Seasonal hydrological and nutrient loading forecasts for watersheds over the Southeastern United States

Satish Bastola[1,#] and Vasubandhu Misra[2,3,4]

[1] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia

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

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

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

[#] Corresponding author email: Satish.bastola@ce.gatech.edu

Abstract

We show useful seasonal deterministic and probabilistic prediction skill of streamflow and nutrient loading over watersheds in the Southeastern United States (SEUS) for the winter and spring seasons. The study accounts for forecast uncertainties stemming from the meteorological forcing and hydrological model uncertainty. Multi-model estimation from three hydrological models, each forced with an ensemble of forcing derived by matching observed analogues of forecasted quartile rainfall anomalies from a seasonal climate forecast is used. The attained useful hydrological prediction skill is despite the climate model overestimating rainfall by over 23% over these SEUS watersheds in December-May period. The prediction skill in the month of April-and May is deteriorated as compared to the period from December-March (zero lead forecast).

A nutrient streamflow rating curve is developed using a log-linear tool for this purpose. The, skill in prediction of seasonal nutrient loading is nearly identical to the skill in predicting the seasonal streamflow.

Keywords: Rainfall-runoff model, Seasonal Hydrologic Forecasting, Southeastern United States, Water Quality, Seasonal Predictability

Software availability

Name: Catchment hydrological model (subroutine written in FORTRAN)

Availability: Available upon request from the authors.

1. Introduction

The Southeastern United States (SEUS) region receives considerable amounts of rainfall ranging spatially between 30 and 100 inches annually relative to the rest of the United States (http://www.nc-climate.ncsu.edu/edu/k12/.SEPrecip). However, the water sector remains vulnerable because the region is exposed to significant climate variability including relatively frequent climate and weather extremes like droughts and landfalling tropical cyclones. There are several studies, which have suggested the benefit of streamflow predictions in managing and regulating water resources (e.g., Broad 2007; Yao and Grergakakos 2001; Obeysekera et al. 1999). For example, Obeysekera et al. (1999) noted the benefit of long-range hydrological forecasts for the complex water management system in South Florida, consisting of large reservoirs, lakes, and water regulating structures. However, Bolson et al. (2013) reported that only about 25% of the water managers use seasonal climate forecasts. Many studies have attributed this infrequent use to lack of awareness and difficulty in understanding, trusting, and applying the forecasts (e.g., Carbone and Dow 2005; Pagano et al. 2001).

The reliability of streamflow forecasts depends on, among other factors, the fidelity of the climate forecast, the reliability of the hydrological models, and the quality of the initial hydrologic conditions used. Over the years considerable progress has been made in improving dynamical seasonal prediction (Kumar et al. 1996; Krishnamurti et al. 1999; Palmer et al. 2004; Zhang et al. 2007; Yang et al. 2009; Kirtman and Min 2009; Saha et al. 2010; Gent et al. 2011; Misra et al. 2013; Li and Misra 2013; Kirtman et al. 2014). Furthermore, Maurer et al. (2004) claim that a better understanding of the teleconnections between large-scale climate features and the hydroclimatology of the region can improve the streamflow forecast for longer lead times. In the SEUS, the El Niño–Southern Oscillation (ENSO) teleconnections are relatively strong, especially during the boreal winter and spring seasons (Ropelewski and Halpert 1987, 1986; Kiladis and Diaz 1989). In a typical El Niño (La Niña) year, the SEUS experiences a cold and wet (warm and dry) winter season. Such robust teleconnections provide an opportunity to improve water resource management at the seasonal to interannual time scales.

To provide predictions of streamflow at long lead times, empirical methods (e.g., multiple linear regression) have long been used in the United States (e.g., Rosenberg et al. 2011; Pagano et al. 2009). These methods use initial conditions and information on future climate condition as predictors. However, ensemble streamflow prediction (ESP; Day 1985) is also being widely considered as an alternative to multiple linear regressions (e.g., Connelly et al. 1999; Franz et al. 2003; Wood et al. 2005; Wood and Schaake 2008; Bohn et al. 2010) and is adopted in this study as well. In the ESP method, multiple realizations of future streamflow are simulated. These realizations are usually generated by independently running multiple calibrated hydrological models forced with multiple realizations of surface meteorological forcing (discussed further in section 3). ESP has some notable advantages over empirical models: ESP makes predictions on physical and conceptual basis, provides an estimate of the forecast uncertainty and offers flexibility in using forcing data from different sources (e.g., climate model, subjective climate outlooks).

A study from NRC (2000) recognized that nutrient loading is the main cause of eutrophication of freshwater bodies and coastal estuaries. The stream nutrient loads depend upon, among other factors, precipitation, temperature, soil, geology, nutrient fate and transport; however, urban and agriculture landscape have the greatest impact (Preston et al. 2011). These sources of nutrient loading can be broadly classified as point source and nonpoint source. Managing and controlling these excessive nutrients relies on instream nutrient concentration data that are only sparsely available. Therefore, mathematical models are widely used in aiding local nutrient management and control. A broad array of models ranging from empirical models such as simple export coefficient to physically based nutrient modeling tools are available to estimate pollution and identify the sources at watershed scale (Shrestha et al. 2008; Arnold et al. 1994; EPA 1987; Smith et al. 1997; Ambrose et al. 1981; Johanson et al. 1981). Most of these studies show strong empirical evidence that streamflow is the single most important variable for estimating the pollution load of namely total nitrogen and total phosphorous. These nutrients are responsible for the impairment of many inland waters as well as coastal bays. Excessive amounts of these nutrients promote profuse algae growth, resulting in unhealthy inland and coastal waters.

Studying over 6300 water bodies located in Florida using a range of chemical and biological parameters, the Florida Department of Environmental Protection (FDEP) in the year 2008 found impairment in 28% of the stream miles, 25% of lake acres, and 59% of square miles of estuaries. The FDEP has recently imposed a numeric nutrient criteria water quality standard specifically for nitrogen and phosphorous (FDEP 2012). In response to adoption of such water quality standards, water quality credit trading is likely to emerge as one of the policy tools to preserve water quality in a cost-effective manner [e.g., pilot water quality credit trading program for the lower St. Johns River (FDEP 2010); establishment of pollutant trading policy advisory committee to assist FDEP in developing a pollutant trading program (FDEP 2006)]. Nutrient trading is especially beneficial in avoiding impairment of waterways and water bodies if the cost involved in controlling pollutants from the various sources within a watershed is considerably different and water quality goals are firmly established.

In this study, we analyze a relatively large set of retrospective seasonal streamflow forecast experiments for the boreal winter and spring months for watersheds in the SEUS using the seasonal climate forecasts produced by the Florida Climate Institute (FCI) of the Florida State University (FSU; Misra et al. 2013 and Li and Misra 2013). Li and Misra (2013) demonstrated that the FCI-FSU Seasonal Hindcasts at 50km grid resolution (FISH50) seasonal mean temperature and precipitation has comparable skill for boreal winter and spring relative to the operational models of the National Multi-Model Ensemble (NMME; Kirtman et al. 2014). FISH50 also offers the highest spatial resolution among the existing seasonal climate hindcast data sources (e.g., NMME). Therefore, through this study, we aim to explore the utility of FISH50 for seasonal hydrologic forecasts. We follow up this analysis with estimations of retrospective seasonal forecasts of nutrient loading, which are based on the empirical nutrient streamflow rating curve and the predicted streamflow. We consider only total nitrogen and total phosphorous as they are widely regarded as predictors of stream-water quality (USEPA 2006).

It may be noted that this is first of such attempt to apply retrospective multi-model seasonal hydrological forecast framework for the boreal winter and spring seasons over these relatively small 28 SEUS watersheds.

2. Study region and data

In this study, a total of 28 watersheds from the Model Parameter Estimation Experiment (MOPEX; Schaake et al. 2006) is selected, which follows from our previous work on the summer seasonal forecasts (Bastola et al., 2013). These watersheds are chosen because they are minimally affected by water management (Schaake et al. 2006). The characteristics of the selected watershed is shown in Table 1.

FISH50 is initialized in late November through early December of each year and integrated through May of the subsequent year for 1982–2008. Each of the seasonal hindscasts of FISH50 has a total of six ensemble members, which are generated by perturbing the initial conditions of the atmosphere (Li and Misra 2013). The data from FISH50 is available at daily time scale from December through May of the following year. It may be mentioned that in this forecast framework of FISH50 December-January-February (DJF) seasonal mean is at zero lead while the March-April-May (MAM) seasonal mean will be at one season lead. We use the unified daily US precipitation analysis of the Climate Prediction Center (CPC; Higgins et al. 2000), available at 0.25^0 grid resolution and from 1948 onward, as the observed rainfall.

3. Hydrological model forcing

Bias correction and the stochastic method are two common approaches that are widely used to bridge the spatial resolution gap that exists between the hydrological and the climate models. To correct for systematic biases in rainfall forecasts, the quantile-based bias correction method has been used extensively in hydrological applications (e.g., Li et al. 2010; Wood et al. 2004). The stochastic method is based on resampling from historical observations (e.g., the Schaake shuffle in Clark et al. 2004). The resampling from analogue years is based on categorical climate forecasts (e.g., forecasts based on tercile or quartile categories of seasonal precipitation anomalies). It preserves the various moments of a time series (e.g., Efron 1979).

In this study, bias correction based on resampling from historical observation is used to circumvent the issue of bias in FISH50. The resampling method to generate a conditioned daily sequence of meteorological forcing for the semi-distributed hydrological models (which includes sub basin average rainfall, temperature, and evapotranspiration) is as follows (see Bastola et al. (2014) for further details):

1. Based on 6-month averaged (Dec-May) forecast rainfall, derive the quartile category for each year for a given watershed,
2. Sample 10 sets of model forcing for hydrological model (a block of 6 month (Dec-May) from historical observation of weather data that has same quartile category as that of forecast seasonal mean (December-May) rainfall.
3. Repeat step 1 and 2 for each of the six-ensemble member of the FISH50.

