ABSTRACT

MAZROOEI, AMIRHOSSEIN. Decomposition of Sources of Errors in Seasonal Streamflow Forecasting Over the US Sunbelt Using Multiple Land Surface Models. (Under the direction of Dr. Sankarasubramanian Arumugam and Dr. Tushar Sinha).

In order to better manage water demands from multiple uses (e.g., municipal water demands, hydroelectric power generation, and agricultural operations), water resources managers and operators are interested to know changes in seasonal streamflow potential based on climate forecasts. However, various sources of uncertainty in streamflow forecasts pose significant challenges to utilize streamflow forecasts in real time operations. Recent studies have shown that streamflow forecasts generated by using multiple hydrologic models can perform better than those developed by using a single model. In this study we systematically evaluate and quantify various sources of errors in developing seasonal streamflow forecasts from multiple Land Surface Models (LSMs) forced with climate forecasts from ECHAM4.5 General Circulation Model (GCM). The objectives of this study are: 1) Develop retrospective seasonal streamflow forecasts in the US Sunbelt using multiple LSMs, 2) Quantify various sources of errors arising from each LSM, climate forecasts, and downscaling/disaggregation techniques employed in developing streamflow forecasts, and 3) Compare the performance and the skill of different LSMs in streamflow forecasting in selected target basins over the study domain.

First, three-month ahead precipitation forecasts from ECHAM4.5 GCM for each season were statistically downscaled to 1/8° scale using Principal Component Regression (PCR) and then temporally disaggregated to daily time step using K-Nearest-Neighbor (KNN) approach. Hourly climatology for other climatic forcing variables were obtained over almost 30 years (from 1979 to 2010) from NLDAS-2 simulations. Then multiple LSMs such as NOAH3.2, Catchment and CLM2 from NASA's LIS framework were forced with precipitation forecasts and climatological forcings to develop retrospective seasonal streamflow forecast over the period 1991-2010. Finally, the performance of different LSMs in forecasting streamflow under different schemes was analyzed to validate forecasted streamflows and decompose sources of errors.

Decomposition of Sources of Errors in Seasonal Streamflow Forecasting Over the US Sunbelt Using Multiple Land Surface Models

by
Amirhossein Mazrooei

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Civil Engineering

Raleigh, North Carolina

2014

APPROVED BY:

Dr. Sankarasubramanian Arumugam Co-Chair of the Advisory Committee	Dr. Tushar Sinha Co-Chair of the Advisory Committee

Dr. Ryan Boyles

DEDICATION

To my parents, Shahnaz and Habib, and my inspiring sister Parisa, due to their endless love and support. I certainly would not be where I am today without your nurture, guidance, and care. I love you.

BIOGRAPHY

Amirhossein Mazrooei was born in Esfahan, Iran. He received his high school diploma in Mathematics and Physics from Ejei high school, Iranian National Organization for Developing Exceptional Talents (NODET). Amirhossein earned his B.S. degree in Civil Engineering at Sharif University of Technology (Tehran, Iran). Then, he decided to join North Carolina State University to pursue his Master's of Science in Water Resources Engineering under the direction of Dr. Sankar Arumugam. Improving streamflow and soil moisture forecasting skill was the main part of his research during his MS.

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to my advisors, Dr. Sankar Arumugam and Dr. Tushar Sinha for their continuous support and guidance and invaluable kindness during my research. It was a privilege to have learned from you. I would like to appreciate my advisory committee member, Dr. Ryan Boyles for his time and consideration. I would also like to thank my friends Dominic, Seung, and JP, whose contribution helped me a lot to complete my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

Water resources managers and operators are interested to know about changes in seasonal streamflow potential based on climate information to ensure supply for multiple uses include municipal water demands, hydroelectric power generation, and agricultural operations. In this regard, considerable progress has been made in reducing the uncertainty in seasonal streamflow forecasts through enhanced understanding of climatic teleconnections (e.g., El Niño Southern Oscillation (ENSO)) as well as developing Land Surface Models (LSMs) that capture land-atmosphere interactions (R. A. Betts et al. 1997; Hamlet et al. 2002; R. D. Koster & Suarez 1995; Mahanama et al. 2012; E. Maurer et al. 2002). Nevertheless, streamflow forecasting skill varies over different geographic regions during different seasons and also due to error propagation resulting from different steps/methodologies employed in streamflow forecasting. For instance, errors in large scale climate forecasts and their spatio-temporal downscaling limit the application of developed streamflow forecasts in real time water management operations (Sinha et al. 2014; Hartmann et al. 2002).

One of the major challenges in developing streamflow forecasts is the spatio-temporal mismatch between the scales of climate forecasts and implementation of land surface models (Luo & E. F. Wood 2008; Yuan et al. 2011; A. W. Wood et al. 2002; A. W. Wood & D. P. Lettenmaier 2006). Climate forecasts from General Circulation Models (GCM) are typically

available at a monthly time step and at large spatial scale (e.g., 2.8°×2.8°) while land surface models (LSMs) require forcings at a finer spatial scale (e.g., 1/8° to 1/4°) and a shorter time-scale (e.g., daily or sub-daily) for developing forecasts. Over the last decade, several strides have been made to bridge the spatial scale mismatch by employing statistical downscaling (A. W. Wood et al. 2004; E. Maurer & Hidalgo 2008) or dynamical downscaling to obtain air temperature and precipitation at the LSMs scale (Leung et al. 2004; Wood et al. 2004; Wilby et al. 2000; Wood et al. 2002). Studies show that there is a limited difference in skill by adopting dynamical or statistical downscaling (Murphy 1999; Wilby et al. 2000) and further research is needed to understand the value added in climate forecasts by using Rigional Climate Models (RGMs) (Di Luca et al. 2012). In this study, a statistical spatial downscaling technique was employed to obtain monthly precipitation forcings at 1/8° spatial scale from the ECHAM4.5 GCM monthly precipitation forecasts available at 2.8°, followed by a temporal disaggregation scheme to estimate daily time series based on monthly forecasts. Nevertheless, both spatial downscaling and temporal disaggregation processes introduce errors (uncertainties) in precipitation forcings, which are propagated into the developed streamflow forecasts (Hayhoe et al. 2004; Prairie et al. 2007). Sinha et al. (2014) provided a systematic approach to decompose various sources of errors arising in monthly streamflow forecasts for a rainfall-runoff dominated basin (Apalachicola River at Chattahoochee, FL) in the Southeast. One of the goals of our study is to quantify various sources of errors in seasonal streamflow forecasts at regional scale, arises from climate forecasts, spatial downscaling, and temporal disaggregation.

