limitations of the regression-based method. Furthermore, ESP is more flexible in the sense that it can use output from different sources (e.g., different climate models). Unlike the deterministic approach, the ESP framework provides information on the uncertainty surrounding the point estimate, which may benefit water resources management (Krzysztofowicz 2001). However, ESP is based on the premise that historical records are representative of the possible future.

Some popular approaches to applying climate forecast in hydrology include using observed resampling from historical data [e.g., the Schaake shuffle (Clark et al. 2004), weather generators (Wilks 2002; Voisin et al. 2011), and climate downscaling (Wood et al. 2004; Bastola and Misra 2013)]. In a resampling-based approach (e.g., the Schaake shuffle), the number of years of historical data potentially restricts the number of ensembles that can be generated. Furthermore, the Schaake shuffle reorders the samples to preserve the observed spatial and temporal structure (Clark et al. 2004). The size of the ensemble further reduces when resampling is made conditional upon the seasonal forecast information. This makes the weather generator attractive. Weather generators can generate daily weather sequences conditioned on climate forecasts. Weather generators, however, may have limitations in reproducing certain observed statistics [e.g., underestimation of interannual variance of monthly mean, the distribution of hot and cold spells (Semenov and Barrow, 1997)]. Furthermore, point weather generators are inappropriate for multisite extension of the meteorological forcing that would become essential if they were to be used for distributed or semi-distributed hydrological model applications (Wilks 2002). In the climate model-based approach, the outputs from global climate models are downscaled to finer resolutions and bias corrected to produce the forcing for the hydrological model (e.g., Wood et al. 2005; Wood et al. 2002). Examining the prospect of the National Centers for Environmental Prediction (NCEP) medium-range forecast (MRF) for streamflow prediction over the United States, Clark and Hay (2004) observed significantly large biases in precipitation and temperature over many areas. However, they observed improvement in streamflow prediction over some watersheds with statistical postprocessing of the NCEP MRF output.

Apart from the skill of the seasonal climate forecasts, the fidelity of a hydrological forecast also relies on the quality of the hydrological model in simulating the streamflow. However, the predictions from hydrological models are also plagued with uncertainties. These uncertainties stem from a variety of sources such as uncertainty in the parameters and structure of hydrological models. In response to such challenges, some studies in hydrological literature have adopted the multimodel approach in which streamflows from different hydrological models are combined using ranges of statistical postprocessing tools (e.g., Ajami et al. 2006; Duan et al. 2006; Bastola et al. 2011; Bastola and Misra 2013).

In this study we examine the results of a comprehensive set of seasonal streamflow predictions over 28 watersheds distributed throughout the SEUS with the aim of exploring the utility of the Florida Climate Institute–Florida State University Seasonal Hindcasts at 50-km grid resolution (FISH50; Misra et al. 2012) as a tool for seasonal hydrologic forecasting. FISH50, by way of its spatial resolution and its overall seasonal prediction skill compared to the NMME models that also include the operational seasonal climate forecasts from the National Centers for Environmental Prediction (NCEP), represents state-of-the-art seasonal prediction skill (Misra et al. 2012). Misra et al. (2012) showed that seasonal prediction skill for the boreal summer and fall seasons is superior to most of the NMME models in the subtropical latitudes of the SEUS. The streamflow forecasting is based on calibrated hydrological models. We examine both the deterministic and the probabilistic forecast skill of rainfall and streamflow and discuss the shortcomings and virtues of our forecast system.

2. Study region and data

In the SEUS region, which is one of the wettest parts of the United States in the boreal summer and fall seasons (Chan and Misra 2010), surface streamflow is a major source of the river system. We have a priori selected 28 watersheds that are minimally affected by water management and are part of the Model Parameter Estimation Experiment

(MOPEX; Schaake et al. 2006). Among the selected watersheds, six, namely 2165000, 2217500, 210200, 3451500, 3454000, and 2156500, are affected to some extent by power plants and small ponds on the tributaries, suggesting that they may be more suitable for analysis at monthly time intervals (see http://pubs.usgs.gov/wri/wri934076/stations/ for details). Figure 1 shows the map of the watersheds (and their sub-basins) included in this study. Further information on the characteristics of these watersheds can be found in Bastola and Misra (2013).

The hydrological models used in this study are forced from the seasonal climate hindcast data of FISH50 (Misra et al. 2012). A brief outline of FISH50 data is provided in Table 1 (see Misra et al. 2012 for further details on FISH50). The retrospective seasonal forecasts of FISH50 are conducted from June 1 to December 1 of each year from 1982 to 2008. For each season there are six ensemble members of the seasonal climate hindcast. The observed rainfall dataset follows from the unified daily US precipitation analysis of the Climate Prediction Center (CPC) available at 0.25° resolution (Higgins et al. 2000). The CPC data, available from 1948 onward, are used as our observed rainfall. The number of meteorological pixels, for both FISH50 and CPC, contained within each basin is shown in Figure 2. It is clear that the resolutions of FISH50 and CPC are relatively coarse and the size of the watersheds is comparatively small. This is a reminder that FISH50, which offers the highest spatial resolution among the existing seasonal climate hindcast data sources (e.g., NMME, EUROSIP), is far from sufficient for hydrological applications.

3. Generation of the forcing for the hydrological models

Hydrological models used for hydrological forecasting usually require meteorological forcing at relatively high (typically daily) temporal resolution. However, given the uncertainty, coarse resolution, and comparatively low fidelity of the climate models, the direct use of the climate model forecast variables to force the hydrological model does not yield reliable results (e.g., Clark and Hay 2004; Wood et al. 2002). Therefore, (either statistical or dynamical) downscaling the output from climate models at locally relevant spatial resolution is often considered for hydrological applications. In this study, three different variants of FISH50 forcing are used for seasonal hydrologic hindcast experiment: FISH50 daily data as are (FISH50), bias-corrected FISH50 (FISH50_BC), and resampling from observed historical record on the basis of the closest analogue of the forecasted climate state (FISH50_Resamp).

In generating the forcing for FISH50_BC, we correct for biases in FISH50 rainfall following the quantile-based bias correction method (Eq 1), which has been widely used in hydrological literature (e.g., Li and Wood 2010; Wood et al. 2004). For each grid, the cumulative distribution function (CDF) of observed and FISH50 datasets is derived.

$$\hat{x}^{m}{}_{i,t} = F^{-1}{}_{obs}(F_{\mathrm{mod}}(x^{m}{}_{i,t})), \qquad (1)$$

where $\hat{x}^{m}{}_{i,t}$ $x^{m}{}_{i,t}$ are the corrected and the uncorrected estimates of a variable i at timestep t for the month m of the selected grid. Fobs (.) and Fmod (.) are the empirical CDFs of the observed and the modeled datasets for the same grid. Maraun (2013) discusses in great detail issues related to techniques based on model output statistics such as quantile mapping (QM; Eq 1). In hydrological applications, correcting for biases in coarse resolution data with fine resolution data implies that QM also serves to downscale the variable to fine resolution, which has the effect of inflating the variance, and this can have severe consequences (Maraun 2013). The author argues that the temporal structure of rainfall derived from QM is based on coarse resolution data, so corrected data may fail to reproduce the local-scale temporal structure. Consequently, QM may not be sufficient to bridge the gap in scale. More importantly, the reproduction of temporal structure is important when the focus of model simulation is on flood risk and small watersheds, and on implementation of a distributed hydrological model that divides the basin into units much smaller than those used in lumped or semi-distributed modeling. Over the SEUS, regional scale phenomena such as sea breeze and thunderstorms are strong; therefore, dynamic downscaling, which aims to resolve subgrid scale processes, is a strong contender for downscaling (Stefanova et al. 2012b). However, because of the high computational demand, only statistical approaches are revisited in this study for downscaling and bias correction.

In lieu of statistical bias correction, resampling historical observations on the basis of matching observed analogues of a certain forecasted climate state is another commonly adopted approach to force hydrological models to circumvent the bias in climate models. FISH50 is shown to produce reasonable conditional probability skill of seasonal anomalies of precipitation (Misra et al. 2012). For FISH50_Resamp we leverage this feature of FISH50 to isolate observed analogues on the basis of the forecasted seasonal (six-month) anomalies of precipitation averaged over the watershed in consideration, which are then used to obtain the meteorological forcing at high (daily) temporal and spatial (sub-basin) resolution. Such resampling methods preserve various moments of a time series (e.g., Efron 1979).

Following the methodology of Prudhomme and Davies (2009), we use block resampling without replacement. Here a block is defined as a continuous six-month duration of daily rainfall. The block could have been made smaller (i.e., one month or three months) for the generation of the time series. However, resampling with such a small block is likely to affect the seasonal structure and consequently introduce large biases (e.g., Prudhomme and Davies 2009). Furthermore, FISH50 are six-month long integrations and therefore we have the ability to extend this block to six months.