This procedure generates 10 resamples for each ensemble member and then propagates them through three hydrological models. We thus obtain 180 (=6 ensemble members of FISH50 x 10 observed resamples per ensemble member of FISH50 x 3 hydrological models) estimates of streamflow for each watershed per season.

In this study, the FISH50 seasonal categorical rainfall forecast (based on quartile rainfall categories) is used to sample from the observed analogue. The resampling of the past observations is done several (10) times per ensemble member of FISH50 using the method of block resampling without replacement (Prudhomme and Davies 2009). Here we define a block as six months of continuous daily rainfall. Though plausible, resampling with a block size of a month or three months is likely to affect the seasonal structure and to introduce biases (Prudhomme and Davies 2009). Furthermore, ten resamples per ensemble member of FISH50 enables us to attain a good sample of the observed near analogues of the forecasted meteorological forcing to make a robust probabilistic streamflow forecast (Fig. 1). It may also be mentioned that this resampling from historical observed data also temporally disaggregates the seasonal forecast total into daily values. In this way, the resampling procedure generates multiple time series of rainfall from a historical record. In addition, semi-lumped [parameters are lumped over the whole river basin with spatial variation in the model (meteorological) forcing] hydrological models are implemented for the hydrological predictions conducted in this study. Therefore, development of a spatially coherent sub-basin average rainfall field is essential. In this context, resampling from historical observation allows development of spatially coherent sub-basin average rainfall (see Bastola et al., 2013).

In this study, we compare the performance of the hydrological forecasts between those that use the FISH50 meteorological forcing directly without any bias correction (named hereafter as FISH50) and those that use the resampled observations using quartile categories of seasonal mean rainfall from FISH50 (FISH50_Resamp). Evaluation of output from enviornmental modeling using suitable perfromance measure is essential before they can be confidently used for their practical application. Hydrologists fundamentally use qualitative (visual) and objective criteria to judge the reliability of output from hydrological simulation. The quantitative criteria most widely used is the efficiency criteria derived from summation of error term normalized by the variability in observation data (Beven and Binley 1992, Bennett et al., 2013). Two commonly used residual methods namely, persistence index and Nash-Sutcliffee Model effeciency (Table 6 of Bennett et al., 2013) is used to evaluate the seasonal hydrological forecasts in this study. Furthermore, receiver operating characteristics curves (ROC) is used to evaluate the probabilistic forecast, again one of the commonly used proabilistic skill metrics (Wilks 2001).

4. Experiment design

i) Hydrological forecasts

In this study, the seasonal hydrological forecast experiment, which is carried out using the FISH50 data for 20 years (1982–2001), is based on ESP methodology (Fig. 1). The FISH50 dataset is available for a six-month period from December through May of the subsequent year. The initial conditions for the hydrological models are obtained by forcing the hydrological models with the observed meteorological forcing up to the start (or initial) time of the forecast (e.g., Wood and Lettenmaier 2006).

Though there is consensus on the importance of uncertainty analysis in hydrological modeling and subsequently on water resource planning and management, there is an intense debate on the framework needed to quantify uncertainty (e.g., Beven et al. 2012; Clark et al. 2012; Kuczera, 1997; Beven and Binley 1992; Duan et al. 1992). The discussion and implementation of different methods is beyond the scope of this study. In this study, we account for the uncertainty of the hydrological forecasts arising from model uncertainty and meteorological forcing (Fig. 1). The latter form of (meteorological) uncertainty is estimated from the 60 ensemble members generated from 10 observed resamples for each of the 6 ensemble members of FISH50 per season (see Section 3). The model uncertainty is accounted for by combining the retrospective predictions derived from three conceptual rainfall-runoff (RR) model structures and their behavioral model parameters. The concept of combining the output from multiple models is growing in the field of climate and environmental modelling. Combining the output from multiple models allows for the characterization of structural uncertainties in the models. Furthermore, multimodel approach may also lead to more skillful simulation/prediction as it accounts for model uncertainty (e.g., Krishnamurti et al., 1999, Georgakakos et al., 2004; Kirtman and Min 2009).

The three RR models used in this study are the Hydrologic MODel (HyMOD; Boyle 2001), the Nedbør-Afstrømnings Model (NAM; Madsen 2000), and the tank model (Sugawara 1995). All three hydrological models are implemented as semi-lumped and are conceptual in the sense that model calibration is essential for the estimation of model parameters. The HyMOD, NAM, and tank have 6, 10, and 16 spatially lumped parameters, respectively.

Hydrological modeling literature, in general, agrees that large combinations of parameters can result in equally acceptable model simulations (Beven 2006). Therefore, we implement a multimodel and multiparameter approach using the generalized likelihood uncertainty estimation method (GLUE; Beven and Binley 1992) to account for uncertainty in hydrological simulation. In GLUE, the ensemble simulation is obtained by weighting the model prediction with model's likelihood measure (Beven and Binley, 1992).

This study builds upon our previous study (Bastola and Misra 2013), which focused on calibration of 28 MOPEX watersheds of the SEUS for the three selected hydrological models. The authors 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 performance of individual model for the entire watershed is measured in terms of the three widely used model performance criteria, namely, the Nash Sutcliffe efficiency index (NSE), Count Effeciency (CE), which is estimated as the percentage of observation included within the Prediction interval (PI) and width of prediction interval. For calibration of each of the three hydrological models, they were used in simulating independently, for each watershed, with 20,000 set of randomly generated model parameters from a uniform distribution. From among these huge set of simulations, only behavioural set of model parameters that result in a value of NSE greater than 0.5 were retained. The model calibration attempt also revealed that the combination of output from the different models improved the reliability and performance of the simulation (not shown). Further details on model calibration and validation can be found in Bastola and Misra (2013).

ii) Simulation of instream water quality

Fernandez et al. (2006) and Shrestha et al. (2008) used a multiple linear regression (log-linear model) to model nutrient loading rate. They reported that a simple log-linear model performs reasonably well in estimating nutrient loadings. A regression model such as load estimator (LOADEST) is traditionally used to predict water quality constituent concentration by linearly relating it with the natural log of streamflow, time, and season (Runkel et al., 2004). Another such regression-based model for modeling nutrient loading rate is the weighted regression on time, discharge and season (WRTDS; Hirsch et al. 2010). More recently, Oh and Sankarasubramanian (2012) looked at the potential application of seasonal forecasts of nutrient loading on a few SEUS watersheds by applying seasonal climate forecasts and the LOADEST. They looked at the variability of the nutrient loadings associated with seasonal climate variability in watersheds that are minimally disturbed. Oh and Sankarasubramanian (2012) conditioned their nutrient prediction on the basis of precipitation as they found a strong correlation between simulated loading and precipitation. They first developed the nutrient loading rating curve by relating nutrient loads with observed streamflow. Then they used empirical orthogonal function and canonical correlation analysis based on simulated loading and the gridded winter precipitation to develop a low-level model to predict loadings for each watershed on the basis of climate forecasts from a climate model. Their study demonstrated useful prediction skills of the winter season total nitrogen behavior in the coastal watersheds of the SEUS.

The streamflow is selected as the predictand as it has been found to be the most important variable in predicting nutrients. Except for the two watersheds in south Florida, all of the watersheds included in this study of nutrient loading forecasts were included in Oh and Sankarasubramanian (2012) study. Among the 28 watersheds used in this study, only seven of these SEUS watersheds have the data on nutrient load. Therefore, water quality forecast is shown only for these seven SEUS watersheds.

Because nutrient measurements are only sparsely available, load estimation using regression-based models has been widely explored for watershed management, especially for watershed planning pollution control (Shrestha et al. 2008). LOADEST, a tool for the estimation of nutrient load based on a number of explanatory variables (e.g., streamflow, decimal time, concentration, etc.) is used in this study. Further details on the calibration and load estimation procedure are found in Runkel et al. (2004). Unlike the watersheds in Oh and Sankarasubramanian (2012), all watersheds included in the nutrient loading forecast experiment are calibrated in this study.

$$\ln(L) = a_0 + a_1 Ln(O) + a_2 \sin(2\pi(T - T')) + a_3 \cos(2\pi(T - T')) \qquad 1$$
$$\ln(O) = \ln(Q) - \ln(Q)' \qquad 2$$

where, L is the load in kg per day; Q is the streamflow in cubic feet per sec; a_0, a_1, a_2, a_3, are the regression coefficients, T is the time measured in years (decimal) and T‘ is the coefficient that defines the center of decimal time; ln(Q)’ is the coefficient that defines the center of the streamflow. The explanatory variable ln(O) is centered to avoid colinearity. Both T‘ and ln(Q)’ are the coefficient estimated from observed data, i.e., load and streamflow data from the U.S. Geological Survey national stream water-quality monitoring networks (http://pubs.usgs.gov/dds/wqn96cd/html/wqn/wq/region03.htm) (WQN). The retrospective seasonal forecast of total nitrogen and phosphorous is computed from (1 & 2). We explored numerous model structure (model option) available in LOADEST. Amongst all available choices, the option 4 (a four-parameter log linear model) of LOADTEST was selected based on its relative performance measured in terms of R^2. The calibrated parameters and performance of model for the selected watersheds is shown in Table 2.

5. Results and discussion

Following Bastola et al. (2013) we will compare the seasonal forecast skill to climatology and one-year lag (persistence) forecast. Furthermore, we will use the Nash Sutcliffe Efficiency (NSE) and Area under the Relative Operating Characteristic Curve (AROC; Marzban 2004) as our deterministic and probabilistic metrics for forecast skill analysis respectively. The NSE derived from the normalized form of root mean square error (2) is used to evaluate the skill of FISH50 with respect to two simple reference forecasts: (a) a climatological forecasts (NSE; Eq 3) and (b) a one-year lag (persistence) forecast (Persistence; Eq 4). The NSE and Persistence are both alternate form of root mean square error. NSE (which range from -∞ to 1) is the normalized root mean square error, which measures the relative magnitude of residual variance to observed variance thereby reflecting how well the simulated value fits observations. An NSE of 1 reflects perfect model and negative value indicates that the skill of the model is worse than using climatology. NSE is one of the most widely used and recommended performance measures for evaluating hydrological predictions and simulations.