In addition to errors due to climate forcings, another source of uncertainty in developed streamflow forecasts is the LSM itself. Several studies have shown that multi-model combination perform better in reducing model uncertainties and errors in forecasted streamflows (Li & Sankarasubramanian 2012; Devineni et al. 2008). Furthermore, implementing LSMs with updated initial hydrologic conditions (IHCs) can reduce uncertainties in forecasted streamflow (Sankarasubramanian et al. 2008; Sinha & Sankarasubramanian 2013). Most studies have focused on quantifying the role of climatological forcings and initial conditions in snow dominated basins (Maurer et al. 2004; Shukla & Lettenmaier 2011; Mahanama et al. 2012). For instance, in snow-melt dominated basins, the role of initial conditions is more dominant over forecasting skill, whereas in rainfall-runoff regimes, climate forcings are more critical (Shukla & Lettenmaier 2011; Maurer et al. 2004; Maurer & Lettenmaier 2003). Fewer studies have exclusively compared the role of climate forecasts and initial conditions and found that seasonal streamflow forecasting skill during summer primarily depends on climate forecasts (Li et al. 2009; Luo et al. 2007).

The main goal of our study is to quantify different sources of errors in developing seasonal streamflow forecasts, such as spatial downscaling, temporal disaggregation, imprecise GCM climate forecasts, and imprecise climatological forcings during different seasons over the US Sunbelt. For this purpose, we considered NASA's Land Information System (LIS) which provides a flexible framework for running multiple LSMs. In this study, we employed three LSMs (NOAH3.2, Catchment, and CLM2) to develop seasonal streamflow forecasts by forcing them with three-month ahead ECHAM4.5 climate forecasts

issued at the beginning of January, April, July, and October. For quantifying different sources of errors in developed forecasts, we prescribe a unique decomposition procedure that compares the streamflow forecasts developed at 1/4°×1/4° resolution with the reference simulated streamflow developed by using observed forcings. The study also compares the performance of three LSMs in forecasting streamflow over four river basins across the US Sunbelt.

This Thesis is organized as follow: Chapter 2 provides details on hydro-climatic data and the LSM modeling framework used in this study. Chapter 3 explains the details of experimental design for quantifying various sources of errors in developing seasonal streamflow forecasts and the associated metrics. Chapter 4 presents the results based on the experimental design. And finally, in Chapter 5, the salient findings from the study are summarized along with directions for future research work.

CHAPTER 2
HYDRO-CLIMATIC DATA AND MODELING FRAMEWORK

2.1. Study Area

In this study, our analysis is focused on the Sunbelt of the United States, which we define it as a region south of the 37°N latitude in the US (Figure 1). The Sunbelt includes 14 states under different hydro-climatic regimes. It experiences a range of desert climates in the Southwest to humid subtropical climates in the Southeast. In this study, we perform detailed analysis on the seasonal streamflow forecasting skill with respect to the US Geological Survey (USGS) observed streamflow by comparing four watersheds across the Sunbelt (Figure 1):

a) **Verde River Basin:** It is located in desert climate with hot summers and mild winters. The Verde River is about 225 Km long and the basin drains about 17,160 Km^2 area in the North-central Arizona. Two major dams, Horseshoe and Bartlett dams are located within the basin.

b) **Guadalupe River Basin:** It flows through Texas and drains about 3,367 Km^2 area. It is located in a region which is classified as semiarid and sub-humid. It has several dams along its reach, among which the Canyon dam is the largest dam.

c) **The Apalachicola-Chattahoochee-Flint (ACF):** It is the largest basin among the selected basins, which has drainage are of 44,500 Km^2. It drains through Eastern Alabama, Northwestern Florida, and Western Georgia.

d) **The Tar River Basin:** It is located in North Carolina with humid subtropical climate and drains around 5,700 Km^2 area.

These basins are selected since they represent different hydro-climatic conditions over the Sunbelt, and have at least one USGS stream gauging station that is listed under the Hydro-climatic Data Network (HCDC) (more details follow in the next section).

2.2. Observed Streamflow Data

We obtained the observed monthly streamflow data from 1981 to 2010 for the selected watersheds from the US Geological Survey (USGS) sites (Table 1) that are included in the Hydro-Climatic Data Network (HCDN) (Slack et al. 1993). These sites are minimally impacted by reservoir operations. The observed monthly streamflow from 1981 to 2010 was obtained from the USGS sites shown in Figure 1.

Table 1. Selected USGS gage stations

Station Name	State	Station Number
Apalachicola River at Chattahoochee	Georgia	02358000
Guadalupe River near Spring Branch	Texas	08167500
Verde River near Camp Verde	Arizona	09506000
Tar River at Tarboro	North Carolina	02083500