Daily meteorological variables resampled from historical record at a single site or watershed is usually used in lumped hydrological modeling that treats the whole watershed as a single unit. However, multisite extension or correction to samples from historical record may be required to generate model forcing for distributed or semi-distributed hydrological models because the rainfall field of each sub-basin or station cannot be treated independently. Hydrological models used in this study (see below) are semi-distributed and require sub-basin average rainfall, temperature, and potential evapotranspiration (PET). Therefore, generation of spatially coherent sub-basin average data is essential. A simple two-step procedure is used in this study. First, spatially averaged rainfall (i.e., at basin scale) is generated by resampling from the observed historical record. Second, on the basis of the calendar date of the resampled data in the previous step, sub-basin average rainfalls are extracted from the relatively high resolution gridded observation. Let $r_{i,j}^{F}, r_{i,j}^{O}$ be the forecasted and observed rainfall for ith day and for

j^{th} sub-basin. The rainfall observations and corresponding FISH50 data are then used to define the limits of very wet, medium wet, medium dry, and very dry conditions (referred to as six-month rainfall quartile categories). Average rainfall over the period of six months (June through November) is sorted in ascending order and then partitioned into four groups of equal size quartiles: bottom, lower middle, upper middle, and top. For a given river system, O and F (Eq 2) are the vectors of observed and forecasted rainfall quartiles for categories x and y. Similarly, rainfall distribution (r_dist; Eq 3) is the vector of q ensemble members resampled from historical precipitation at the i^{th} time step and for the j^{th} sub-basin.

$$O = \{x_1, x_2, \ldots. x_N\}; F = \{y_1, y_2, \ldots. y_N\}, \qquad (2)$$

Where $x_1 = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{p}\sum_{j=1}^{p} r_{i,j}^{O}; \;\; y_1 = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{p}\sum_{j=1}^{p} r_{i,j}^{F}$, n is the number of the time step, p is the number of the sub-basin

$$r\_dist = \{z_{i,j,1}, z_{i,j,2} \ldots. z_{i,j,q}\} \;\; \forall i = 1, n; j = 1, p \qquad (3)$$

where n and p are as defined in Eq (2), and q is the number of ensemble members. For each watershed, rainfall category in Eq 2 is defined on the basis of a six-month (June through November) block of spatially averaged rainfall for a period of 20 years (1982–2001) for both CPC and FISH50. In this study, the rainfall total is categorized as very wet (VW), very dry (VD), medium wet (MW) and medium dry (MD).

The procedure to generate a conditioned daily sequence of weather data for the semi-distributed hydrological models (Eq 3) is as follows:

1. Obtain the quartile category of the six-month averaged (June–November) rainfall forecast (x in Eq 2) for a given year and watershed region.

2. Select q (=10 for this study) times a block (six months, from June–November) of weather sequence from observed historical records that has the same forecast category as defined in step 1. Repeat this procedure independently for each ensemble member of FISH50 for a given year. At each instance (q) of selecting the observed analogue, store the corresponding date (i.e., year, month, and day).
3. In correspondence with the selected date from step 2, select for each sub-basin the precipitation, temperature, and PET from the observed historical record so that the disaggregation from watershed to sub-basin scale is limited by the resolution of the observed rainfall. (For very small watersheds, the values of daily rainfall may be alike for the sub-basins).

In effect, in FISH50_Resamp, we generate 10 resamples for each ensemble member, and then propagate them through three hydrological models. We thus obtain 180 (=3x6x10) estimates of streamflow for each watershed per season.

4. Description of the hydrological models and experimental setup

In hydrological modeling, uncertainty stems from a variety of sources. Readers are directed to Beven et al. (2012) and Clark et al. (2012) for an ongoing debate on how to assess hydrological model uncertainty. Implementation of a robust statistical framework for assessing uncertainty in a hydrological model is vital for water resource planning and management (e.g., Steinschneider et al. 2012). Despite their known limitations, conceptual rainfall-runoff (RR) models continue to be widely used for assessing the impacts of climate change on water resources and for projecting potential ranges of future climate change impacts (e.g., Bastola et al. 2011). In this study the uncertainty in simulation is accounted for by combining simulations obtained from three RR models [the Hydrologic MODel (HyMOD; Wagener et al. 2001), the Nedbør-Afstrømnings Model (NAM; Madsen 2001), and the tank model (Sugawara 1995)]. These models have been widely used to simulate hydrological responses in watersheds located in various climatic regions (e.g., Bastola et al. 2011;Wagener et al. 2001; Tingsanchali and Gautam 2000). All three models quantify different components of the hydrological cycle by accounting for moisture content in different but mutually interrelated storage.

The RR models discussed above require the calibration of key parameters to yield reliable predictions (Gupta et al. 2003). In the present application, HyMOD, NAM, and tank require 6, 10, and 16 parameters, respectively, to be estimated through model calibration. Several studies discuss problems associated with model calibration and parameter uncertainty (e.g., Kuczera 1997; Beven et al. 1992; Duan et al. 1992). There is growing agreement in the hydrological modeling community that a large combination of parameters results in reasonable model simulation (Beven 2005). We implement a multimodel and multiparameter simulation to conduct the hydrological forecast experiments. The generalized likelihood uncertainty estimation method (GLUE; Beven and Binley 1992) is implemented to account for parametrically and structurally different hydrological models (i.e., multiparameter and multimodel). The individual model's likelihood measure is used to weight the model prediction for the ensemble simulation in GLUE (Beven and Binley 1992).

The calibrated parameters for the 28 watersheds (Fig. 1) and the three selected models are taken from Bastola and Misra (2013). Bastola and Misra (2013) calibrated the three conceptual models for the period 1948–1968 using CPC rainfall data. The hydrological models were then validated for the period of 1969–1979. The GLUE method was used to calibrate a suite of models based on a selected threshold that differentiates good and poor models according to selected likelihood measures. They used prediction interval (width of the prediction interval), median model performance, and count efficiency (percentage of observation points lying within the prediction interval) to demonstrate the performance of the selected hydrological model. Bastola and Misra (2013) also reported that on average (across watersheds), the performance of the model measured in terms of Nash Sutcliffe efficiency (NSE) is 0.72 and the average volume error is around 10% for both calibration (20-year block) and validation (10-year block). Furthermore, the large fractions of observation points were well encapsulated within the simulated prediction interval. Therefore, the same set of model parameters are used as calibrated model parameters in this study. The output from these calibrated hydrological models forced with CPC rainfall data is used as control or truth for verification of the hydrological forecasts.

The seasonal hydrological forecast experiments for a six-month period (June–November) are initialized in summer (first week of June) for a period of 20 seasons (1982–2001) that coincides with the hydrometeorological data available for verification from MOPEX. The seasonal hydrological forecast experiment in this study is carried out following the ESP framework (Resamp hereafter; Wood and Lettenmaier 2006). ESP entails initializing the hydrological models by forcing them with observed meteorological forcing up to the start of the forecast. Shukla and Lettenmaier (2011) and Wood and Lettenmaier (2006) suggested that initialization of hydrological models is important, especially for short-lead hydrologic forecasts such as those attempted here.

5. Analysis of forecast skill

Performance evaluation is an important part of any forecast experiment. There are many examples of objective skills scores, which measure the quality of a forecast with respect to, say, climatology (e.g., Murphy 1988; Wilks 2001). The deterministic accuracy measures, such as mean average error, mean squared error, or normalized root mean square [e.g., Nash Sutcliffe efficiency (NSE)] have also been commonly used to assess forecast skill. The NSE (Eq 4) is used as a skill score metric in this study. The skill of FISH50 seasonal prediction is assessed with respect to two reference forecasts: (a) climatology and (b) persistence (one-year lag).

$$NSE = 1 - \frac{\sum_{i=1}^{n}(A_{obs,i} - A_{pred,i})^2}{\sum_{i=1}^{n}(A_{obs,i} - A_{\mathrm{Re}f,i})^2} \quad \text{where } Q_{\mathrm{Ref,i}} = \begin{cases} \overline{A_{obs}} & \text{(climatological forecat)} \\ A_{obs,i-1} & (\textit{One} \text{ year lag foreast)} \end{cases} \qquad 4$$

The NSE is a deterministic skill metric that uses the ensemble mean and ignores the ensemble spread, a measure of the forecast uncertainty. In the above equation (4), the observed streamflow for a given year I ($A_{obs,i}$) is obtained from the control simulation of the individual hydrological models, which is forced with the observed meteorological forcing and then averaged across the three model estimates using GLUE. Similarly, $A_{pred,i}$ is the predicted streamflow from each of the three hydrological models forced with a variant of FISH50 (e.g., FISH50_BC, FISH50_Resamp) and then averaged using GLUE.

Brier score (Wilks 1995), ranked probability score (Wilks 1995), and relative operating characteristic or receiver operating characteristic curves (ROC) (Wilks 2001), are a few widely used probabilistic skill scores. These measures show the skill of forecast in discriminating occurrence and non-occurrence of events (Zhang and Casey 2000; Wilks 2001). ROC, which is defined as the plot of sensitivity to specificity, is constructed from a plot of probability of hit rate against probability of false alarm rate for a given probability threshold. The area under the ROC (AROC) is a probabilistic forecast skill metric (Marzban 2004). The AROC values range between 0 and 1, with 1 being a perfect forecast and AROC values ≤ 0.5 regarded as no better than climatology.

6. Results and discussion

i) Skill analysis: Deterministic

In this section we focus on the multiensemble member mean of FISH50 and the likelihood weighted average of the hydrological multimodel mean based on GLUE (multimodel mean hereafter). The probabilistic skill is examined in the subsequent subsection.