The persistence model efficiency is also a normalized statistic as NSE, but the sum of the square of the error is normalized differently (i.e. with respect to one lag forecast). Both NSE and persistence are however biased towards large data values (or high flow period). However these skill metrics also penalize the prediction if they underestimate the high flow period.

$$NSE = 1 - \sum_{i=1}^{n} (Q_{obs,i} - Q_{sim,i})^2 / \sum_{i=1}^{n} (Q_{obs,i} - \overline{Q})^2 \qquad 3$$

$$Persistence = 1 - \sum_{i=1}^{n} (Q_{obs,i} - Q_{sim,i})^2 / \sum_{i=1}^{n} (Q_{obs,i} - Q_{obs,i-1})^2 \qquad 4$$

The Relative (or sometimes referred as Receiver) Operating Characteristic Curve (ROC) describes the relation between the probability of correct (hit rate) and incorrect (false alarm rate) forecasts from a given model. In ROC, models ability to correctly predict the event is plotted versus models ability to exclude a condition correctly. The area under the curve reflects the performance of the predictive models.

Unlike NSE, it takes into account the forecast uncertainty as described by the forecast spread of the individual ensemble members. In the adopted methodology of using resampled historical observations for forcing the multiple hydrological models, we generate a relatively large number of ensemble members (=180; see Section 3) per season, which provides a robust measure of AROC for the streamflow predictions. It may be noted that the value of AROC < 0.5 suggests that the skill is no better than observed climatology.

i) Deterministic skill analysis

In this section, ensemble spreads of the predictions are ignored, and skills are evaluated on the basis of the ensemble mean. The experiment design explained earlier produces an ensemble of predicted flows from the FISH50 ensemble and the multiple calibrated hydrological models (Fig. 1). Hydrological model output simulated with the observed CPC rainfall data is used as a control, or truth, to verify the fidelity of the hydrological predictions.

For the 28 SEUS watersheds included in this study, the spatially averaged daily rainfall from December through May from FISH50 is significantly higher compared to rainfall from CPC (Fig. 2), the reference rainfall dataset. Over most of the watersheds, FISH50 overestimates observed rainfall by nearly 23%. Such a high bias may have a greater impact on the SEUS watersheds characterized by high precipitation elasticity of streamflow (e.g., Sankarasubramanian et al. 2001).

Fig. 3 shows that the bias (i.e., the volume error in simulated flow associated with the FISH50 forcing), which is high for some watersheds, is significantly reduced with the use of resampled observations (FISH50_Resamp). Figure 4 shows climatological streamflow only for the selected six watersheds that broadly span our study region of the SEUS. The raw FISH50 forcing data produces significant bias in the seasonal cycle of the streamflow over the majority of the watersheds shown in Fig. 4. The resampling from historical observations based on analogues of the forecasted quartile rainfall category of the December–May season from FISH50 seems to ameliorate some of this bias in the seasonal cycle relative to the control flow (Fig. 4).

Fig. 5 shows the normalized root mean square errors of the ensemble average streamflow based on two-reference forecasts: the climatological (the NSE) and the lag one-year forecast [or persistent forecast, wherein the observed flow anomalies from the previous year are continued through the following year; persistence efficiency measure (PEM)]. The predicted flow forced with raw FISH50 and FISH50_Resamp shows some skill against both reference forecasts. Skills, however, tend to decrease with lead time.

ii) Probabilistic skill analysis

It is prudent to examine the probabilistic skill of these forecasts given the non-deterministic nature of these seasonal forecasts (Palmer et al. 2000). The probabilistic skill score (as measured by AROC) of FISH50 shows some skill in discriminating different (quartile) categories of rainfall. The FISH50 forecasted monthly mean rainfall shows superior skill than corresponding climatology on 27, 19, 12, and 11 watersheds for very wet, wet, dry, and very dry rainfall categories respectively (Fig. 6). In Fig. 6 each bubble represents an AROC value greater than 0.5 for that watershed, and the size of the bubble indicates a relative value of the AROC that is larger than 0.5. The watersheds in Florida, especially the St. Johns and the Peace River, show skill over the climatological forecast in the wet and very wet categories (Fig. 6). Similarly, on the basis of the average value of AROC across watersheds, the very wet and wet categories are more skillful than the dry and very dry rainfall categories (Fig. 6). In addition, most of the watersheds in Georgia and Alabama also show skill in the very wet quartile (Fig. 6).

The probabilistic skill of the hydrological predictions derived from FISH50 is examined, given the skill of FISH50 in discriminating the different quartile categories of the seasonal rainfall. The AROC for streamflow is calculated as the probabilistic measure to evaluate the experimental hydrological forecast using all 180 realizations per season for each of the 28 watersheds of the SEUS. Fig. 7 shows that on average, very wet and wet categories show more skill (on the basis of the number of watersheds with AROC values $>$ 0.5) compared to the other two quartile categories (i.e., very dry and dry). The AROC values for the forecasted streamflow for each of the 28 watersheds and each month of the season for all the four quartiles are shown in Figs. 8 and 9. The results shown in these figures are consistent with similar analysis of FISH50 precipitation (Fig. 6), which shows that the very wet and wet quartile precipitation categories have higher skill than the other two quartile categories. In other words, the results suggest that for the majority of the SEUS watersheds, the streamflow is sensitive to the quality of the precipitation forecasts. In comparison to the FISH50 summer and fall seasonal hydrological forecasts (Bastola et al., 2013), the winter and spring seasons in Fig. 7 show significantly higher skill.

iii) Skill of seasonal nutrient loading simulation

It is expected that the skill in streamflow will be translated into skill in predicting nutrient loading as the log-linear model (equation 1) is used to predict the nutrient load from the flow and the day of the year, The AROC value for the seven watersheds for which nutrient loading data are available is shown for only two categories, i.e., very wet and very dry quartiles (Figs. 10 and 11). As the skills of the nutrient loading forecasts are identical to the skills in seasonal streamflow prediction, the AROC values for the middle quartiles are not shown. Any differences between the skills in streamflow and nutrient loading can be attributed to the performance of the log linear model during calibration. In Figs. 10 and 11 the skills for both total Nitrogen and total Phosphorous loading for both extreme quartile categories are nearly similar to their corresponding streamflow prediction skills. However, it is apparent from comparing Figs. 10 and 11 that there is more skill in the seven watersheds in forecasting monthly nutrient loadings for the very wet quartile (Fig. 10) than for the very dry quartile (Fig. 11), which follows from similar features observed in streamflows (Figs. 8 and 9). Only seven watersheds which had required water quality data were only considered compared to the 28 watersheds used to evaluate the seasonal predictability of stream flow. This skill in the seasonal forecast of the monthly nutrient loading can be exploited in revising the total maximum daily load that waterways can carry without being impaired. In Florida, the FDEP has imposed a numeric nutrient criteria water quality standard specifically for nitrogen and phosphorous. The skill in seasonal prediction of nutrient loading is likely to promote nutrient trading, especially since nutrient trading has been recently proposed as a major policy to address impairment of waterways and water bodies in Florida.

The simulation of nutrient loading is based on the forecasted streamflow and a relationship between nutrient load and streamflow calibrated from historical observed streamflow and nutrient loading data. The nutrient prediction model only accounts for the influence of the variability in rainfall and streamflow on nutrient dynamics and does not account for the influence of land use explicitly in the dynamics of nutrient.

7. Conclusion

The seasonal climate retrospective forecasts for the boreal winter and spring seasons of FISH50 are evaluated over the SEUS region in simulating streamflow across 28 watersheds and nutrient loading for a small subset (6) of these watersheds. A seasonal hydrological forecast experiment is designed on the basis of an ESP framework, forced with FISH50 meteorological forcing. Three semi-distributed hydrological models are adopted for this study. The experiment setup allows for sampling the hydrological model uncertainty and the meteorological forcing uncertainty. The first uncertainty is handled by using a multimodel approach to predict the streamflow. It is found that over the 28 watersheds, FISH50 overestimated the winter and spring rainfall total by nearly 23%. Therefore, some form of bias correction of rainfall is essential for the application of FISH50 in hydrology. The selected watersheds are characterized by high precipitation elasticity of streamflow, which makes bias correction of forecasted rainfall essential. Bias correction of rainfall from FISH50 is accomplished by resampling the observed seasonal (December–May) historical record with quartile categories similar to those of the FISH50 forecast, which also serve in sampling the uncertainty of the meteorological forcing to the hydrological models.

The experimental setup, therefore, entails 180 ensemble members per season, which includes three hydrological models and 10 samples of observed analogues of meteorological forcing per ensemble member of FISH50 (which has 6 ensemble members per season).

In this study, we examine both the deterministic and the probabilistic skill measures of the meteorological forcing, the predicted streamflow, and the nutrient loading. The former skill measure entails examining the ensemble mean, which ignores the ensemble spread and the forecast uncertainty therein, whereas the latter uses the forecast from all ensemble members.

The seasonal hydrologic forecasts based on ensemble average show superior skill relative to the climatological and lag one-year forecast based on the measures of the NSE. However, these prediction skills show a clear decrease with lead time. In this study, we also use AROC value as a measure of the probabilistic skill of the forecast. The probabilistic skill score of the predicted streamflow is encouraging for the selected watersheds. Especially for the top and middle top quartiles (i.e., the very wet and wet quartile categories), the FISH50 rainfall product for December–May shows comparatively higher skill than the climatology over most of the SEUS watersheds.

For the subset of seven watersheds selected for studying nutrient loading, the log-linear model appears to perform well in modeling the total nitrogen and total phosphorous load. The strong relationship between nutrient loading and streamflow implies that forecast skill in winter streamflow can be potentially exploited in predicting the nutrient load. This advance information on nutrient loading based on seasonal climate forecasts will prove to be essential for maintaining the water quality standards in waterways and water bodies by helping watershed managers plan the total maximum daily load for the season and promote water quality trading.