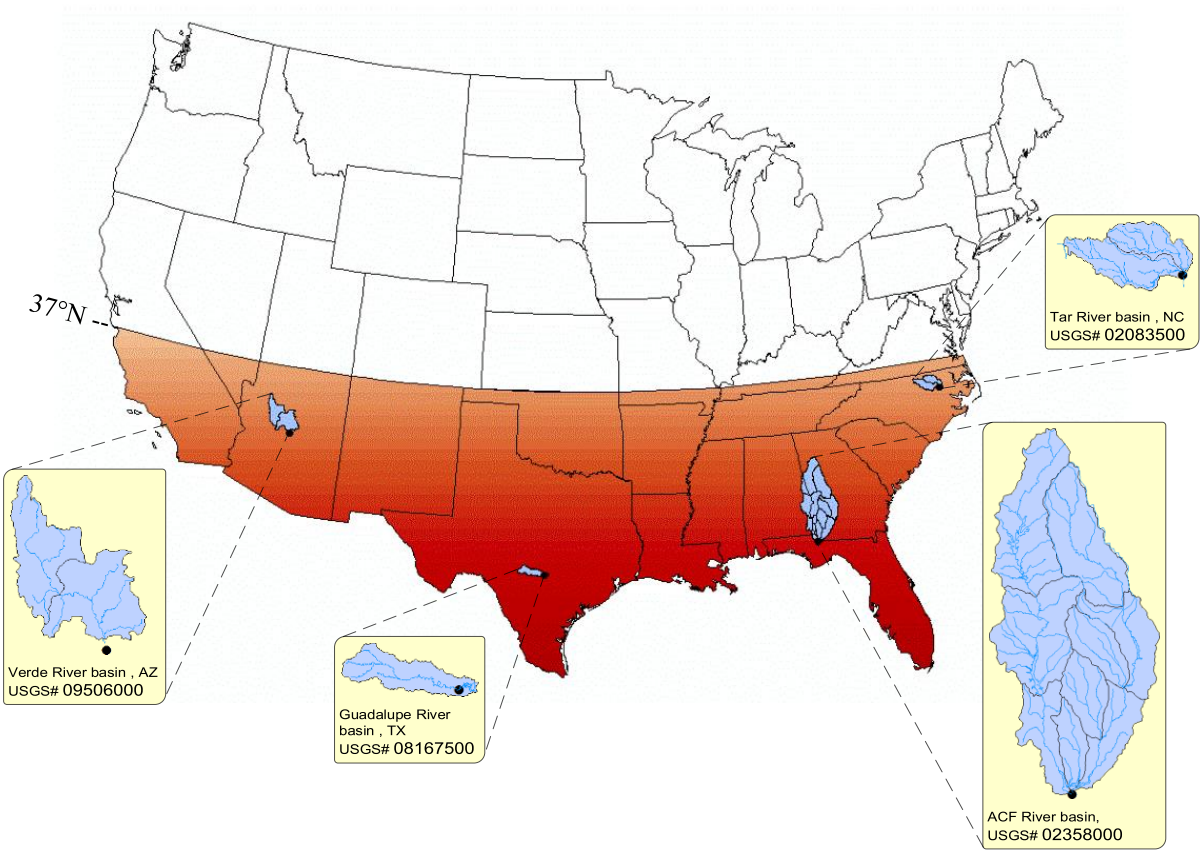

Figure 1. Location of the selected river basins across the Sunbelt of the US (shaded area) and the considered USGS gauging stations for evaluating the forecast skill.

2.3. Forcings Data

2.3.1. (NLDAS-2) Dataset

North American Land Data Assimilation System version 2 (NLDAS-2) comprises of meteorological forcings data for phase 2 of the NLDAS project (K. E. Mitchell et al. 2004). The data is available at 1/8° spatial resolution and hourly temporal resolution from 01 Jan 1979 to present. These data are available at Goddard Earth Science Data and Information Services Center (GES DISC) (http://disc.sci.gsfc.nasa.gov/hydrology/data-holdings). Except precipitation, other land surface forcings in the NLDAS-2 dataset are derived from the analysis fields of the National Centers of Environmental Predictions (NCEP) North American Regional Reanalysis (NARR) which are estimated based on multiple land surface models (e.g., SAC, Mosaic, Noah, and VIC). The NLDAS-2 dataset (Table 2) served as the primary forcings to update the LSMs as well as for obtaining the simulated flows from them.

Table 2. Meteorological forcing variables in the NLDAS-2 dataset used in this study.

Full Name	**Unit**
180-0 mb above ground Convective Available Potential Energy	J/kg
Fraction of total precipitation that is convective	unitless
Long wave radiation flux downwards (surface)	W/m^2
Short wave radiation flux downwards (surface)	W/m^2
Potential evaporation	kg/m^2
Surface pressure	Pa
2-m above ground Specific humidity	kg/kg
2-m above ground Temperature	K
10-m above ground Zonal wind speed	m/s
10-m above ground Meridional wind speed	m/s

2.3.2. Hourly Climatological Data

Under streamflow forecasting scheme, LSMs were fed with climatological forcings as well. In this regard, we computed the mean hourly forcings of all the variables at 1/8°×1/8° spatial scale based on the NLDAS-2 hourly forcings over the 31 years period (from 1979 to 2010).

2.3.3. Observed Precipitation Data

Observed daily precipitation data during the years 1957 to 2010 were obtained from Maurer et al. (2002), which are available at 1/8° spatial scale for the entire US. This observed precipitation along with the NLDAS-2 forcings were used to force LSMs to develop updated initial hydrological conditions as well as for obtaining the simulated flow over the grid point for each LSM.

Maurer et al (2002) used interpolation between observations from weather gauges across the country to develop gridded observed precipitation dataset. However, the errors due to interpolation can be significant under specific circumstances particularly in mountainous regions. Further, it is important to note that the gauge densities across our study region are not uniform. For example urban regions have higher gauge density than rural and hilly terrains which can result in different error characteristics in the interpolated precipitation. Also, complex topographies experienced by the weather gauges can result in significant errors depending on the form of the precipitation (e.g. in the Western US). In addition, seasonality characteristics of precipitation can also increase the interpolation errors

particularly during the summer season when the localized convection results in increased spatial variability in rainfall.

2.3.4. ECHAM4.5 Precipitation Forecasts

We obtained monthly updated precipitation forecasts from the ECHAM4.5 GCM from the International Research Institute of Climate and Society (IRI) Climate Data Library. These updated forecasts are available from January 1957 to present at spatial scale of 2.8° for up to 7-month lead time. Constructed analogue Sea Surface Temperatures (SSTs) forecasts have been used to develop the ECHAM4.5 climate forecast data and it consists of 24 ensemble members. In this study, we used the ensemble average of monthly precipitation forecasts issued at the beginning of each season. For instance, we obtained precipitation forecasts of January, February, and March (JFM) issued in January for developing JFM streamflow forecasts. Thus, we utilized seasonal climate forecasts issued at the beginning of January, April, July, and October during the period from 1991 to 2010 in order to develop seasonal streamflow forecasts for winter (JFM), spring (AMJ), summer (JAS), and fall (OND) respectively.