Over the 28 watersheds, the estimates of spatially averaged climatological rainfall (averaged from June through November) from FISH50 are significantly higher than the corresponding observations (Fig. 3). On average, over most of these SEUS watersheds, FISH50 overestimates the summer and fall rainfall total by nearly 65%. For hydrological applications, this error could prove to be grave. For example, in Fig. 4 it can be seen that using FISH50 rainfall produces extremely large biases in the forecasted streamflow. It should be mentioned that the volume error is computed with respect to the multimodel mean of the streamflow in the control simulation (i.e., flow simulated with the CPC precipitation dataset). In comparison, the bias-corrected FISH50 rainfall (FISH50_BC) and FISH50_Resamp appreciably reduce the bias. The results for FISH50 and FISH50_BC are based on average output from six ensemble members of FISH50 and the multimodel mean (based on GLUE) from three hydrological models. However, the result for FISH50_ Resamp is based on six-ensemble members of FISH50, the three hydrological models, and 10 realizations of observed resampling from historical records for each ensemble member. This means that there are 180 realizations of streamflow prediction for every season using the FISH50_Resamp as opposed to 18 realizations in FISH50 and FISH50_BC forcing.

The elasticity of rainfall on streamflow (calculated as proportional change in the mean annual streamflow to proportional change in mean annual rainfall) varies from 1.5 to 3.5 with an average value of 2.7. This is consistent with the estimates of the elasticity of rainfall on the streamflow for the SEUS region from Sankarasubramanian et al. (2001).

Although we examine only six months of the season, most of the annual mean rainfall in the SEUS occurs in the boreal summer and fall seasons. Therefore, given this large elasticity, the use of the raw FISH50, with its large bias in rainfall, may amplify the errors further when it is propagated through a hydrological model of a watershed. For the sake of brevity, in Fig. 5 we show the monthly mean climatological streamflow for six watersheds distributed across our study region (two in Florida, two in Georgia, one in North Carolina, and one in South Carolina). These watersheds have been chosen across latitude bands in the SEUS domain (Fig. 1) that exhibit the largest contrasts of seasonality in rainfall (Misra and DiNapoli 2012) and therefore are most representative of the SEUS region. The bias in predicted flow measured with respect to control simulation (i.e., flow simulated with the reference precipitation dataset) is significantly large over the selected watersheds irrespective of their latitudinal location. This bias is greatly reduced both in FISH50_Resamp and FISH50_BC. Interestingly, the volume errors of FISH50_Resamp are comparable to those of FISH50_BC. In other words, this result suggests that despite the large wet bias in FISH50, the seasonal climate state as defined by the precipitation anomaly (in quartile categories) is skillful enough to isolate observed analogues that can yield relatively more realistic rainfall time series for these watersheds than the raw FISH50 can.

Figure 6 shows the normalized root mean square errors of the ensemble average streamflow based on two reference forecasts: the climatological (referred to as NSE hereafter) and the one-year lag forecast [or persistent forecast, wherein the forecast from the previous year is persisted through the following year, defined as the persistence efficiency measure (PEM) hereafter]. With respect to the climatological forecast, the flow simulated with FISH50_BC has little or no skill (i.e., NRMSE has a value less than or equal to zero in Eq 4, which means that the forecast is inferior to the reference forecast). FISH50 and FISH50_Resamp forcing are equally poor (not shown). However, FISH50_BC and FISH50_Resamp exhibit relatively higher skill with respect to persistence (Fig. 6). In fact in many of the watersheds, PEM_FISH50_Resamp shows higher skill than PEM_FISH50_BC (Fig. 6). But there is no systematic decrease in NRMSE with increase in lead time of the forecast[1] to suggest a significant impact of the initial conditions on the forecast errors. The results in Fig. 6 suggest that the observed (or control simulation) variability of streamflow is very large, which makes the climatological forecast of streamflow superior to the persistence forecast. However, the deterministic skill of streamflow forecast using variants of FISH50 forcing (FISH50, FISH50_BC, FISH50_Resample) is unskillful, which is not what we expected.

ii) Skill analysis: Probabilistic

FISH50 shows some skill in discriminating different (quartile) categories of the seasonal (June–November) rainfall. The bubble plot in Figure 6 shows the distribution of AROC (with the size of the circles representing relative values above 0.5) for the seasonal mean (June–November) rainfall over each of the 28 watersheds. The latitudinal gradient of AROC in Fig. 7 is likely artificial as there are more watersheds in the northern latitudes of the SEUS than in the southern latitudes. It is apparent in Fig. 7 that the extreme

22

[1] In Fig. 6, June has zero lead while July has a one-month lead and so on till November, which has a five-month lead.

quartiles (for both wet and dry seasonal anomalies) have higher skill than the medium quartile seasonal rainfall categories. In fact, this result is consistent with the seasonal prediction skill of the tercile rainfall anomalies analyzed in Misra et al. (2012). However, there are exceptions, such as the watersheds in Florida (e.g., the St Johns and the Peace River watersheds), which show higher skill for the medium quartile categories of seasonal rainfall anomalies than for the extreme quartile categories. But more watersheds display higher skill for the extreme quartiles than for the middle quartiles (Fig. 7). Most of the watersheds in Georgia, North Carolina, and South Carolina show higher skill than climatology in extreme quartiles. Of the 28 SEUS watersheds, FISH50 rainfall shows skill (better than climatology) in the very wet (18 watersheds), medium wet (16 watersheds), medium dry (13 watersheds), and very dry (18 watersheds) rainfall categories.

The 180 realizations of FISH50_Resamp for each watershed are used to calculate the AROC to evaluate the experimental hydrological streamflow forecasts probabilistically. The AROC values for the predicted seasonal (June–November) streamflow with FISH50_Resamp are shown in Figs. 8 and 9. The results in Figs. 8 and 9 are summarized in Fig. 10. In Figs. 8–10 it is clear that AROC is better for the extreme quartiles than for the middle quartiles. The AROC for upper medium (i.e., wet medium) is the least skillful. Considering that the SEUS watersheds have high elasticity of rainfall on streamflow, the high value of the AROC for the extreme quartile of streamflow can be attributed to the fact that the majority of the watersheds in the SEUS showed similar rainfall skill for extreme wet quartiles. Contrary to our expectation that skill would decrease with lead-time, the relationship between skill and lead-time is not apparent in Figs. 8, 9, and 10. The comparable AROC values for streamflow are consistently less than 0.5 for FISH50 and FISH50_BC (not shown). These results suggest that despite the bias in FISH50 rainfall, it could be usefully applied for predicting the extreme quartiles, while for medium categories the forecasts remain more uncertain.

7. Conclusion

In this study we have examined the deterministic and probabilistic skill of seasonal hydrological predictions over 28 watersheds in the SEUS forced by an extensive set of retrospective seasonal climate forecasts known as FISH50.

It is observed that FISH50 exhibits a relatively large wet bias in June–November over all 28 watersheds in the SEUS. Because of the large elasticity exhibited by most of the 28 watersheds, the errors in rainfall are translated to a disproportionate response in streamflow. Therefore, direct application of FISH50 is not suggested for hydrological application in the SEUS. The deterministic skill analysis as depicted by the NSE reveals that the streamflow predictions forced with FISH50 have lower skill than those forced with observed climatological rainfall over all 28 watersheds. This result holds true irrespective of whether FISH50, FISH50_BC, or FISH50_Resamp is used to force the hydrological models.

FISH50_Resamp, however, offers a relatively large ensemble of streamflow estimation for every seasonal forecast from the climate model that potentially could be exploited for probabilistic forecasts. This large ensemble stems from the multiple hydrological models used in this study and the multiple (10 in our case) plausible matching observed analogues for the forecasted climate state (which in our case is the June–November rainfall quartile anomaly over the watershed). This gives rise to 180 independent estimations of streamflow per season, considering that there are six ensemble members of FISH50 for each season. Our study reveals that for extreme quartiles, there is significantly higher probabilistic streamflow skill from FISH50_Resamp than from the climatological forecast for the majority of the 28 watersheds in the SEUS.

In conclusion, this study provides evidence of useful probabilistic hydrological prediction over the SEUS watersheds during the boreal summer and fall seasons. However, the experiment yielded poor corresponding deterministic forecast skills. We contend that using a multimodel framework with multiple ensemble members for the meteorological forcing to generate a total of 180 forecasts per watershed, per forecast period allows for robust estimates of hydrological forecast uncertainty. This is portrayed in the probabilistic prediction skill, in contrast to its total neglect in the deterministic forecast prediction. Furthermore, probabilistic hydrological prediction augurs well when we realize that seasonal climate prediction is indeterminate, especially in the summer season when external factors like El Niño and the Southern Oscillation have little influence on the SEUS climate.

In other words, our study shows that despite the grave wet bias in FISH50, it still could be exploited for harvesting useful streamflow predictions over a majority of these watersheds. In fact, for the medium quartile categories, the uncertainty of the rainfall and streamflow forecast is higher, which can also serve as useful information for managing water resources. We believe that this is a significant result given the challenge of seasonal climate prediction of the boreal summer and fall seasons over the SEUS.