Acknowledgements

We acknowledge Kathy Fearon of COAPS for help with editing this manuscript. This work was supported by grants from NOAA (NA12OAR4310078, NA10OAR4310215, NA11OAR4310110) and USGS G13AC00408. All computations for integration of the seasonal climate model (FISH50) used in this paper were done on the computational resources provided by the Extreme Science and Engineering Discovery Environment (XSEDE) under TG-ATM120017 and TG-ATM120010.

References

Ambrose, R.B., Wool, T.A., Martin, J.L., Connolly, J.P., Schanz, R.W., 1981: WASP4: A hydrodynamic and water quality model—Model Theory. In: User's Manual and Programmer's Guideunknown:book, US-EPA, Athens,GA.

Arnold, J., A. Williams, R. Srinivasan, B. King, and A. Griggs. 1994: SWAT, soil and water assessment tool. Temple, TX 76502. ARS, USDA.

Bastola, S. and Misra, V. 2013: Evaluation of dynamically downscaled reanalysis precipitation data for hydrological application. Hydrol. process. doi: 10.1002/hyp.9734

Bastola, S., Misra, V., Li, H 2013: Seasonal hydrological forecasts for watersheds over the Southeastern United States for boreal summer and fall seasons. Earth Interactions, 17(25), 1-22, doi:10.1175/2013EI000519.1.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D. and Andreassian, V. (2013) Characterising performance of environmental models. In: Environmental modelling and software, 40 (2013) pp. 1-20.

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.

Beven K. 2006: A manifesto for the equifinality thesis. J. Hydrol. 320:18–36. doi:10.1016/j.jhydrol.2005.07.007.

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.

Bolson, J., C. Martinez, N. Breuer, P. Srivastava, and P. Knox, 2013: Climate information use among Southeast US water managers: An Assessment of Opportunities. Reg Environ Change. In press.

Boyle, D. 2001: Multicriteria calibration of hydrological models. PhD dissertation. Tucson, AZ: Department of Hydrology and Water Resources, University of Arizona; 2001.

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.

Carbone, G. J., and K. Dow, 2005: Water resource management and drought forecasts in South Carolina. J Am Water Resour As. 41: 145–155.

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.

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.

Connelly, B. A., D. T. Braatz, J. B. Halquist, M. M. DeWeese, L. Larson, and J. J. Ingram. 1999: Advanced hydrologic prediction system. J Geophys Res. 104(D16): 19655–19660.

Day, G.N. 1985: Extended Streamflow Forecasting Using NWSRFS. J of Water Res Pl-ASCE, 111(2): 157-170. doi: http://dx.doi.org/10.1175/JCLI3812.1

Dai, Aiguo, 2006: Precipitation Characteristics in Eighteen Coupled Climate Models. J. Climate, 19, 4605–4630. doi: http://dx.doi.org/10.1175/JCLI3884.1

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.

EPA, 1987: The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Document and User’s Manual. US EPA, EPA/600/3-87/007, Athens, GA.

FDEP 2006: Water Quality Credit Trading: A Report to the Governor and Legislature. (http://www.dep.state.fl.us/water/watersheds/docs/WQ_CreditTradingReport_final_December2006.pdf, Feb/2013)

FDEP 2010: The Pilot Water Quality Credit Trading Program for the Lower St. Johns River: A Report to the Governor and Legislature (http://www.dep.state.fl.us/water/wqssp/docs/WaterQualityCreditReport-101410.pdf, Feb/2013).

FDEP 2012: Development of Numeric Nutrient Criteria for Florida Lakes, Spring Vents and Streams (http://www.dep.state.fl.us/water/wqssp/nutrients/docs/tsd-nnc-lakes-springs-streams.pdf, Feb/2013).

Fernandez, GP., Chescheir, GM., Skaggs, RW., Amatya, DM. 2006: DRAINMOD-GIS: a lumped parameter watershed scale drainage and water quality model. Agr Water Manage. 81: 77–97

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.

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

Georgakakos, K. P., Seo, D.-J., Gupta, H., Schaake, J., and Butts, M. B. 2004: Characterizing streamflow simulation uncertainty through multimodel ensembles, J. Hydrol., 298, 222–241

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

Hirsch, R. M., Moyer, D. L. and Archfield, S. A. 2010: Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs. J Am Water Resour As. 46: 857–880. doi: 10.1111/j.1752-1688.2010.00482.x

Johanson, R.C., Imhoff, J.C., Davis, H.H., Kittle, J.L., Donigian, A.S. 1981: User's Manual for Hydrologic Simulation Program-Fortran (HSPF), Release 7.0, US-EPA, Athens, GA.

Kiladis GN, Diaz HF. 1989: Global climate extremes associated with extremes of the Southern oscillation. J Climate. 2:1069–1090

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

Kirtman, B.P. et al. 2014: The North American Multi-Model Ensemble (NMME): Phase-1 Seasonal to Interannual Prediction, Phase-2 Toward Developing Intra-Seasonal Prediction. Bull. Amer. Meteor. Soc., doi: 10.1175/BAMS-D-12-00050.1

Krishnamurti,T.N., Kishtawal,C.M., LaRow,T., Bachiochi,D., Zhang,Z, Williford,E., Gadgil,S. and Surendran,S. 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. Science. 285 No 5433: 1548-1550.

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

Kumar, A., and M. P. Hoerling, M. Ji, A. Leetmaa, and P. D. Sardeshmukh, 1996: Assessing a GCM's suitability for making seasonal predictions. J Climate. 9: 115–129.

Li H and Misra V. 2013: Global Seasonal Climate Predictability in a Two Tiered Forecast System. Part II: Boreal winter and spring seasons. Clim Dynam. Available from http://floridaclimateinstitute.org/images/document_library/publications/fish50-paper-partII-final.pdf

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.

Madsen H. 2000: Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. J Hydrol. 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

Maurer, E.P., D.P. Lettenmaier, and N. J. Mantua. 2004: Variability and predictability of North American runoff. Water resour res. 40(9), W09306 doi:10.1029/2003WR002789.

Misra V, Li H, Wu Z, DiNapoli S. 2013: Global seasonal climate predictability in a two tiered forecast system. Part I: Boreal summer and fall seasons. Clim Dyn. doi:10.1007/s00382-013-1812-y

NRC 2000: Clean Coastal Waters. National Academy Press, Washington, D.C

Obeysekera JA, Trimble P, Cadavid L, Santee R, White C. 1999: Use of Climate Outlook for Water Management in South Florida, USA. South Florida Water Management District: West Palm Beach, FL

Oh, J., and Sankarasubramanian A. 2012: Climate, Streamflow and Water Quality Interactions over the Southeastern US, Hydrol Earth Syst Sc. 17: 2285-2298.

Palmer, T. N, C. Brankovic, and D. S. Richardson, 2000: A probability and decision model analysis of PROVOST seasonal multi-model ensemble integrations. Quart. J. Roy. Meteor. Soc., 126, 2013-2034.

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 As. 45(3): 767–778.

Pagano, T. C., H. C. Hartmann, and S. Sorooshian, 2001: Using climate forecasts for water management: Arizona and the 1997-2001 El Niño. J Am Water Resour As. 37: 1139–1153.

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.

Preston SD, Alexander RB, Schwarz GE, Crawford CG. 2011: Factors Affecting Stream Nutrient Loads: A Synthesis of Regional SPARROW Model Results for the Continental United States. J Am Water Resour As. 47(5): 891-915. doi: 10.1111/j.1752-1688.2011.00577.x.

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. Climatic Change. 93: 177–195.

Ropelewski CF, Halpert MS. 1986: North American precipitation and temperature patterns associated with the El Nino/Southern Oscillation (ENSO). Monthly Weather Review. 114:2352–2362

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

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.

Runkel, R.L., Crawford, C.G., and Cohn, T.A. 2004: Load Estimator (LOADEST): A FORTRAN Program for Estimating Constituent Loads in Streams and Rivers: U.S. Geological Survey Techniques and Methods Book 4, Chapter A5, 69 p.

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

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).

Shrestha, S., Kazama, F., Newham, LTH. 2008: A framework for estimating pollutant export coefficients from long-term in-stream water quality monitoring data. Environ Modell Softw. 23:182–194

Smith, R.A., Schwarz, G.E., Alexander, R.B. 1997: Regional Interpretation of water-quality monitoring data. Water resour res. . 33 (12), 2781–2798.

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

USEPA 2006: Reassessment of Point Source Nutrient Mass Loadings to the Mississippi River Basin, November, 2006, Mississippi River/Gulf of Mexico Watershed Nutrient Task Force. 2006a pp.

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

Wood A. W., Leung LR, Sridhar V, Lettenmaier DP. 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., John C. Schaake, 2008: Correcting Errors in Streamflow Forecast Ensemble Mean and Spread. J. Hydrometeor. 9, 132–148. doi: http://dx.doi.org/10.1175/2007JHM862.1

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

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

List of Figures

Fig. 1. Schematic of hydrological simulation based on multiple ensembles of climate model forecasted meteorological forcing (60 members per seasonal forecast) and multiple hydrological models (3 models), for a total of 180 seasonal hydrological simulations per season per watershed.

Fig. 2. Climatological rainfall averaged over the 28 watersheds in the Southeastern U.S. (identifiers indicated along the x-axis).

Fig. 3. Volume error of the flow predicted with forcing from raw FISH50 data (FISH50), resampled from historical observational analogues of Dec-May mean rainfall from FISH50 (FISH50_Resamp).

Fig. 4. Predicted monthly mean flow with raw FISH50 (FISH50), resampled from historical observational analogues of Dec-May mean rainfall from FISH50 (FISH50_Resamp) forcing.

Fig. 5. 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 lag one-year as a reference forecast and NSE is the Nash Sutcliffe efficiency measured with respect to climatological value as a reference forecast.

Fig. 6. Area under ROC (AROC) for FISH50 Dec–May mean precipitation averaged over the respective watersheds in the Southeastern U.S. (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 from 0.5 to 1.0.