2.4. NASA's Land Information System (LIS)

The LIS is a flexible and powerful framework used for high performance land surface modeling and data assimilation (Kumar et al. 2006). The LIS has been developed at NASA GSFC with the goal of integrating satellite and ground-based observational data products and advanced land surface modeling techniques to produce optimal fields of land surface states

and fluxes. In this study, we considered three LSMs included in the LIS framework: 1) NOAH3.2, 2) CATCHMENT, and 3) CLM2 to quantify different sources of errors in seasonal streamflow forecasting over the Sunbelt. A brief overview of these models is provided below:

2.4.1. NOAH3.2 Land Surface Model

NOAH LSM has been developed from 1993 through multi-institutional cooperation (Chen et al. 1996; Niu et al. 2011). NOAH model has been widely used in operational weather and climate predictions. NOAH is a 1-D column hydrological model which can be executed in either coupled or uncoupled mode. NOAH3.2 is the latest version, which requires near-surface atmospheric forcings as inputs and computes surface energy and water balance variables. The output of the NOAH LSM is soil moisture, soil temperature, snowpack depth, snowpack water equivalent, canopy water content, and the energy flux and water flux terms (MB Ek et al. 2003).

2.4.2. CATCHMENT Land Surface Model

The Catchment LSM (CLSM) is a semi-distributed LSM that describes the coupled water and energy budgets based on near-surface meteorology (Koster et al. 2000). The CLSM uses the concept of the TOPMODEL (Beven 1997) to account for lateral water fluxes along with topography and their influence on small scale variability of soil moisture, and runoff. The important feature of the catchment model is the representation of various topography conditions in spatial scale. Thus it is easier to parameterize downslope

streamflows and spatial variability in surface fluxes and soil moisture (Stieglitz et al. 1997; Famiglietti & Wood 1994; Famiglietti & Wood 1991).

2.4.3. CLM2 Land Surface Model

The Community Land Model (CLM) is a 1-D columnar (snow-soil-vegetation) model of the land surface (Bonan et al. 2002). The CLM2 is a newer version which is coupled to an atmospheric model or a climate system model (CSM) (Bonan et al. 2002; Hoffman et al. 2005). Therefore, it represents exchange of mass, energy, and momentum over the land surface with the atmosphere. CLM2 quantifies physical states and land-surface fluxes over 10 layers for soil and up to 5 layers for snow. The unique feature of CLM2 is its utilization of more developed biophysics, carbon cycle, and vegetation dynamics.

Thus, these three LSMs were forced with monthly precipitation forecasts from ECHAM4.5 GCM and the hourly climatological forcings obtained from NLDAS-2 dataset, to develop seasonal streamflow forecasts at 1/4° spatial scale over the entire study domain.

CHAPTER 3
EXPERIMENTAL DESIGN AND ANALYSES METRICS

In this chapter, first a description of the employed methodologies related to downscaling and disaggregation of GCM forecasts is provided . Then, I describe the experimental design considered to analyze different sources of errors in developing seasonal streamflow forecasts followed by forecasts-evaluation metrics.

3.1. Spatial Downscaling

As mentioned before, monthly ECHAM4.5 precipitation forecasts available at 2.8° spatial scale, and daily observed precipitation data at 1/8° spatial scale were obtained over the period from 1961 to 2010. In order to spatially downscale the ECHAM4.5 GCM monthly precipitation forecasts from 2.8° to 1/8° spatial scale, we employed Principal Component Regression (PCR). Given a specific month, for each 1/8° grid, four nearest 2.8° grids were selected as predictors while the observed monthly precipitation at 1/8° grid was assigned as predictand. Since the predictors are correlated, we developed a PCR model to estimate the downscaled monthly precipitation by considering previous 25 years of data as training period prior to the forecasting year. For example, to obtain spatially downscaled value for a given 1/8° grid for January 1991, previous 25 years of January observations (i.e., from 1966 to

1990) were served as training period. Thus, we consider the downscaled monthly precipitation forecasts for temporal disaggregation approach, which is described next.

3.2. Temporal Disaggregation

In order to derive daily time series of precipitation for the given monthly downscaled precipitation forecasts from ECHAM4.5 GCM, we used temporal disaggregation method explained in Prairie et al. (2007). This method uses Kernel nearest neighbor approach (K-NN) to identify "K" numbers of months that have similar historical monthly conditions to the downscaled monthly forecasts based on leave-one out cross-validation. Based on the identified "K" neighbors of monthly precipitation, the disaggregation approach basically resamples the daily precipitation using Lall and Sharma kernel (Lall & Sharma 1996), which gives the highest (lowest) weights to the historical monthly precipitation that is the closest (farther) to the forecasted monthly precipitation. Thus we develop disaggregated daily precipitation forecasts for each forecasting season (from 1991 to 2010), based on the three-month ahead precipitation forecasts issued at the beginning of the season over the evaluation period (1991-2010).

It is important to note that the disaggregated precipitation at each 1/8° grid point is obtained independently to the precipitation at its neighbor grid points. This may result in reduced spatial correlation in the ensembles of disaggregated precipitation. This could also potentially add error to the forecasted precipitation particularly during the winter season when there is a high spatial correlation across the observed daily precipitation. This issue

could be potentially improved by adding spatial probability functions in the disaggregation procedure in order to preserve the spatial correlation of daily precipitation time series between neighboring grid points. But, this improved spatial correlation in disaggregated time series may not have a significant effect in the streamflow forecasting skill since we analyze the skill scores on a seasonal time step and also add up 1/4° forecasted streamflows to a basin scale.

3.3. Precipitation Schemes

Various combinations of different precipitation forcings as well as other meteorological forcings (identified in Table 2, either from NLDAS-2 dataset or its climatology) were used in this study to implement the three LSMs (NOAH3.2, Catchment, and CLM2). Figure 2 illustrates an overview of different schemes considered for developing the precipitation forcings. Consequently, using different schemes of precipitation forcings provide different streamflow forecasts which decompose different sources of errors in developing streamflow forecasts.

3.3.1. Observed Precipitation:

Daily 1/8° observed precipitation forcings from Maurer et al. (2002) are used directly in developing streamflow under observed forcings scheme. This provides the simulated streamflow at each grid point.

3.3.2. Temporally Disaggregated Observed Precipitation:

Observed daily precipitation is first aggregated to monthly time series and then disaggregated to daily time series based on the disaggregation method described in the section 3.2. Forcing this disaggregated time series from observed precipitation on LSMs provide the error due to temporal disaggregation.