Acknowledgements

This work was supported by grants from NOAA (NA12OAR4310078, NA10OAR4310215, NA11OAR4310110), USGS (06HQGR0125), and USDA (027865). FISH50 integrations used in this paper were completed on the computational resources provided by the Extreme Science and Engineering Discovery Environment (XSEDE) under TG-ATM120017 and TG-ATM120010. We acknowledge Kathy Fearon of COAPS for the help with editing this manuscript.

References

Ajami, N.K., Q. Duan, X. Geo, and S. Sorooshian, 2006. MultiModel Combination Techniques for Analysis of Hydrological Simulations: Application to Distributed Model Intercomparison Project Results. J. Hydrometeor., 8:755-768.

Balmaseda, M., A. Vidard, and D. Anderson, 2008. The ECMWF System 3 ocean analysis system. Mon. Wea. Rev., 136: 3018–3034.

Balmaseda, M.A., and D. Anderson, 2009. Impact of initialization strategies and observations on seasonal forecast skill. Geophysical Research Letters 36, L01701, doi:10.1029/2008GL035561.

Bastola, S., Misra, V. 2013. Evaluation of Dynamically Downscaled Reanalysis Precipitation Data for Hydrological Application. Hydrol Processes. doi: 10.1002/hyp.9734.

Bastola, S., Murphy, C., Sweeney, J. 2011. Role of hydrological model uncertainty in climate change impact studies, Adv water Resour., 34(5): 562-576.

Beven, K. 2005. A manifesto for the equifinality thesis. J Hydrol. 20: 1-19.

Beven, K. and Binley, A. 1992. The future of distributed models: model calibration and uncertainty prediction, Hydrol Process., 6: 279-298.

Beven, K., P. Smith, I. Westerberg, and J. Freer. 2012. Comment on ‘‘Pursuing the method of multiple working hypotheses for hydrological modeling’’ by P. Clark et al., Water Resour. Res., 48, W11801, doi:10.1029/2012WR012282.

Bohn, T. J., M. Y. Sonessa, and D. P. Lettenmaier, 2010. Seasonal hydrologic forecasting: Do multimodel ensemble averages always yield improvements in forecast skill? J. Hydrometeor.,11: 1358–1372.

Broad, K., Pfaff, A., Taddei, R., Sankarasubramanian, A., Lall, U., de Assis de Souza Filho, F. 2007. Climate, stream flow prediction and water management in northeast Brazil: societal trends and forecast value Climatic Change, 84: 217–239.

Cazes-Boezio, G., D. Menemenlis, and C.R. Mechoso. 2008. Impact of ECCO ocean-state estimates on the initialization of seasonal climate forecasts. J. Climate 21:1929–1947.

Chan, S. Misra, V. 2010. A diagnosis of the 1979-2005 extreme rainfall events in the southeastern United States with Isentropic Moisture Tracing Mon. Wea. Rev., 138, 1172-1185.

Clark, M. P., Gangopadhyay, S., Brandon, D., Werner, K., Hay, L., Rajagopalan, B. and Yates, D. 2004. A resampling procedure for generating conditioned daily weather sequences, Water Resour. Res., 40, W04304, doi:10.1029/2003WR002747.

Clark, M. P., Hay, L. 2004. Use of Medium-Range Numerical Weather Prediction Model Output to Produce Forecasts of Streamflow. J., Hydrometeor, 5: 15–32.

Clark, M. P., D. Kavetski, and F. Fenicia, 2012. Reply to comment by K. Beven et al. on ‘‘Pursuing the method of multiple working hypotheses for hydrological modeling’’, Water Resour. Res., 48, W11802, doi:10.1029/2012WR012547.

Day, G.N. 1985. Extended Streamflow Forecasting Using NWSRFS. Journal of Water Resources Planning and Management. ASCE, 111(2): 157-170. doi: http://dx.doi.org/10.1175/JCLI3812.1

Derber, J., and A. Rosati, 1989. A global oceanic data assimilation system. J. Phys. Oceanogr., 19, 1333–1347.

Duan, Q., N.K. Ajami, X. Gao, and S. Sorooshian, 2006. MultiModel Ensemble Hydrologic Prediction Using Bayesian Model Averaging. Adv Water Resour., 30:1371-1386.

Duan, Q., Sorooshian, S., and Gupta, V.K. 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour Res., 28(4): 1015– 1031.

Efron, B. 1979. Bootstrap methods: Another look at the jackknife, Ann. Stat., 7: 1–26.

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. Hydrometeorol., 4: 1105–1118.

Gent PR, G Danabasoglu, LJ Donner, et al. 2011. The Community Climate System Model Version 4. J Clim 24: 4973-4991

Gent PR, SG Yeager, RB Neale, et al. 2010. Improvements in a half degree atmosphere/land version of the CCSM. Clim Dyn., 34:819-833

Gupta, H.V., Beven, K., and Wagener, T. 2003. In Model Calibration and Uncertainty Estimation, Anderson MG (ed.). John Wiley and Sons:Chichester, United Kingdom.

Hansen, J. W., A. W. Hodges, and J. W. Jones, 1998. ENSO influences on agriculture in the southeastern United States. J. Climate., 11: 404–411.

Higgins RW, Shi W, Yarosh E, Joyce R. 2000. Improved United States precipitation quality control system and analysis. NCEP/ CPC ATLAS No. 7. Also available at:

http://www.cpc.ncep. noaa.gov/research_papers/ncep_cpc_atlas/7/index.html

Kirtman, B. P. and D. Min, 2009. Multimodel Ensemble ENSO prediction with CCSM and CFS. Mon. Wea. Rev., 137: 2908-2930.

Krzysztofowicz, R. 2001. The case for probabilistic forecasting in hydrology, J. Hydrol., 249: 2–9.

Kuczera, G. 1997. Efficient subspace probabilistic parameter optimization for catchment models. Water Resour Res., 33(1): 177–185.

Li, H., Sheffield, J., and Wood, E.F. 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, J. Geophys. Res., 115, D10101, doi:10.1029/2009jd012882.

Lobell, D.B., Field, C. B., Cahill, K. N. and Bonfils, C. 2006. Impacts of future climate change on California perennial crop yields: model projections with climate and crop uncertainties Agric. Meteorol. 141: 208–18

Madsen, H. 2000. Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. Journal of Hydrology 235: 276–288.

Marzban, C. 2004. The ROC Curve and the Area under It as Performance Measures. Wea. Forecasting, 19: 1106–1114. doi: http://dx.doi.org/10.1175/825.1

Maraun, D. 2013. Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue. J. Climate, 26: 2137–2143.

Mcenery, John, John Ingram, Qingyun Duan, Thomas Adams, Lee Anderson. 2005. NOAA'S Advanced Hydrologic Prediction Service: Building Pathways for Better Science in Water Forecasting. Bull. Amer. Meteor. Soc., 86: 375–385.

Misra, V. and S. DiNapoli, 2012. Understanding the wet season variations over Florida. Clim. Dyn. In press. Available from http://coaps.fsu.edu/~vmisra/wetFL.pdf

Misra, V., H. Li, Z. Wu, S. DiNapoli, and Z. Wu. 2012. Global Seasonal Climate Predictability in a Two Tiered Forecast System. Part I: Boreal Summer and Fall Seasons. Climate dynamics (submitted) (available from http://floridaclimateinstitute.org/images/document_library/publications/fish50-paper-partI-final.pdf)

Murphy, Allen H. 1988. Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. American Meteorological Society -- Monthly Weather Review (Allen Press) 116 (12): 2417–2424

Voisin, N., Pappenberger, F., Lettenmaier, D.P., Buizza, R., Schaake. J. 2011. Application of a Medium-Range Global Hydrologic Probabilistic Forecast Scheme to the Ohio River Basin. Wea. Forecasting, 26: 425–446. doi: http://dx.doi.org/10.1175/WAF-D-10-05032.1

Pagano, T. C., D. C. Garen, T. R. Perkins, and P. A. Pasteris. 2009. Daily updating of operational statistical seasonal water supply forecasts for the western US, J. Am. Water Resour. Assoc., 45(3): 767–778.

Palmer, T. N., Alessandri, A., Andersen, U. et al. 2004. Development of a European multimodel ensemble system for seasonal to interannual prediction (DEMETER). Bull. Amer. Meteor. Soc., 85: 853-872.

Paolino, D. A., J. L. Kinter III, B. P. Kirtman, D. Min, D. M. Straus, 2012. The impact of land surface initialization on seasonal forecasts with CCSM. J. Climate*,* **25**: 1007-1021.

Prudhomme, C., and H. Davies. 2009. Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1: Baseline climate, Clim. Change., 93: 177–195.

Ropelewski CF, Halpert MS. 1986. North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). Mon Wea Rev 114:2352–2362

Ropelewski CF, Halpert MS. 1987. Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. Mon Wea Rev 115:1606–1626

Rosati, A., K. Miyakoda, and R. Gudgel, 1997. The impact of ocean initial conditions on ENSO forecasting with a coupled model. Mon. Wea. Rev., 125: 754–772.