Fig. 7. Summary of probabilistic assessment of flow predicted with FISH50_Resamp for four selected quartile categories of December-May mean rainfall over the various watersheds in the Southeastern U.S (Very Dry, Very Wet, Medium Wet, and Medium Dry). An AROC greater than 0.5 suggests that the prediction skill is better than the climatology.

Fig. 8. AROC for the 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 from 0.5 to 1.0. Values of AROC below 0.5 are not plotted.

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

Fig. 10. AROC value for simulated (a) total nitrogen (upper quartile) and (b) total phosphorous (upper quartile). Size of the blue circle represents the skill (AROC) of streamflow and size of the red circle represents the skill of nutrient loading.

Fig. 11. AROC value for simulated (a) total nitrogen (lower quartile) and (b) total phosphorous (lower quartile). Size of the blue circle represents the skill (AROC) of streamflow and size of the red circle represents the skill of nutrient loading.

List of Tables

Table 1 General characteristics of the selected watershed

SN	Basin (USGS ID)	Lon	Lat	Area (Sq mile)	Annual Rain (mm)	Annual Ave runoff (Cumecs)	River system
1	2456500	-87.0	33.7	885	1425	45.6	LOCUST FORK AT SAYRE, AL.
2	3574500	-86.3	34.6	320	1467	17.3	PAINT ROCK RIVER NEAR WOODVILLE AL
3	2414500	-85.6	33.1	1675	1425	86.3	TALLAPOOSA RIVER AT WADLEY AL
4	2296750	-81.9	27.2	1367	1248	44.8	PEACE RIVER AT ARCADIA, FLA.
5	2329000	-84.4	30.6	1140	1349	49.3	OCHLOCKONEE RIVER NR HAVANA, FLA.
6	2365500	-85.8	30.8	3499	1425	171.9	CHOCTAWHATCHEE RIVER AT CARYVILLE, FLA.
7	2375500	-87.2	31.0	3817	1493	201.2	ESCAMBIA RIVER NEAR CENTURY, FL
8	2236000	-81.4	29.0	3066	1293	97.7	ST. JOHNS RIVER NR DELAND, FLA.
9	2192000	-82.8	34.0	1430	1333	65.7	BROAD RIVER NEAR BELL, GA.
10	2202500	-81.4	32.2	2650	1189	85.4	OGEECHEE RIVER NEAR EDEN, GA.
11	2217500	-83.4	33.9	392	1385	19.9	MIDDLE OCONEE RIVER NEAR ATHENS, GA.
12	2347500	-84.2	32.7	1850	1317	78	FLINT RIVER NEAR CULLODEN, GA.
13	2383500	-84.8	34.6	831	1528	48	COOSAWATTEE RIVER NEAR PINE CHAPEL, GA.
14	2339500	-85.2	32.9	3550	1475	189.3	CHATTAHOOCHEE RIVER AT WEST POINT, GA.
15	2387000	-84.9	34.7	687	1433	37.2	CONASAUGA RIVER AT TILTON, GA.
16	2387500	-84.9	34.6	1602	1480	87.6	OOSTANAULA RIVER AT RESACA, GA.
17	2102000	-79.1	35.6	1434	1171	51	DEEP RIVER AT MONCURE, N.C.
18	2118000	-80.7	35.8	306	1257	13.9	SOUTH YADKIN RIVER NEAR MOCKSVILLE N C
19	2126000	-80.2	35.1	1372	1173	48.9	ROCKY RIVER NEAR NORWOOD, N. C.
20	2138500	-81.9	35.8	67	1436	4	LINVILLE RIVER NEAR NEBO N C
21	3443000	-82.6	35.3	296	2156	33.5	FRENCH BROAD RIVER AT BLANTYRE N C
22	3451500	-82.6	35.6	945	1544	70.7	FRENCH BROAD RIVER AT ASHEVILLE, N. C.
23	3504000	-83.6	35.1	52	1895	4.6	NANTAHALA RIVER NEAR RAINBOW SPRINGS, NC
24	3512000	-83.4	35.5	184	1720	14	OCONALUFTEE RIVER AT BIRDTOWN, N. C.
25	3550000	-84.0	35.1	104	1846	8.5	VALLEY RIVER AT TOMOTLA, N. C.
26	2156500	-81.4	34.6	2790	1319	139.1	BROAD RIVER NEAR CARLISLE, S. C.
27	2165000	-82.2	34.4	236	1340	10.6	REEDY RIVER NEAR WARE SHOALS,S.C.
28	3455000	-83.2	36.0	1858	1340	114.5	FRENCH BROAD RIVER NEAR NEWPORT, TN

Table 2. Parameters and corresponding performance of nutrient loading rating curve developed using LOADEST.

Sno	Nutrient	Station	No of data points	T'	Ln(Q)'	a	a1	a2	a3	R^2	Mean Load (Ton/day)	SE
1	Total nitrogen	2296750	143	1984.77	6.30	7.76	1.04	0.17	0.13	0.90	4.89	0.30
2		2329000	133	1983.44	6.32	7.34	0.85	0.00	0.00	0.92	5.96	0.37
3		2365000	118	1983.52	8.82	9.07	0.93	-0.09	0.12	0.83	11.31	0.51
4		2375500	144	1983.14	8.61	8.87	1.04	0.12	-0.09	0.87	10.25	0.47
5		2236000	61	1985.65	7.41	8.55	1.11	-0.06	-0.06	0.90	10.69	0.58
6		2202500	141	1983.95	7.35	7.72	1.07	-0.08	-0.32	0.92	4.77	0.16
7		2126000	64	1984.76	7.25	8.92	1.08	-0.21	0.08	0.96	19.32	1.68
8	Total phosphorous	2296750	143	1984.77	6.30	7.60	0.75	-0.01	-0.02	0.72	3.09	0.13
9		2329000	133	1983.44	6.32	5.55	0.73	-0.03	0.03	0.82	0.81	0.07
10		2365000	118	1983.52	8.82	6.31	1.21	-0.02	0.11	0.82	0.77	0.06
11		2375500	144	1983.14	8.61	6.17	1.21	0.05	-0.10	0.85	0.80	0.09
12		2236000	61	1985.65	7.41	5.91	1.06	-0.15	0.22	0.80	0.84	0.11
13		2202500	141	1983.95	7.35	5.28	0.94	-0.13	-0.29	0.86	0.35	0.01
14		2126000	64	1984.76	7.25	6.95	1.03	-0.43	0.39	0.88	2.20	0.32

T' and Ln(Q)' are the centering parameter for time and log of flow

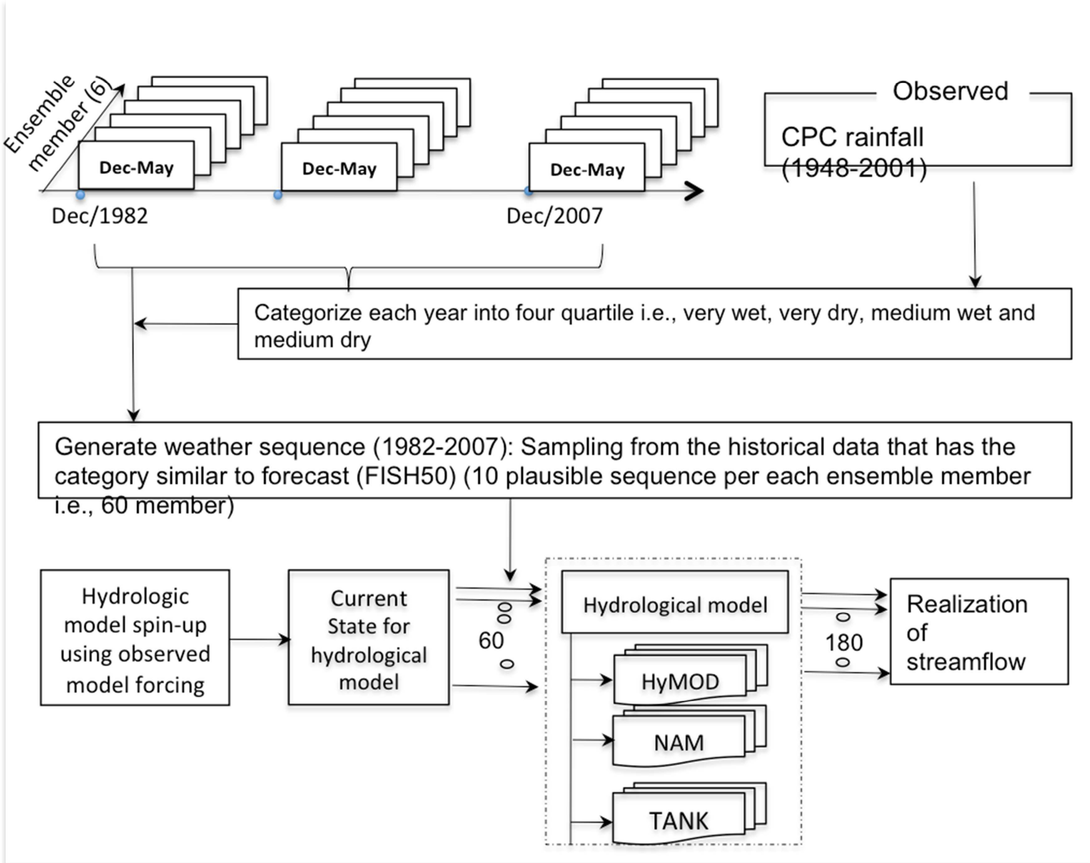

Fig. 1. Schematic of hydrological simulation based on multiple ensembles of climate model forecasted meteorological forcing (60 members per seasonal forecast) and multiple hydrological models (3 models), for a total of 180 seasonal hydrological simulations per season per watershed.