3.3.3. Spatially Downscaled and Temporally Disaggregated Observed Precipitation:

Precipitation forcings under this scheme were obtained in three steps:

i) The 1/8° observed daily precipitation was aggregated to monthly precipitation and then upscaled to 2.8° scale to be consistent with the resolution of the ECHAM4.5 precipitation forecasts.

ii) Then the monthly observed precipitation at 2.8° was spatially downscaled to monthly 1/8° resolution using Principal Component Regression (PCR) approach described in "Spatial Downscaling" section (3.1).

iii) Finally, temporal disaggregation of the precipitation is performed using the K-NN scheme described in section 3.2.

This way, we get the simulated flows from LSMs forced with observed monthly precipitation which incorporate downscaling errors (due to PCR) and disaggregation errors (due to K-NN).

3.3.4. ECHAM4.5 Precipitation Forecasts:

Using spatial downscaling and temporal disaggregation approaches described earlier, daily precipitation forecasts at 1/8° spatial scale were obtained from monthly ECHAM4.5 precipitation forecasts available at 2.8° spatial scale.

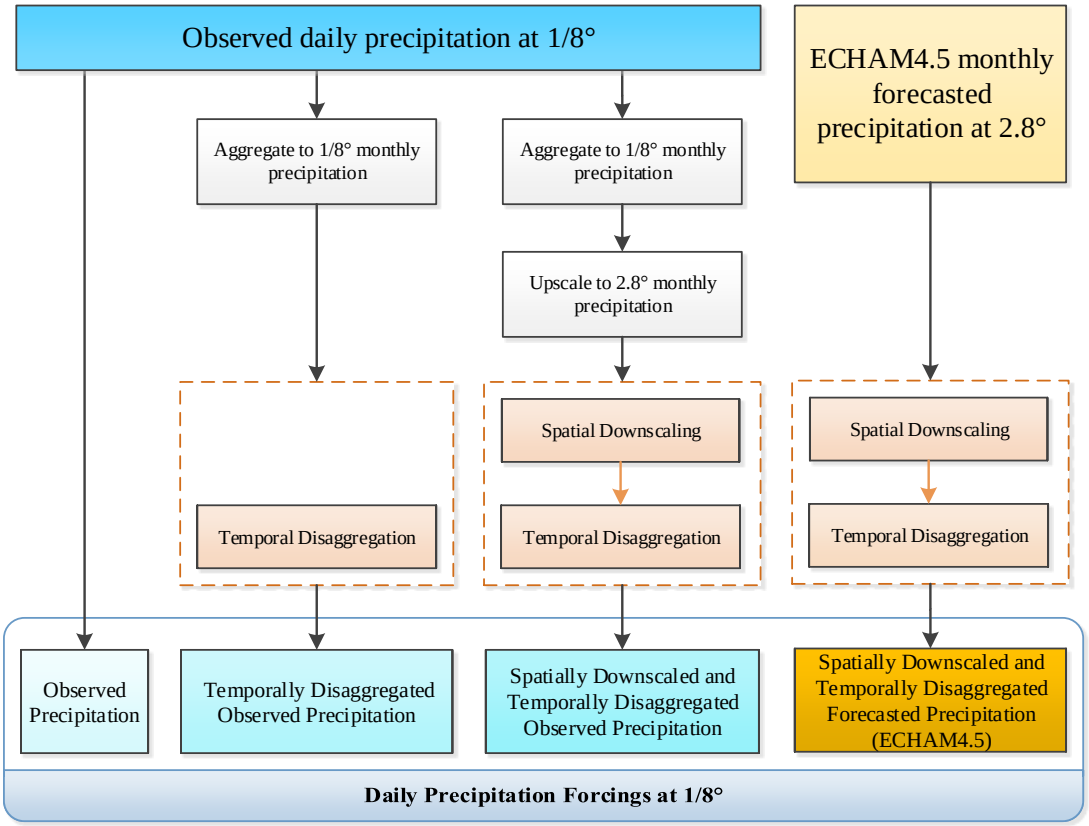

Figure 2. Schematic diagram of different precipitation forcing schemes for quantifying the sources of errors in developing seasonal streamflow forecasts.

3.4. Seasonal Streamflow Simulation/Forecast Schemes

All the streamflow estimates/forecasts were obtained by feeding LSMs with different precipitation forcings obtained in section 3.3 as well as other forcings (identified in Table 2, either from NLDAS-2 or its climatology) along with updated initial conditions. Figure 3 illustrates the schematic of the overall experimental design considered in quantifying the sources of errors arising in developing seasonal streamflow forecasts.

3.4.1. Streamflow Simulation Using Observed Forcings (Q^{ref})

To begin with, we obtained streamflow simulation from 1988 to 2010 using all observed forcings. Under this scheme, the NLDAS-2 hourly precipitation is superimposed by the observed precipitation to simulate streamflow by LSMs (Q^{ref} in Figure 3). We discarded the first three years of the simulation, which served as the spin up period to reach steady-state for each LSM. Since we used the observed forcings to implement each of the three LSMs included in this study, we considered this simulated streamflow as the "reference streamflow". In addition, these simulations were served as the basis to update the initial conditions at the beginning of seasons for each LSM to develop seasonal streamflow forecasts.

3.4.2. Seasonal Streamflow Forecast Using ECHAM4.5 (Q^{fcst})

Under the forecasting mode, we used monthly updated spatially downscaled and temporally disaggregated ECHAM4.5 precipitation forecasts as the precipitation forcing and the NLDAS-2 hourly climatological values for forcings other than precipitation. Using these forcings and initial conditions updated at the beginning of each season, the seasonal

streamflow forecasts were obtained for each 1/4° grid point over the period 1991-2010 for each LSM (Q^{fcst} in Figure 3).