Rosenberg, E. A., A. W. Wood, and A. C. Steinemann. 2011. Statistical applications of physically based hydrologic models to seasonal streamflow forecasts, Water Resour. Res., 47, W00H14, doi:10.1029/2010WR010101.

Saha, S., and co-authors, 2010. The NCEP Climate Forecast Reanalysis. Bull. Amer. Meteor. Soc., 91: 1015-1057.

Saha, S., and Coauthors, 2006. The NCEP Climate Forecast System. J. Climate, 19: 3483–3517.

Sankarasubramanian, A., Vogel, R.M., Limbrunner, J.F.. 2001. The climate elasticity of streamflow in the United States Water Resour. Res. 37(6): 1771-1781

Schaake, J., Cong, S., and Duan, Q. 2006. U.S Mopex Datasets, IAHS publication series (https://e-reports-ext.llnl.gov/pdf/333681.pdf).

Seager R., A. Tzanova, and J. Nakamura, 2009. Drought in the southeastern United States: Causes, variability over the last millennium, and the potential for future hydroclimatic change. J. Climate, 22: 5021–5045.

Semenov MA, Barrow EM. 1997. Use of Stochastic Weather Generator in the Development of Climate Change Scenarios. Climatic Change 35: 397–414.

Shukla, S. and Lettenmaier, D. P. 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.

Stefanova L, Misra V, O'Brien J J, Chassignet E P and Hameed S. 2012a. Hindcast skill and predictability for precipitation and two-meter air temperature anomalies in global circulation models over the Southeast United States. Climate Dynamics 38: 161-173.

Stefanova, L., Misra, V., Chan, S., Griffin, M., O'Brien J. J., and SmithIII, T. J. 2012b. A proxy for high-resolution regional reanalysis for the southeast united states: assessment of precipitation variability in Dynamically downscaled reanalyses. Climate Dynamics 38: 2449-2446. DOI:10.1007.s00382-011-1230-y

Steinschneider, S., A.Polebitski, C.Brown, and B. H.Letcher. 2012. Toward a statistical framework to quantify the uncertainties of hydrologic response under climate change, Water Resour. Res., 48, W11525, doi:10.1029/2011WR011318.

Sugawara, M. 1995. Tank model. In: Singh VP, editor. Computer models of watershed hydrology. Littleton, Co.: Water Resource Publication: 165–214.

Tingsanchali, T. and Gautam, MR. 2000. Application of Tank, NAM, ARMA and Neural

Network Models to Flood Forecasting. Hydrol. Processes, 14:2473-2487.

Wagener T., Boyle DP., Lees MJ., Wheater HS., Gupta H. V. and Sorooshian S. 2001. A framework for development and application of hydrological models. Hydrol. Earth Syst. Sci., 51: 13–26.

Wilks, D. S.. 2002. Realizations of Daily Weather in Forecast Seasonal Climate. J. Hydrometeor., 3: 195–207.

Wilks, D. S., 1995. Statistical Methods in the Atmospheric Sciences. Cambridge Press. 547 pp

Wilks, D.S. 2001. A skill score based on economic value for probability forecasts. Meteorol. Appl., 8: 209-219.

Wood AW, Leung LR, Sridhar V, Lettenmaier D. P. 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change 62: 189–216.

Wood, A. W., A. Kumar, and D. P. Lettenmaier. 2005. A retrospective assessment of National Centers for Environmental Prediction climate model–based ensemble hydrologic forecasting in the western United States. J. Geophys. Res., 110, D04105, doi:10.1029/2004JD004508.

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

Wood, A. W., E. P. Maurer, A. Kumar, and Lettenmaier. D.P. 2002. Long-range experimental hydrologic forecasting for the eastern United States. J. Geophys. Res., 107(D20), 4429, doi:10.1029/2001JD000659.

Xue, Y., and co-authors, 2011. Comparative analysis of upper ocean heat content variability from ensemble operational ocean analysis. U.S. CLIVAR, 9 (1): 7-10.

Yang, S. -C., C. Keppenne, M. Rienecker, E. Kalnay, 2009. Application of coupled bred vectors to seasonal-to-interannual forecasting and ocean data assimilation. J. Climate, 22: 2850-2870.

Yao, H. and Georgakakos, A. 2001. Assessment of Folsom Lake Response to Historical and Potential Future Climate Scenarios. J Hydrol, 249: 176-196.

Zhang H, Casey T. 2000. Verification of categorical probability forecasts. Weather Forecasting, 15: 80–89.

Zhang, S., M.J. Harrison, A. Rosati, and A. Wittenberg, 2007. System design and evaluation of coupled ensemble data assimilation for global oceanic studies. Monthly Weather Review 135:3541–3564.

List of Tables

List of Figures

Fig. 3. Climatological rainfall averaged over the 28 watersheds in the Southeastern US (identifiers indicated along the *x*-axis).

Fig. 4. Volume error of the flow predicted with FISH50 daily data as are (FISH50), bias-corrected FISH50 (FISH50_BC), and resampling from observed historical record on the basis of the closest analogue of the forecasted climate state (FISH50_Resamp)

Figure 5: Flow predicted with FISH50 daily data as are (FISH50), bias-corrected FISH50 (FISH50_BC), and resampling from observed historical record on the basis of the closest analogue of the forecasted climate state (FISH50_Resamp.

Fig. 6. Skill scores of the hydrological prediction based on normalized root mean square errors. PEM (persistence efficiency measure) is the Nash Sutcliffe efficiency criteria measured with respect to a one-year lag as the reference forecast. NSE is the Nash Sutcliffe efficiency measured with respect to climatological value as the reference forecast.

Fig. 7. Area under ROC (AROC) for FISH50 June–November mean precipitation averaged over the respective watersheds in the Southeastern US. (a) AROC value for very wet (blue circle) and very dry (red circle) rainfall categories, (b) AROC value for medium wet (blue) and medium dry (red) categories. Only AROC values over 0.5 are shown. The size of the bubble indicates the relative magnitude of the AROC, which can range between 0.5 and 1.0. Values of AROC below 0.5 are not plotted.

Fig. 8. AROC for very wet (blue) and very dry (red) categories of streamflow predicted with FISH50_Resamp forcing. The size of the bubble indicates the relative magnitude of

the AROC, which can range between 0.5 and 1.0. Values of AROC below 0.5 are not plotted.

Fig. 9. Same as Fig. 8 but for AROC values for medium wet (blue) and medium dry (red) categories of streamflow.

Fig. 10. Summary of probabilistic assessment of flow predicted with FISH50_Resamp for four selected categories of June–November mean rainfall over the various watersheds in the Southeastern US (very dry, very wet, medium wet, and medium dry). AROC greater than 0.5 suggests that the prediction skill is better than climatology.

Table 1: A brief outline of FISH50

Number of years of the seasonal hindcast	Number of ensemble members per seasonal hindcast	Start date of the seasonal hindcast	End date of the seasonal hindcast
27 (1982–2008)	6	1 June	31 December