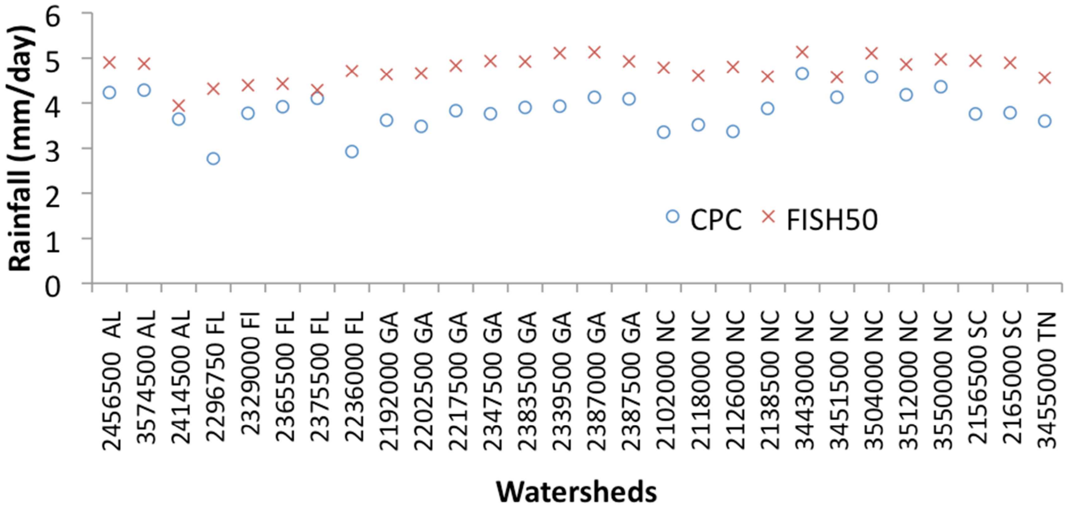

Fig. 2. Climatological rainfall averaged over the 28 watersheds in the Southeastern U.S. (identifiers indicated along the *x*-axis).

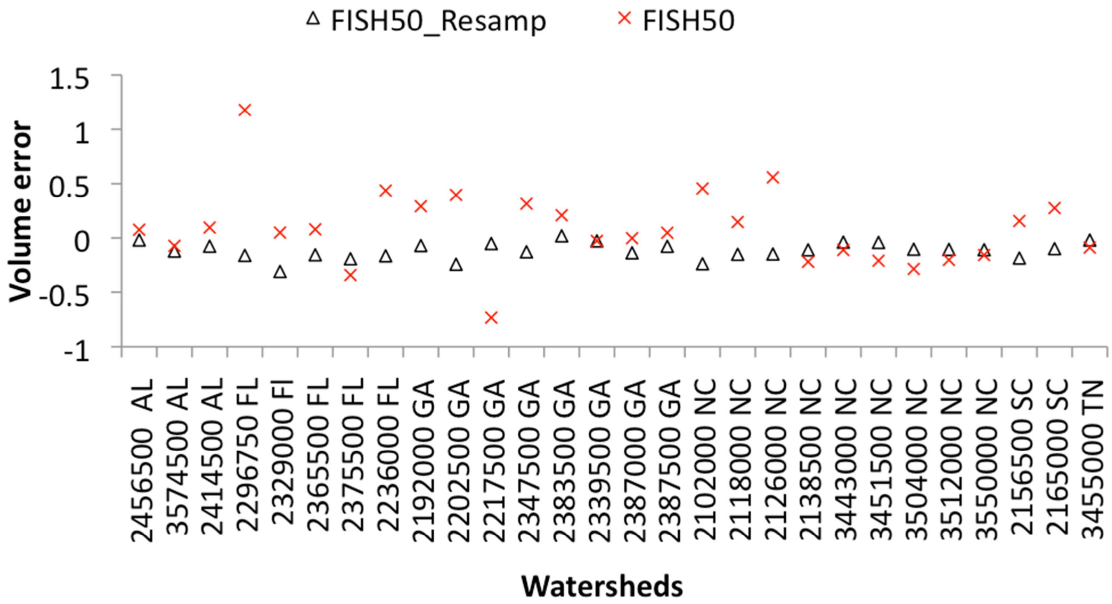

Fig. 3. Volume error of the flow predicted with forcing from raw FISH50 data (FISH50), resampled from historical observational analogues of Dec-May mean rainfall from FISH50 (FISH50_Resamp).

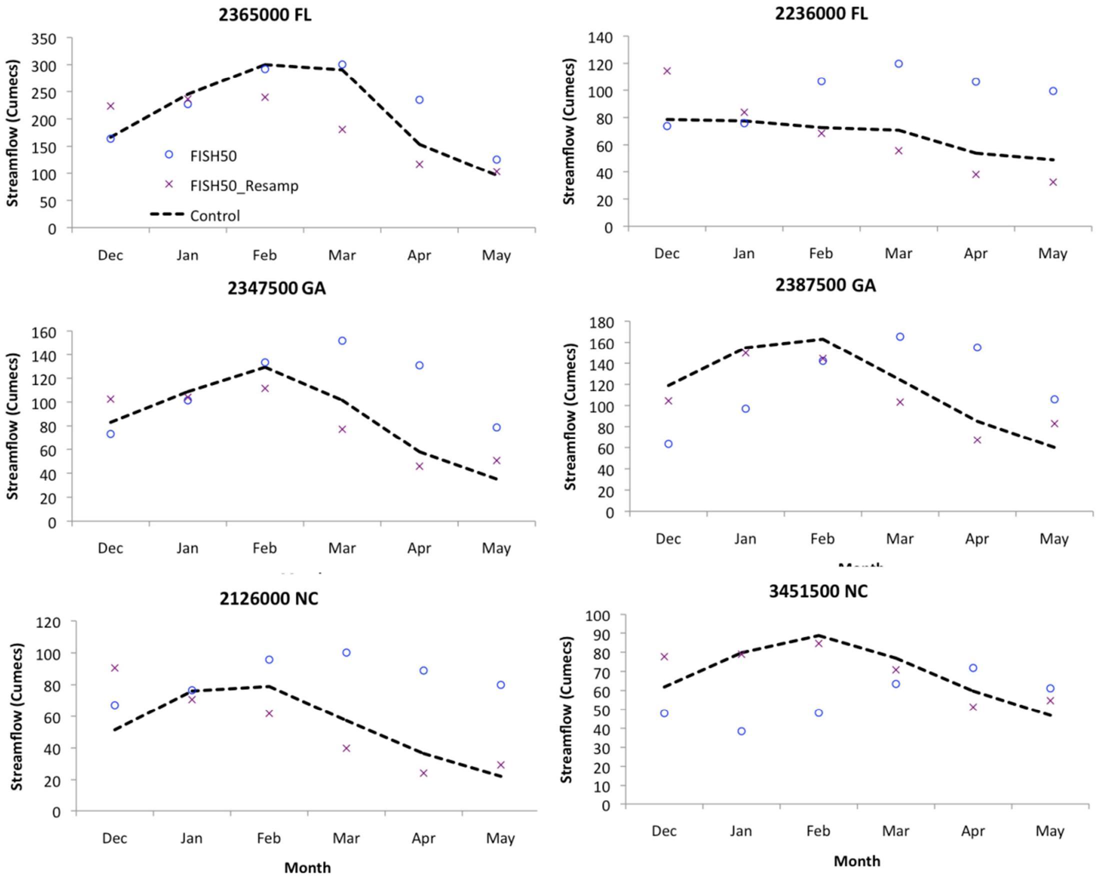

Fig. 4. Predicted monthly mean flow with raw FISH50 (FISH50), resampled from historical observational analogues of Dec-May mean rainfall from FISH50 (FISH50_Resamp) forcing.

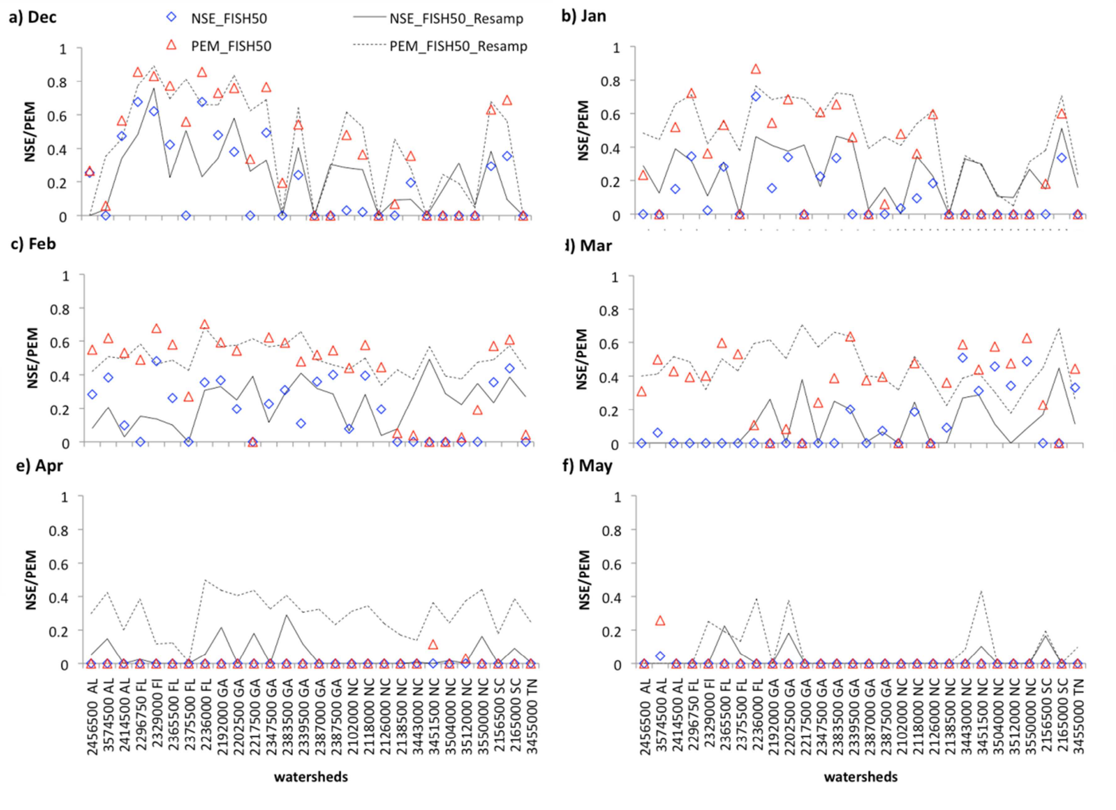

Fig. 5. Skill scores of the hydrological prediction based on normalized root mean square errors. PEM (persistence efficiency measure) is the Nash Sutcliffe efficiency criteria measured with respect to lag one-year as a reference forecast and NSE is the Nash Sutcliffe efficiency measured with respect to climatological value as a reference forecast.