3.4.3. Quantifying Various Sources of Errors in Streamflow Forecasts

In order to quantify various sources of errors in forecasting streamflow arising from spatial downscaling and temporal disaggregation schemes, using precipitation forecasts, and also using climatological forcings, all of these streamflow estimates were developed by updating the initial conditions in each LSM at the beginning of each season. Error decomposition metrics used to quantify sources of errors are described in details in section 3.5. We first define the following streamflow forecast schemes:

i. "Temporally disaggregated" scheme (Q^{dis}): LSMs are fed with disaggregated observed precipitation along with NLDAS-2 forcings other than precipitation to obtain streamflows having error due to temporal disaggregation.

ii. "Spatially downscaled and temporally disaggregated" scheme (Q^{down_dis}): By feeding LSMs with downscaled and disaggregated observed precipitation and NLDAS-2 forcings excluding precipitation, we obtained seasonal streamflow estimates which includes errors due to both downscaling and disaggregation.

iii. "ECHAM4.5 forecasted precipitation" scheme ($Q^{ECHAM4.5}$): Using downscaled and disaggregated ECHAM4.5 forecasted precipitation along with NLDAS-2 forcings, seasonal streamflow forecasts were made.

iv. "Climatological forcings" scheme (Q^{clim}): In this scheme, hourly climatological forcings which were estimated over a 32 years period (from 1979 to 2010) for all the NLDAS-2 variables are used to implement each of the three LSMs. This indicates the ability of forecast purely based on the updated initial conditions and the expected climate.

v. "NLDAS-2 climatological forcings excluding precipitation climatology" scheme ($Q^{clim\_excl\text{-}prcp}$): Land surface models are fed with NLDAS-2 hourly climatological forcings but instead of climatological precipitation forcings we used observed daily precipitation forcings. Thus, this scheme quantifies the error in streamflow simulations arises from the climatological forcings of the NLDAS-2 variables except precipitation.

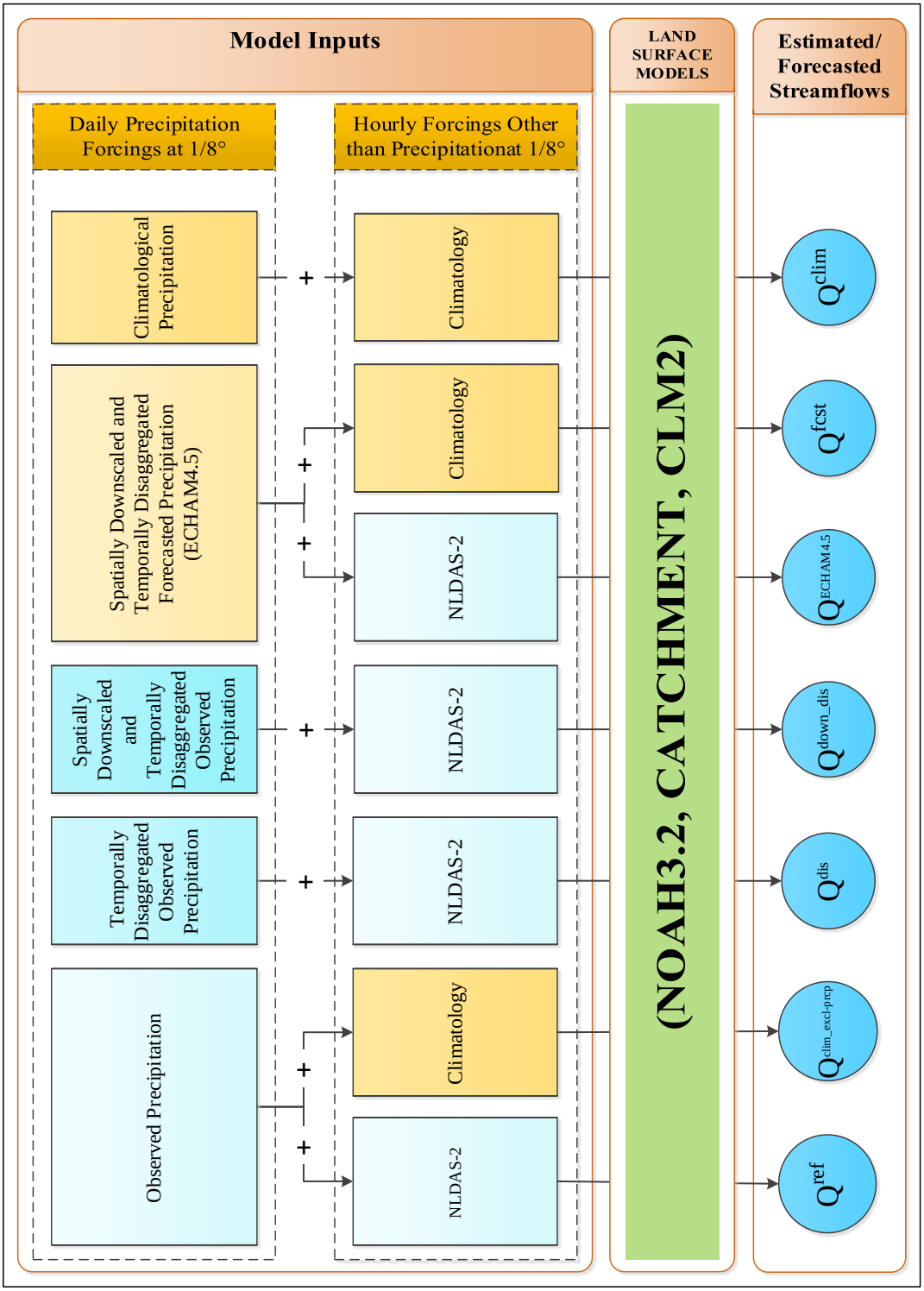

Figure 3. Experimental Design for quantifying different sources of errors in developing seasonal streamflow forecasts.