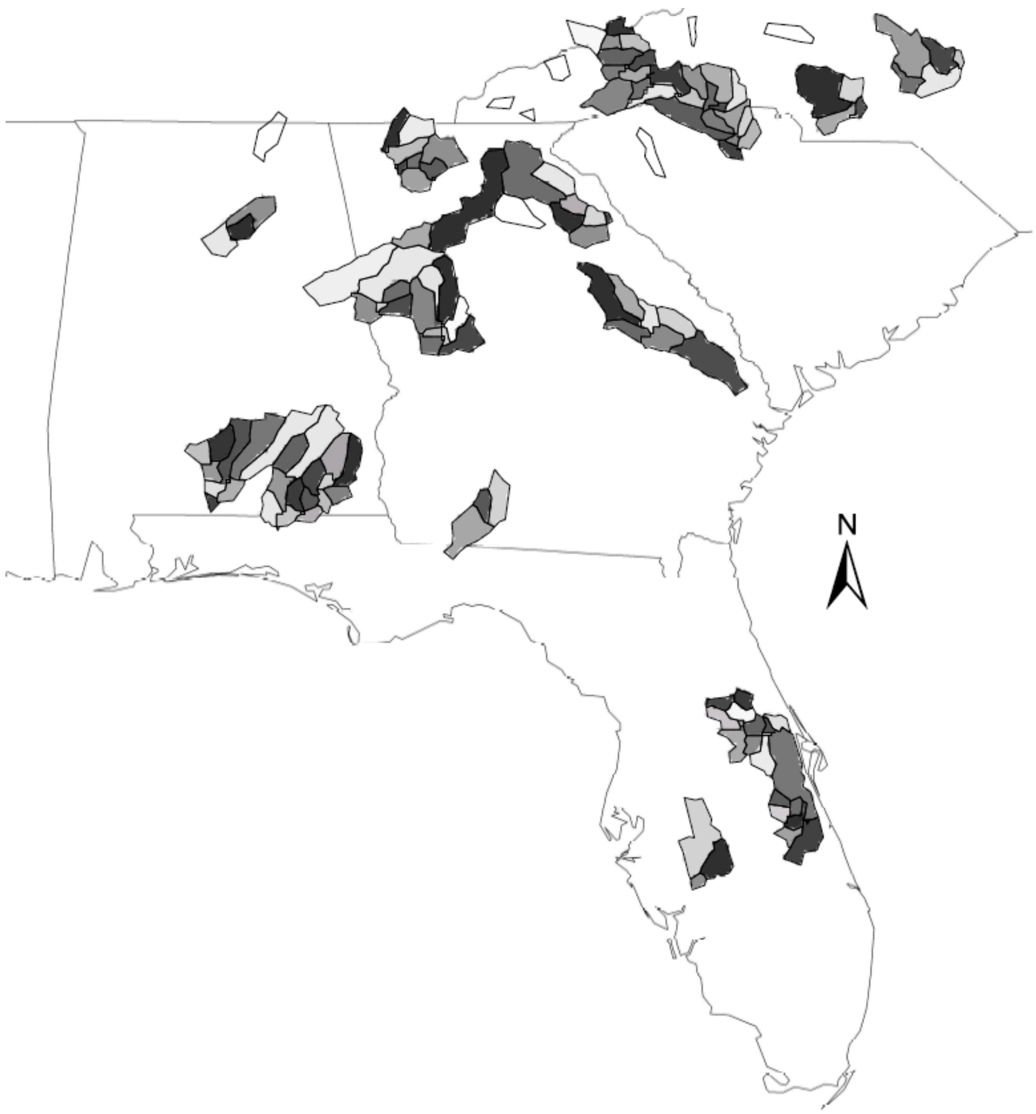

Fig. 1. Outline of the 28 watersheds in the Southeastern US for which the seasonal hydrological forecasts are made. The sub-basins within each watershed are also indicated.

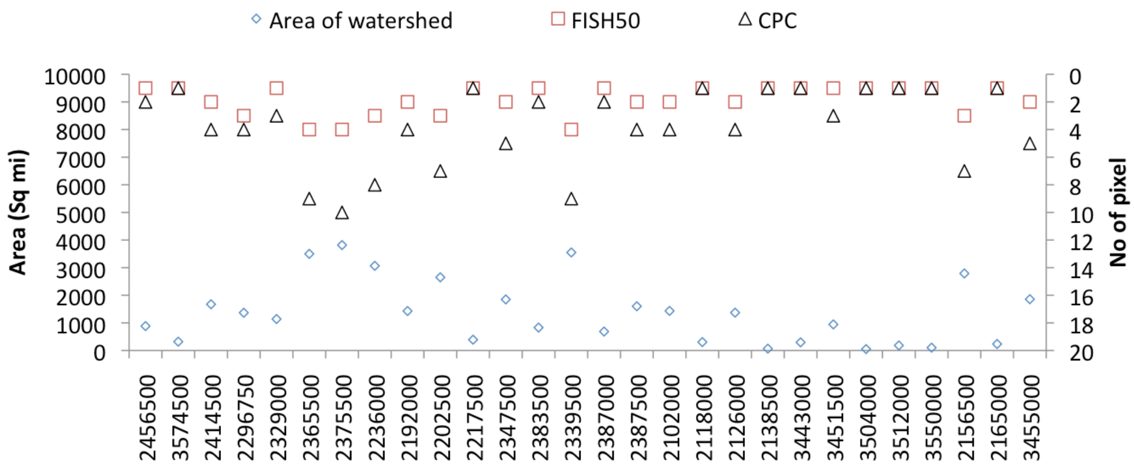

Fig. 2. Area and number of meteorological pixels contained in each watershed for FISH50 and CPC observed rainfall datasets.

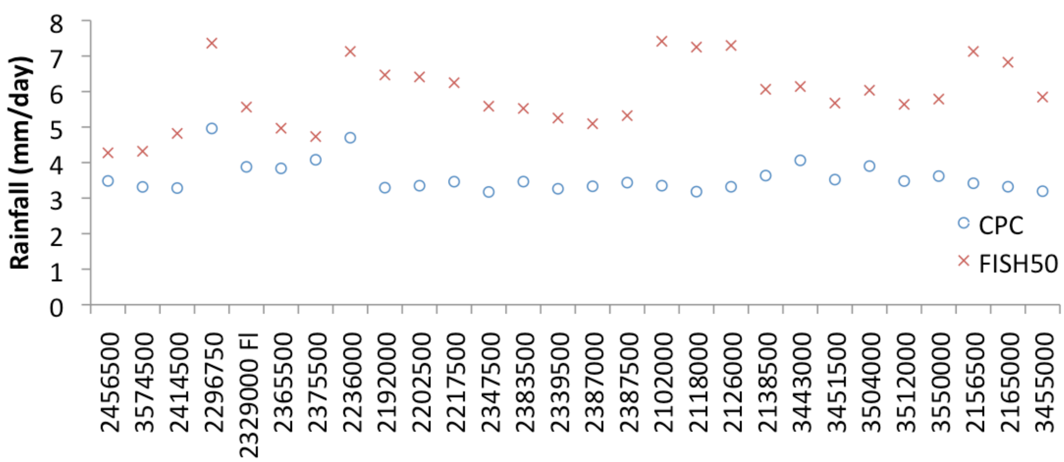

Fig. 3. Climatological rainfall averaged over the 28 watersheds in the Southeastern US (identifiers indicated along the *x*-axis).

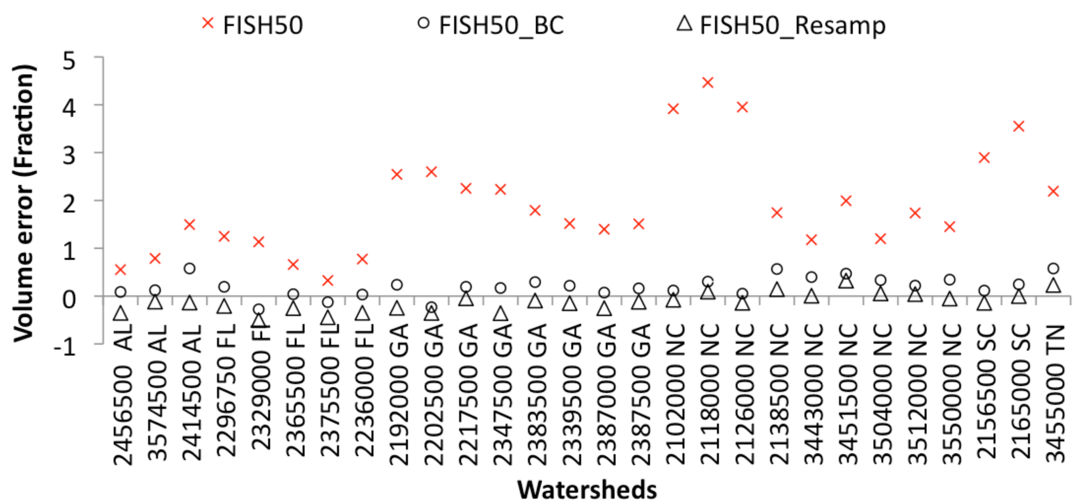

Fig. 4. Volume error of the flow predicted with FISH50 daily data as are (FISH50), bias-corrected FISH50 (FISH50_BC), and resampling from observed historical record on the basis of the closest analogue of the forecasted climate state (FISH50_Resamp)