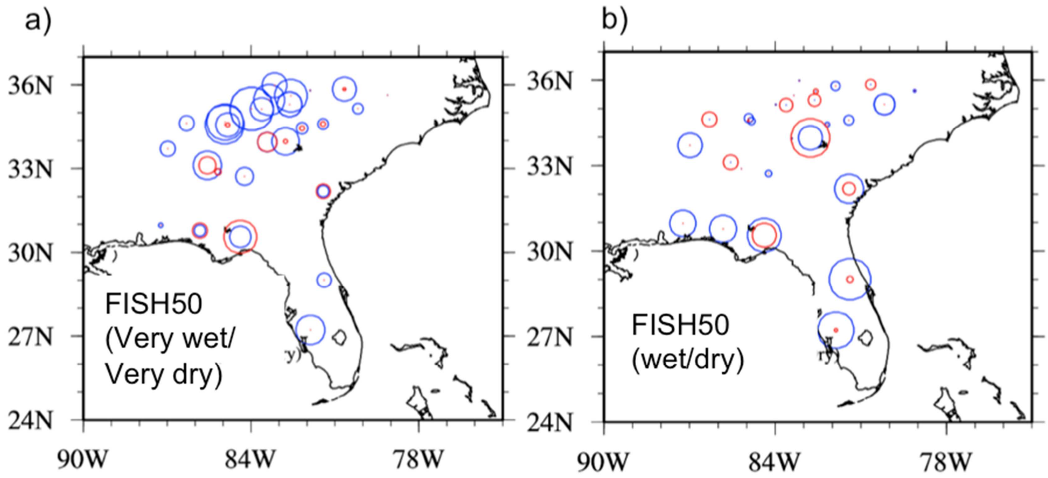

Fig. 6. Area under ROC (AROC) for FISH50 Dec–May mean precipitation averaged over the respective watersheds in the Southeastern U.S. (a) AROC value for very wet (blue circle) and very dry (red circle) rainfall categories, (b) AROC value for medium wet (blue) and medium dry (red) categories. Only AROC values over 0.5 are shown (for watershed with no skill, a dot is used to represent the location of the watershed). The size of the bubble indicates the relative magnitude of the AROC, which can range from 0.5 to 1.0.

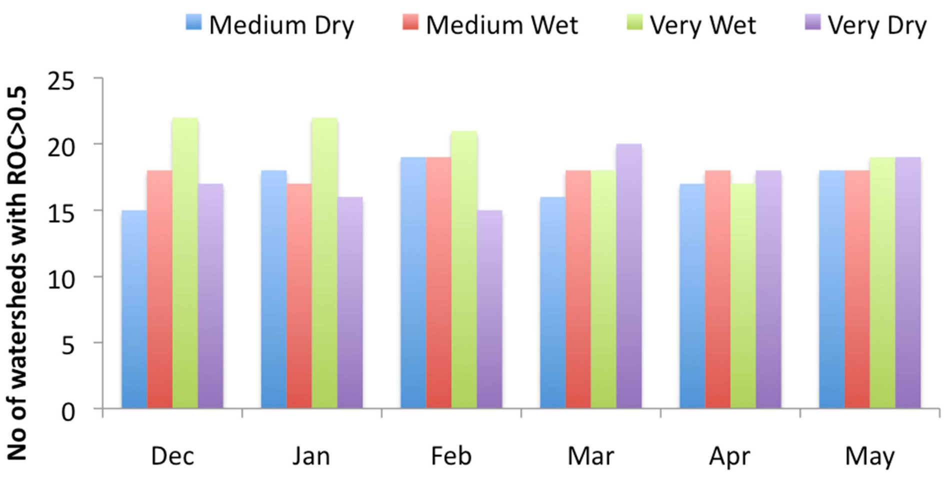

Fig. 7. Summary of probabilistic assessment of flow predicted with FISH50_Resamp for four selected quartile categories of December-May mean rainfall over the various watersheds in the Southeastern U.S (Very Dry, Very Wet, Medium Wet, and Medium Dry). An AROC greater than 0.5 suggests that the prediction skill is better than the climatology.

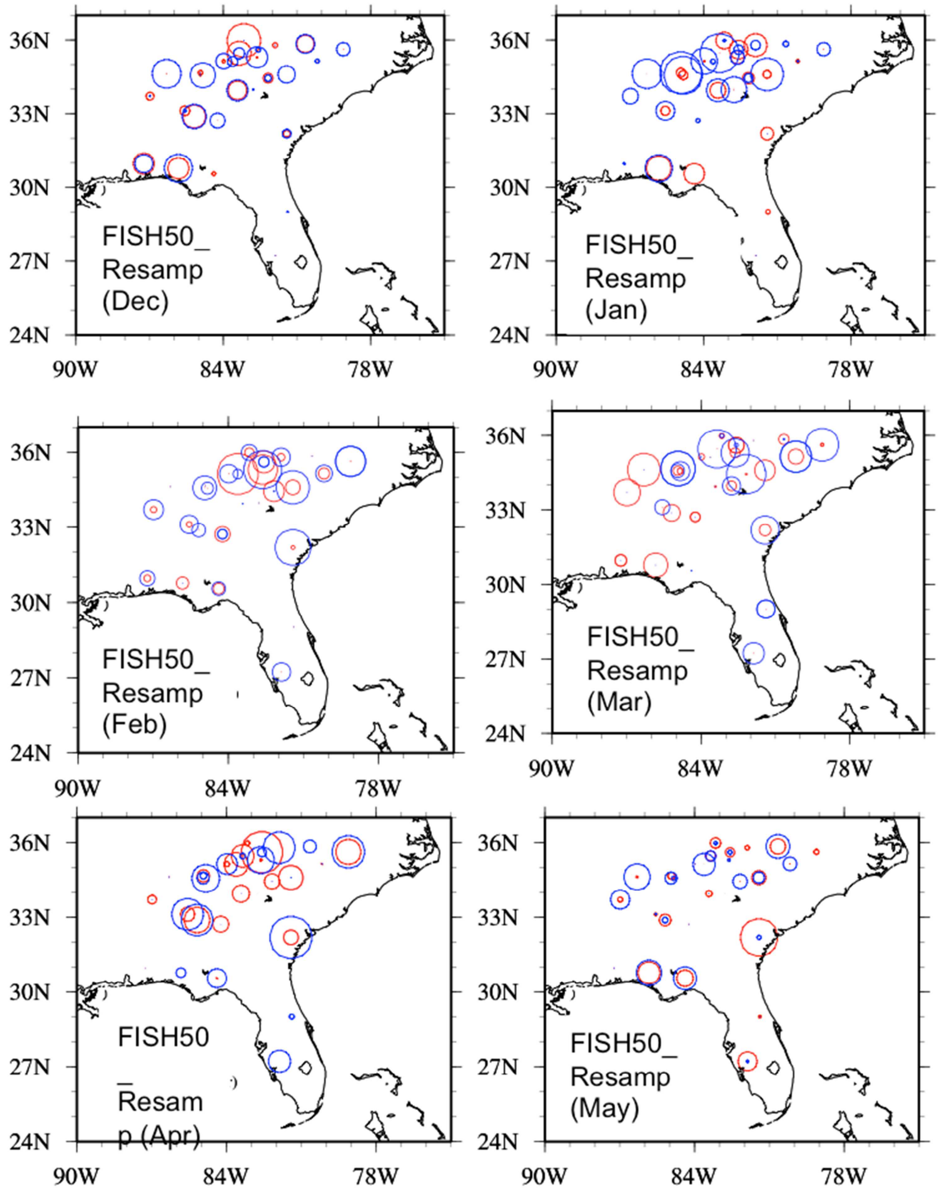

Fig. 8. AROC for the very wet (blue) and very dry (red) categories of streamflow predicted with FISH50_Resamp forcing. The size of the bubble indicates the relative magnitude of the AROC, which can range from 0.5 to 1.0. Values of AROC below 0.5 are not plotted.