3.5. Error Decomposition Metrics

Streamflow estimates/forecasts were obtained at daily time scale over 1/4° spatial scale covering the US Sunbelt. From these daily streamflow values, we computed seasonal mean streamflow for 20 years (from 1991 to 2010) over each grid (1/4°×1/4° spatial resolution). Therefore, the errors are computed in form of Root Mean Square Error (RMSE) with respect to the reference seasonal streamflow (Q^{ref}) in the following equations where *t* refers to a given year, *m* denotes LSM, *s* denotes each season. , and *n* refers to the number of years used to compute RMSE which is equal to 20 in this study. The error decomposition metrics for each source of errors are described below:

3.5.1. Errors due to Temporal Disaggregation

In order to quantify errors due to temporal disaggregation technique alone, we computed Root Mean Square Error (RMSE) between reference streamflow (Q^{ref}) and simulated streamflow in "Temporally disaggregated" scheme (Q^{dis}). Thus, based on the assumption in equation (3.1), errors due to disaggregation (MSE_m^{dis}) can be quantified as equation (3.2).

$$Q_{t,m,s}^{dis} = Q_{t,m,s}^{ref} + \varepsilon_{t,m,s}^{dis} \tag{3.1}$$

$$RMSE_{m,s}^{dis} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Q_{t,m,s}^{ref} - Q_{t,m,s}^{dis})^2} \tag{3.2}$$

3.5.2. Errors due to Spatial Downscaling

Based on equation (3.3) we know that the developed streamflow under "Spatially downscaled and temporally disaggregated" scheme (Q^{down_dis}) has both errors due to disaggregation and downscaling techniques. Thus, in equation (3.4), errors associated with spatial downscaling technique alone are detected by comparing Q^{down_dis} with streamflow values simulated in "Temporally disaggregated" scheme (Q^{dis}).

$$Q_{t,m,s}^{down\_dis} = Q_{t,m,s}^{ref} + \varepsilon_{t,m,s}^{down} + \varepsilon_{t,m,s}^{dis} \tag{3.3}$$

$$RMSE_{m,s}^{down} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Q_{t,m,s}^{down\_dis} - Q_{t,m,s}^{dis})^2} \tag{3.4}$$

3.5.3. Errors due to ECHAM4.5 Precipitation Forecasts

Given that the ECHAM4.5 precipitation forecasts are obtained over 2.8° spatial scales and at monthly time scales, both downscaling and disaggregation processes were used to implementing LSMs. Therefore, developed streamflows under this scheme also include errors due to disaggregation and downscaling as well as errors due to large scale ECHAM4.5 precipitation forecasts alone (equation (3.5)). Thus, by removing disaggregation and downscaling errors, we can estimate errors due to ECHAM4.5 precipitation forecasts alone (equation (3.6)).

$$Q_{t,m,s}^{ECHAM\,4.5} = Q_{t,m,s}^{ref} + \varepsilon_{t,m,s}^{ECHAM\,4.5} + \varepsilon_{t,m,s}^{dis} + \varepsilon_{t,m,s}^{down} \tag{3.5}$$

$$RMSE_{m,s}^{ECHAM\,4.5} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Q_{t,m,s}^{ECHAM\,4.5} - Q_{t,m,s}^{down\_dis})^2} \tag{3.6}$$

3.5.4. Errors due to Climatological forcings

In the "Climatological forcings scheme" we used all forcings from NLDAS-2 hourly climatology to implement LSMs (Q^{clim}). Therefore, by comparing results with reference streamflow (Q^{ref}) we can estimate mean square error due to climatological forcings(equation (3.7)). In this scheme, there are no errors due to downscaling or disaggregation methods since we used hourly climatological data at 1/8° spatial resolution to force the models.

$$RMSE_{m,s}^{clim} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Q_{t,m,s}^{ref} - Q_{t,m.s}^{clim})^2} \tag{3.7}$$

On the other hand, we can assume that errors due to climatological forcings consist of two sources of errors. As shown in equation (3.8), errors due to climatological forcings consist of errors due to climatological precipitation forcings ($\varepsilon_{t,m,s}^{clim\_prcp}$) and errors due to climatological forcings excluding precipitation ($\varepsilon_{t,m,s}^{clim\_excl-prcp}$).

$$Q_{t,m,s}^{clim} = Q_{t,m,s}^{ref} + \varepsilon_{t,m,s}^{clim\_prcp} + \varepsilon_{t,m,s}^{clim\_excl-prcp} \tag{3.8}$$

Based on this, we can use streamflow values simulated under "NLDAS-2 climatological forcings excluding precipitation climatology" scheme ($Q^{clim\_excl-prcp}$) to estimate each of the sources of error due to climatology.

$$Q_{t,m,s}^{clim\_excl-prcp} = Q_{t,m,s}^{ref} + \varepsilon_{t,m,s}^{clim\_excl-prcp} \tag{3.9}$$

$$RMSE_{m,s}^{clim\_excl-prcp} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Q_{t,m,s}^{ref} - Q_{t,m,s}^{clim\_excl-prcp})^2} \tag{3.10}$$

$$RMSE_{m,s}^{clim\_prcp} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Q_{t,m,s}^{clim} - Q_{t,m,s}^{clim\_excl\text{-}prcp})^2} \tag{3.11}$$

$RMSE_{m,s}^{clim\_excl\text{-}prcp}$ indicates the error due to using climatological forcing variables excluding precipitation and $RMSE_{m,s}^{clim\_prcp}$ indicates the error due to using climatological precipitation forcings in streamflow simulations.

3.5.5. Errors due to Land Surface Models

All LSMs, based on their structure, result in simulated streamflows different from observed streamflow values, even when they are fed with observed forcings. In this study, since we want to quantify portion of uncertainties at each grid point, we consider simulated streamflow under observed forcings scheme as reference streamflow which is the basis of comparison between streamflows developed under other schemes.

$$Q_{t,m,s}^{Obs} = Q_{t,m,s}^{ref} + \varepsilon_{t,m,s} \tag{3.12}$$

Since observed streamflow data is not obtainable at 1/4°×1/4° grid resolution over the US Sunbelt, this analyses focused on quantifying the model errors by comparing the streamflow simulations/forecasts with the observed flows recorded in each basin. In the next chapter, we present the results and discuss the sources of errors in developing seasonal streamflow forecasts over the sunbelt.