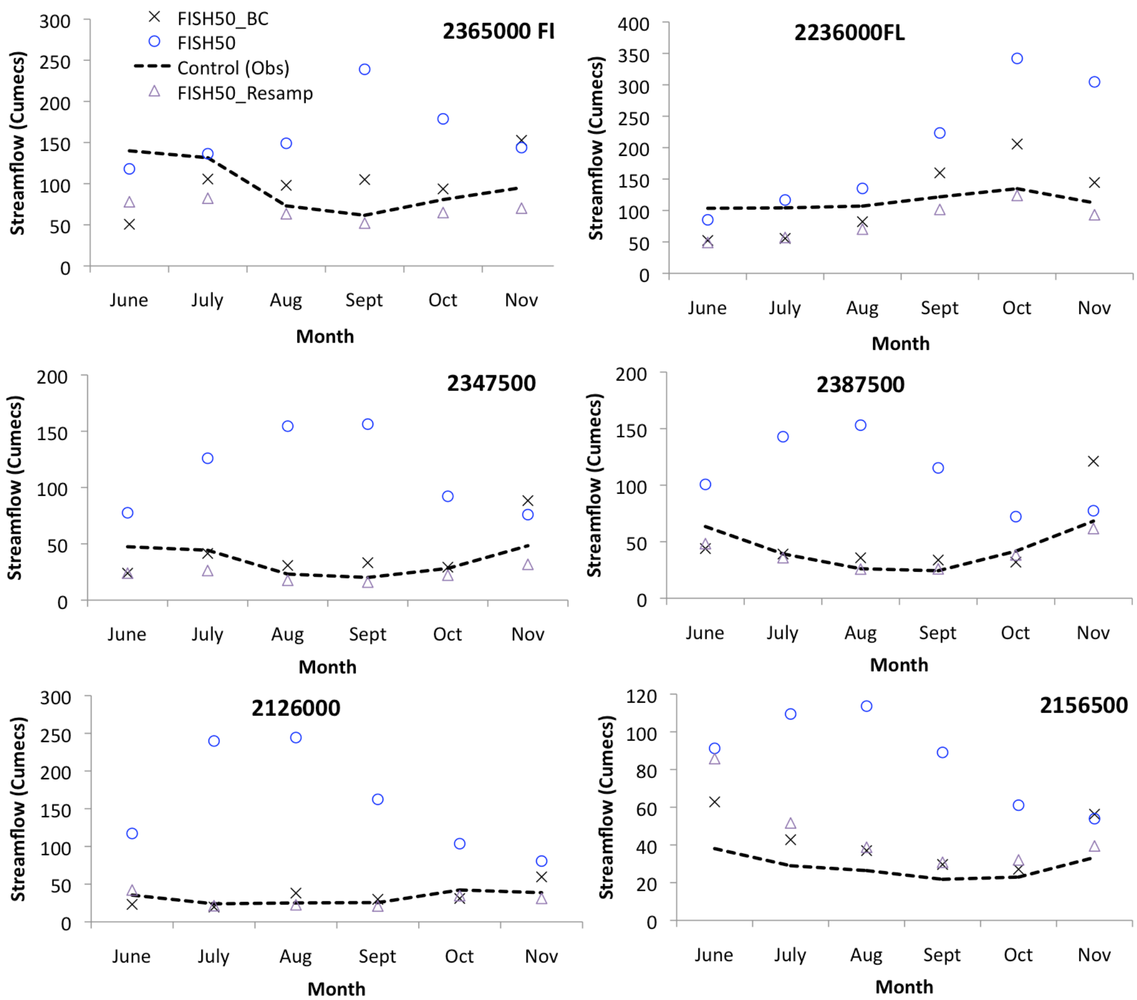

Figure 5: Flow predicted with FISH50 daily data as are (FISH50), bias-corrected FISH50 (FISH50_BC), and resampling from observed historical record on the basis of the closest analogue of the forecasted climate state (FISH50_Resamp.

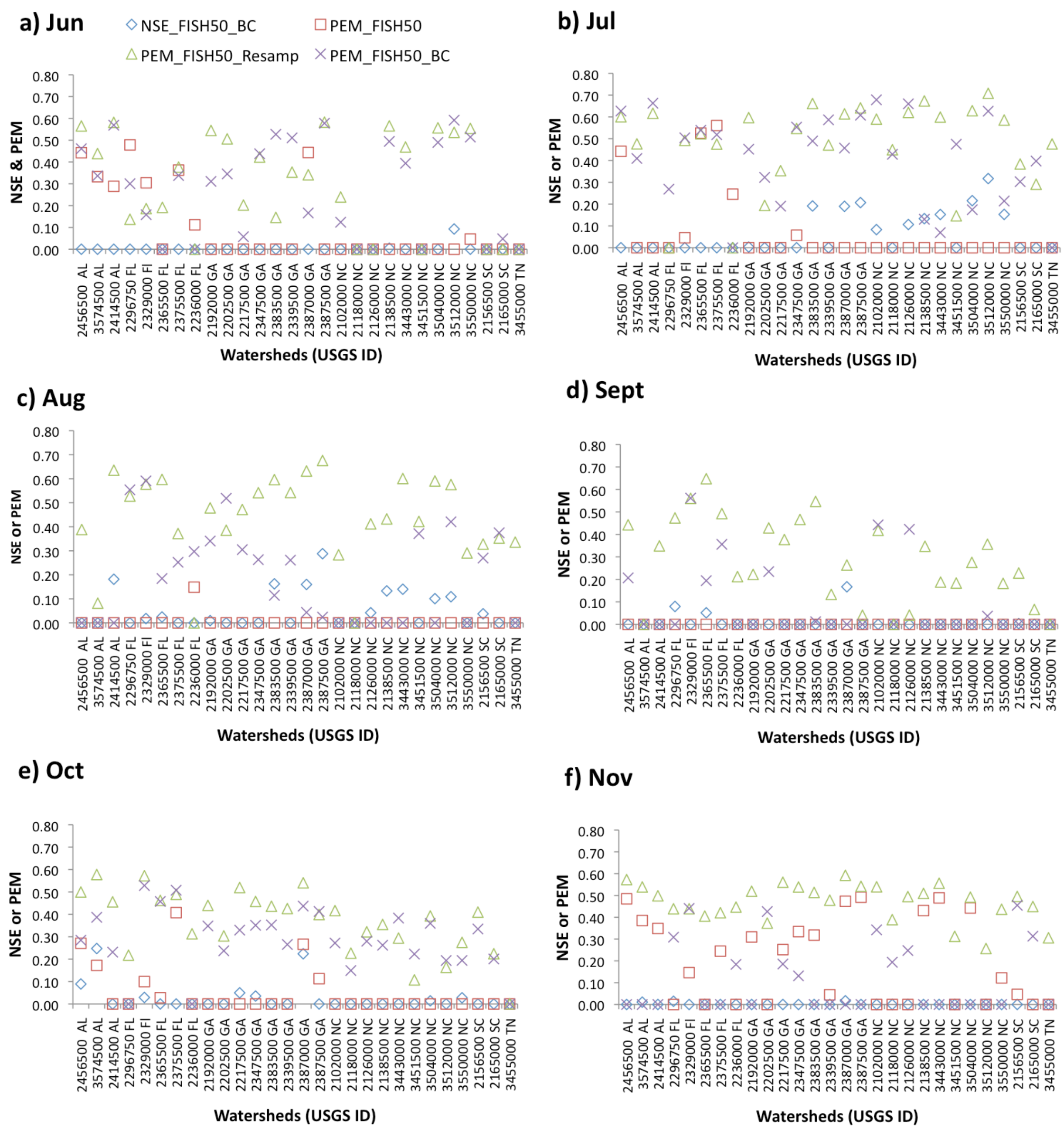

Fig. 6. Skill scores of the hydrological prediction based on normalized root mean square errors. PEM (persistence efficiency measure) is the Nash Sutcliffe efficiency criteria measured with respect to a one-year lag as the reference forecast. NSE is the Nash Sutcliffe efficiency measured with respect to climatological value as the reference forecast.

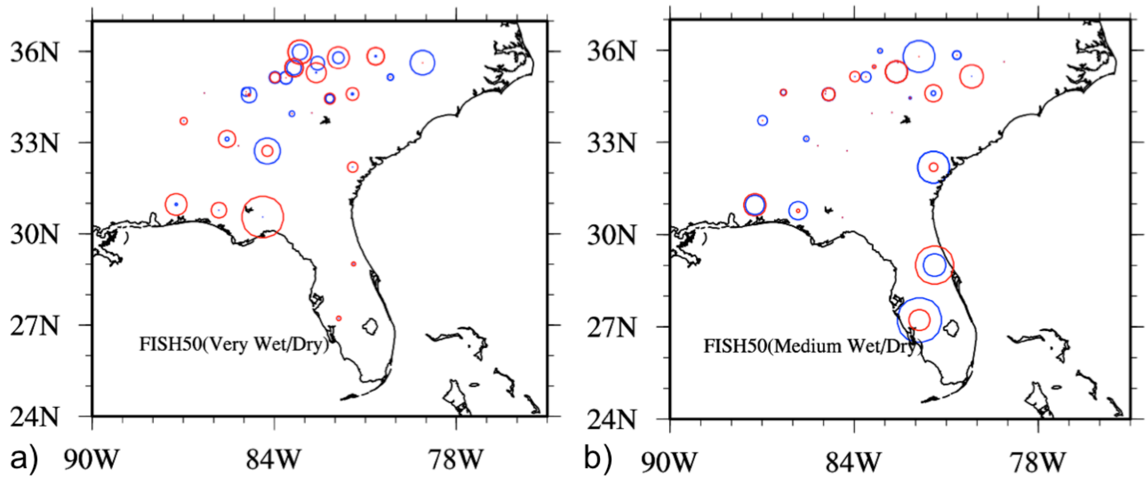

Fig. 7. Area under ROC (AROC) for FISH50 June–November mean precipitation averaged over the respective watersheds in the Southeastern US. (a) AROC value for very wet (blue circle) and very dry (red circle) rainfall categories, (b) AROC value for medium wet (blue) and medium dry (red) categories. Only AROC values over 0.5 are shown. The size of the bubble indicates the relative magnitude of the AROC, which can range between 0.5 and 1.0. Values of AROC below 0.5 are not plotted.