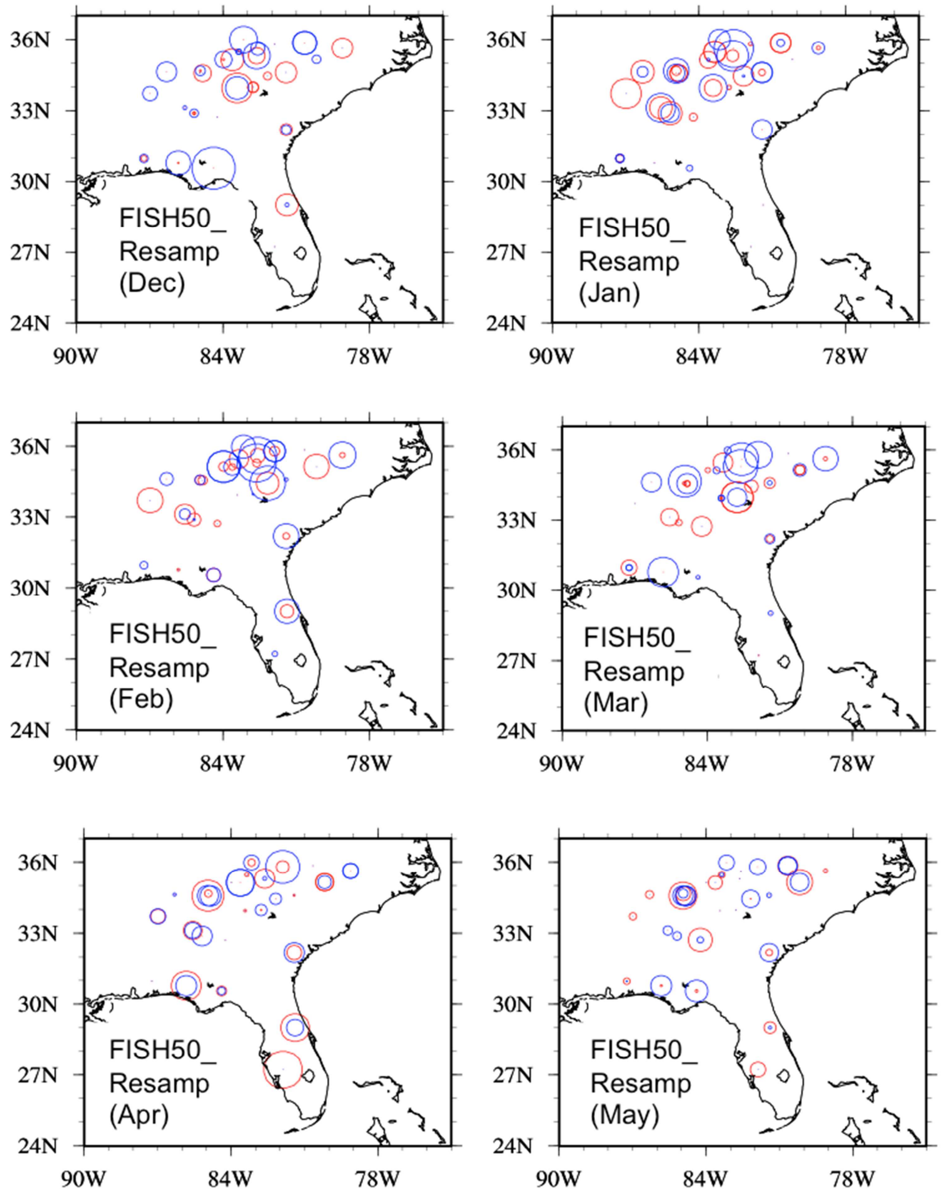

Fig. 9. Same as Fig. 9 but for AROC values for the medium wet (blue) and medium dry (red) categories of streamflow.

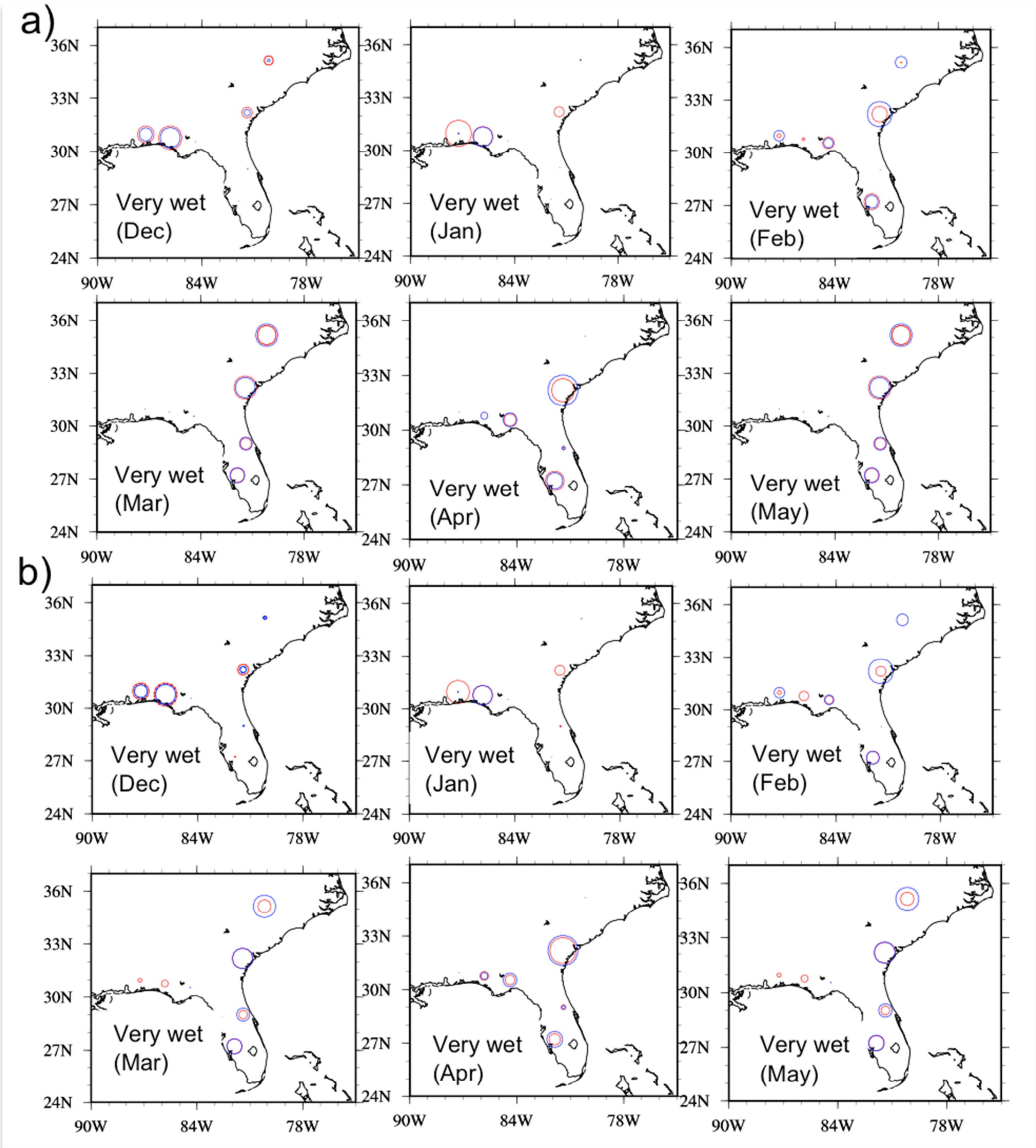

Fig. 10. AROC value for simulated (a) total nitrogen (upper quartile) and (b) total phosphorous (upper quartile). Size of the blue circle represents the skill (AROC) of streamflow and size of the red circle represents the skill of nutrient loading.

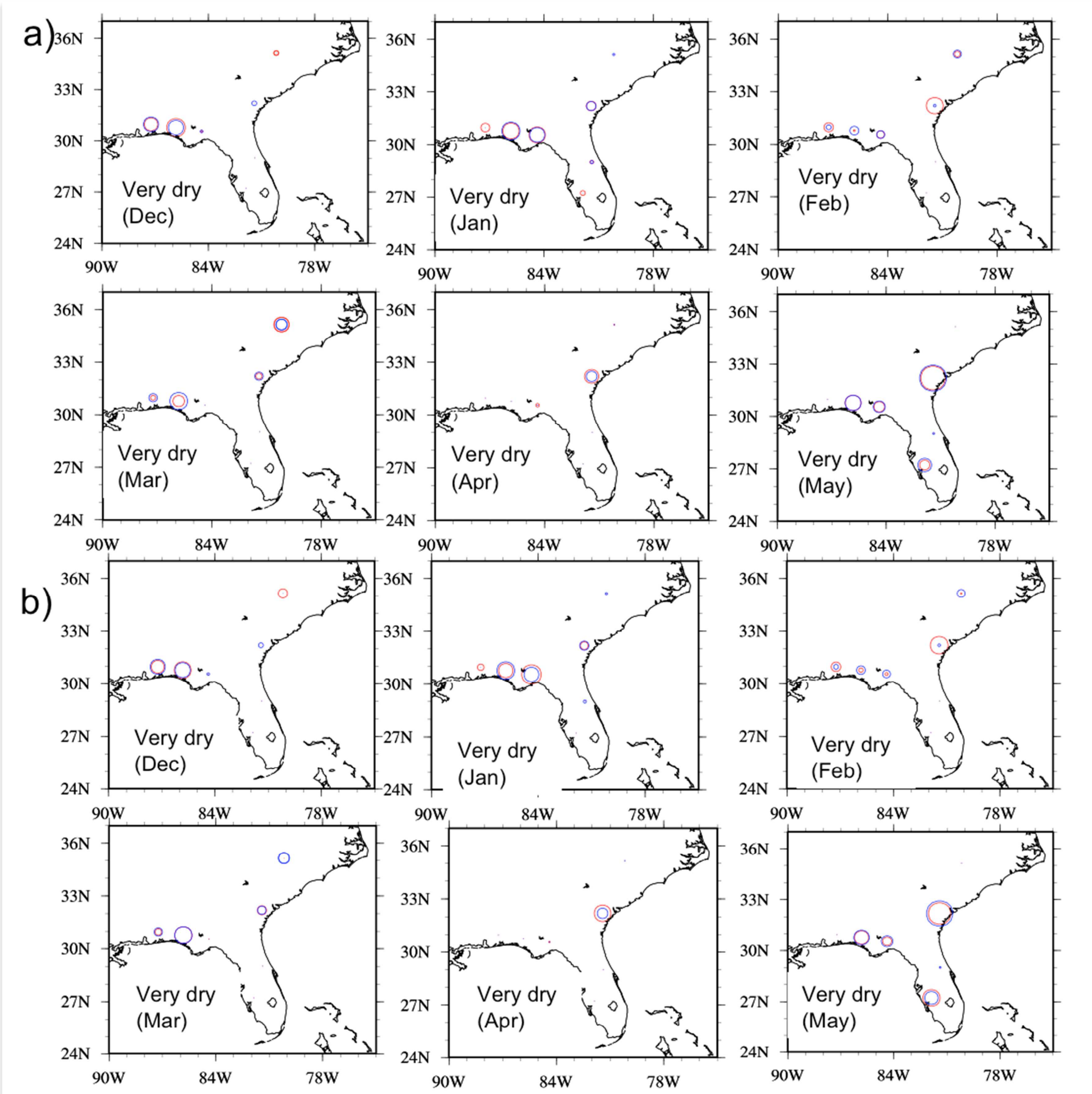

Figure 11: AROC value for simulated a) total nitrogen (lower quartile) and b) total phosphorous (lower quartile). Size of the blue circle represents the skill (AROC) of streamflow and red represents the skill of nutrient loading.