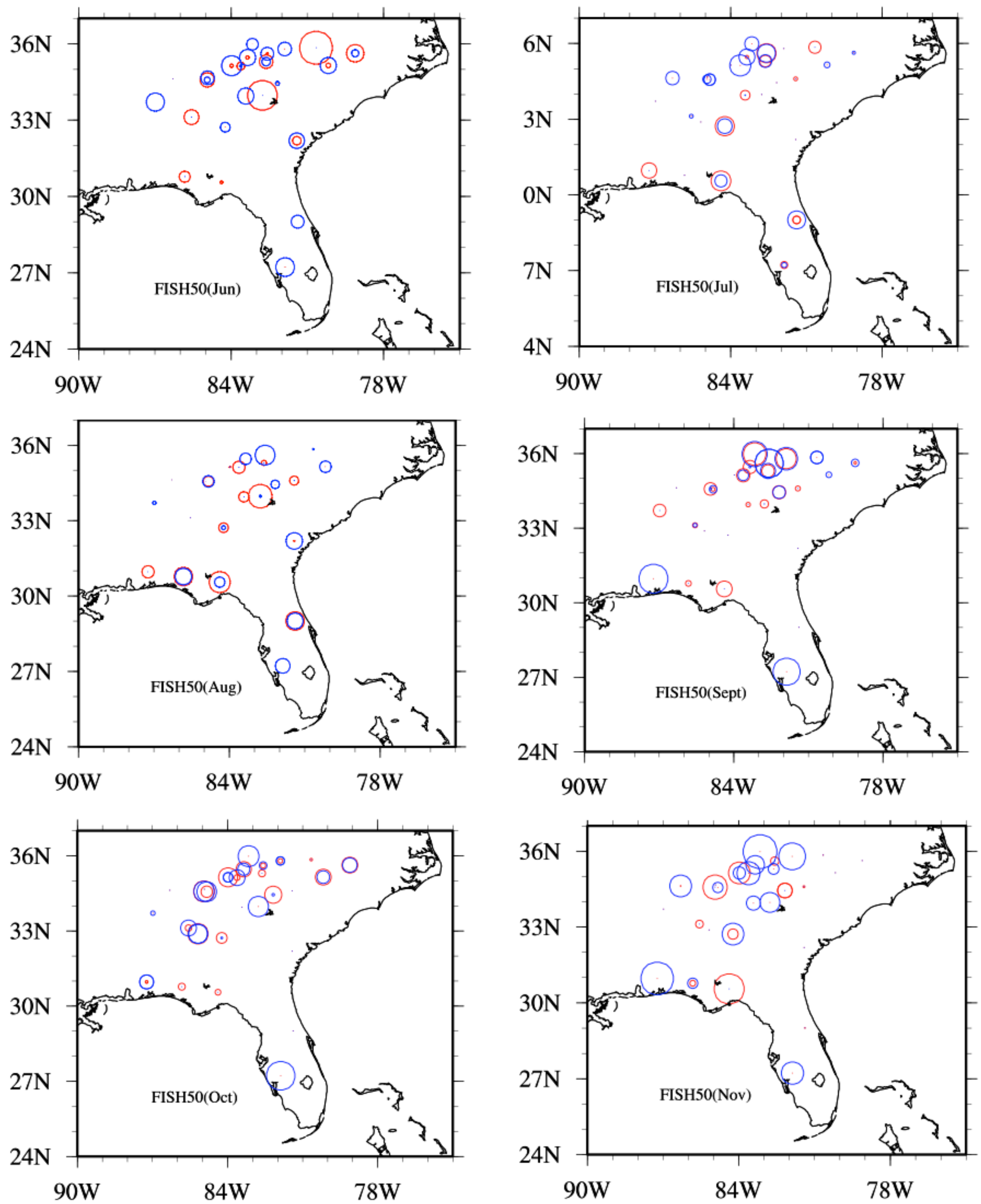

Fig. 8. AROC for very wet (blue) and very dry (red) categories of streamflow predicted with FISH50_Resamp forcing. The size of the bubble indicates the relative magnitude of the AROC, which can range between 0.5 and 1.0. Values of AROC below 0.5 are not plotted.

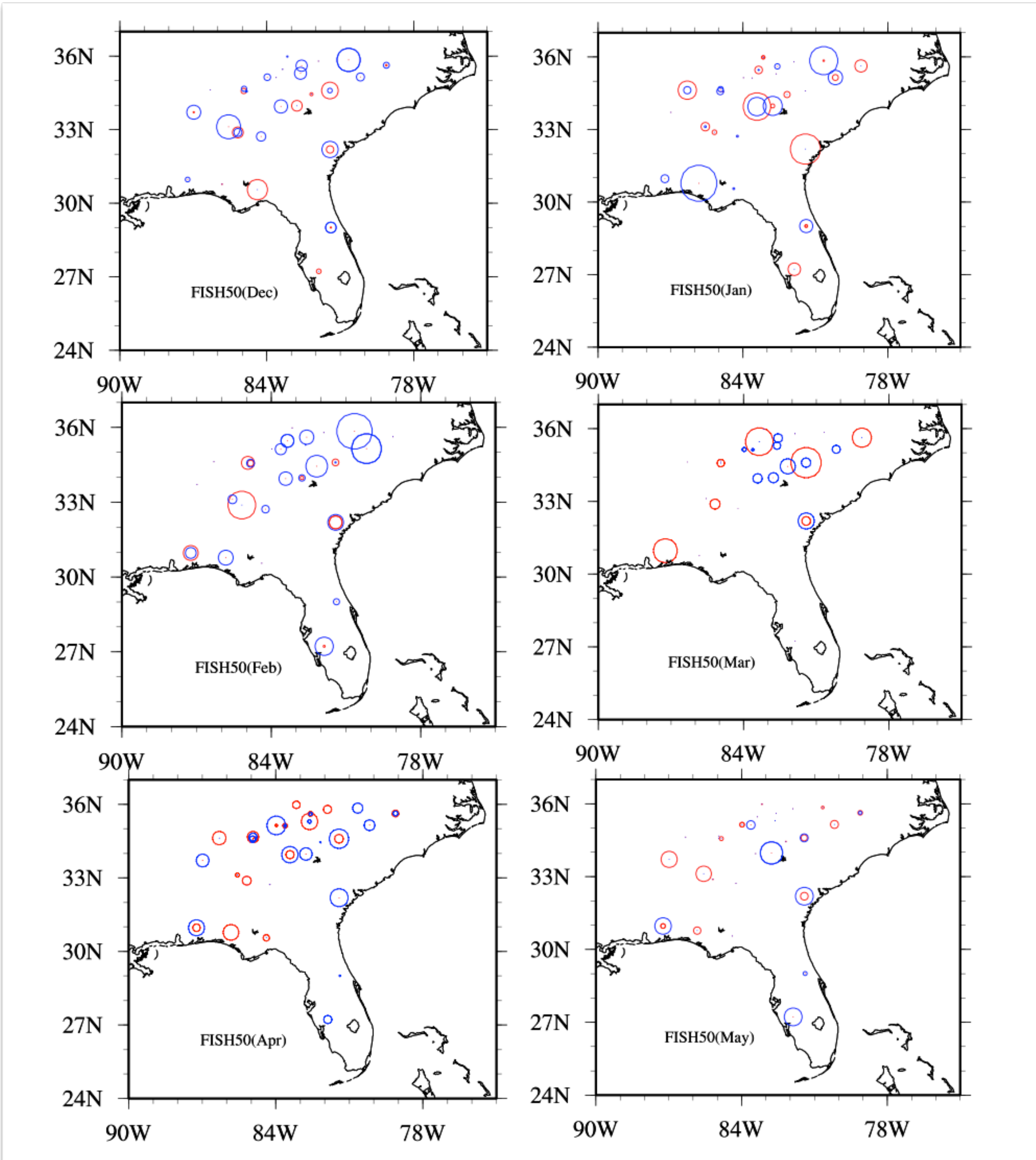

Fig. 9. Same as Fig. 8 but for AROC values for medium wet (blue) and medium dry (red) categories of streamflow.

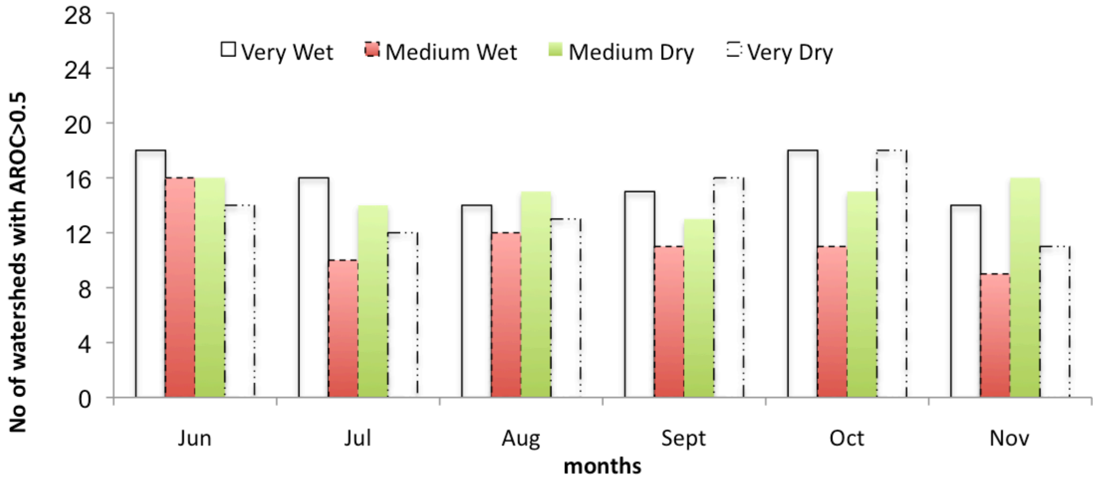

Fig. 10. Summary of probabilistic assessment of flow predicted with FISH50_Resamp for four selected categories of June–November mean rainfall over the various watersheds in the Southeastern US (very dry, very wet, medium wet, and medium dry). AROC greater than 0.5 suggests that the prediction skill is better than climatology.