CHAPTER 4
RESULTS AND ANALYSES

4.1. Seasonal Streamflow Forecasts - Skill Assessment

4.1.1. Skill of ECHAM4.5 Precipitation Forecasts

In order to evaluate the skill of the precipitation forecasts, we computed Spearman's rank correlation between the forecasted precipitation and the observed precipitation for each season over the 1991-2010 time period. The precipitation forecasts capture the observed variability in precipitation over the Southeast, Northeast, and the Southern coastal plains during winter and fall seasons (Figure 4). In addition, the Mean Square Skill Score (MSSS) of the forecasted precipitation was computed based on the Mean Square Error (MSE) between the forecasted precipitation and the observed precipitation for each season over the period of 20 years (equation (4.1)). The MSSS (Figure 5) represents the accuracy of the forecast against climatological precipitation. The positive MSSS values indicate that forecasted precipitations have higher skill (less error) in comparison to climatological precipitation.

$$MSSS_s = 1 - \frac{MSE(P_{t,s}^{fcst}, P_{t,s}^{obs})}{MSE(\overline{P_{t,s}^{obs}}, P_{t,s}^{obs})} \tag{4.1}$$

Where in this equation *t* refers to a given year and *s* denotes the season. Figure 5 clearly demonstrates that the ECHAM4.5 precipitation forecasts offer added value over

climatological precipitation during winter and fall seasons in most of the study area. However, improvement in precipitation forecasts over climatology is limited over spring and summer because of smaller internal variability of precipitation during these seasons.

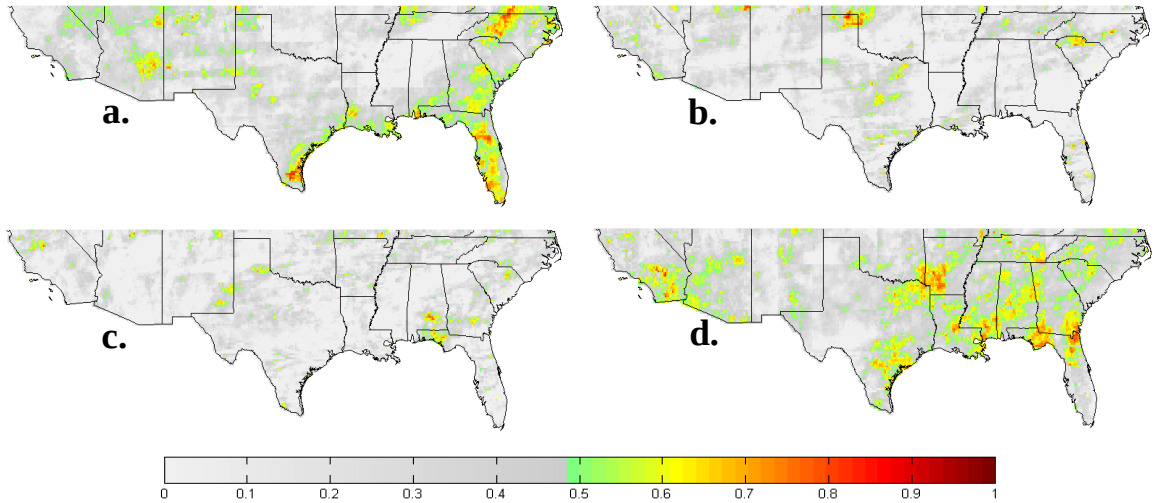

Figure 4. Spearman rank correlation between seasonal observed precipitation and seasonal (spatially downscaled and temporally disaggregated) forecasted precipitation at 1/8° spatial scale during (a) Winter, (b) Spring, (c) Summer, and (d) Fall. Gray values denote correlation (<0.47) that is not significant at 5% significance level.

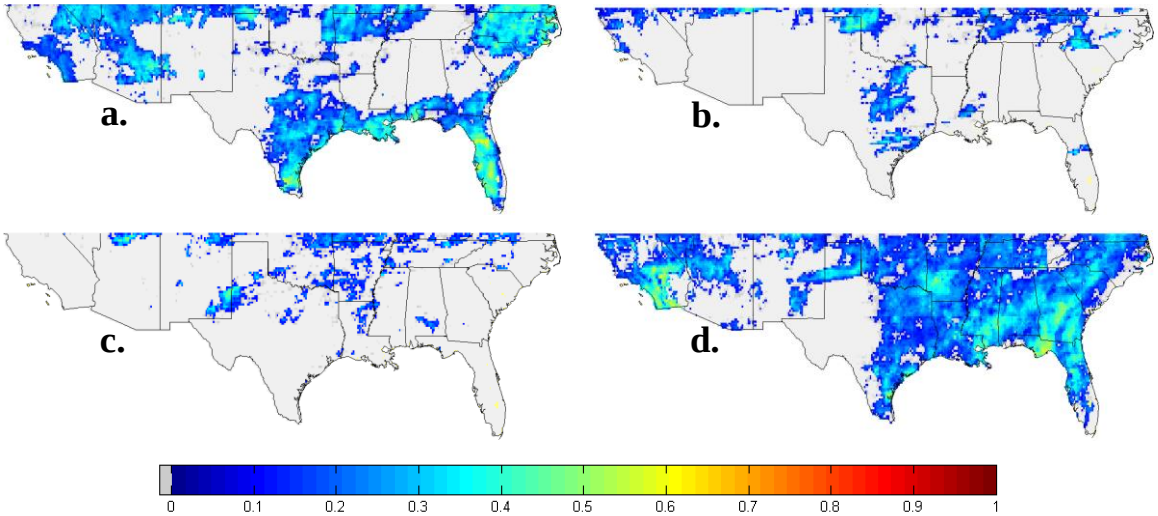

Figure 5. Mean Square Skill Score of ECHAM4.5 precipitation forecasts (after spatial downscaling and temporal disaggregation) against climatology during (a) Winter, (b) Spring, (c) Summer, and (d) Fall seasons. Gray color means the forecasted precipitation does not have skill in comparison with climatology.

4.1.2. Skill of Streamflow Forecasts

As described earlier, seasonal streamflow forecasts were developed by implementing three different LSMs using climatological non-precipitation forcings with downscaled and disaggregated ECHAM4.5 forecasted precipitation forcings. In order to estimate the skill of forecasted streamflows, we computed Mean Square Skill Score (MSSS) (Figure 6) based on the Mean Square Error (MSE) between the forecasted streamflow (Q^{fcst}) and the reference streamflow (Q^{ref}) over 20 years (equation (4.2)). The MSSS was computed against climatological streamflows, which were computed based on the mean seasonal streamflow for each 1/4°×1/4° grid.

$$MSSS_{m,s} = 1 - \frac{MSE(Q_{t,m,s}^{fcst}, Q_{t,m,s}^{ref})}{MSE(\overline{Q_{t,m,s}^{ref}}, Q_{t,m,s}^{ref})} \tag{4.2}$$

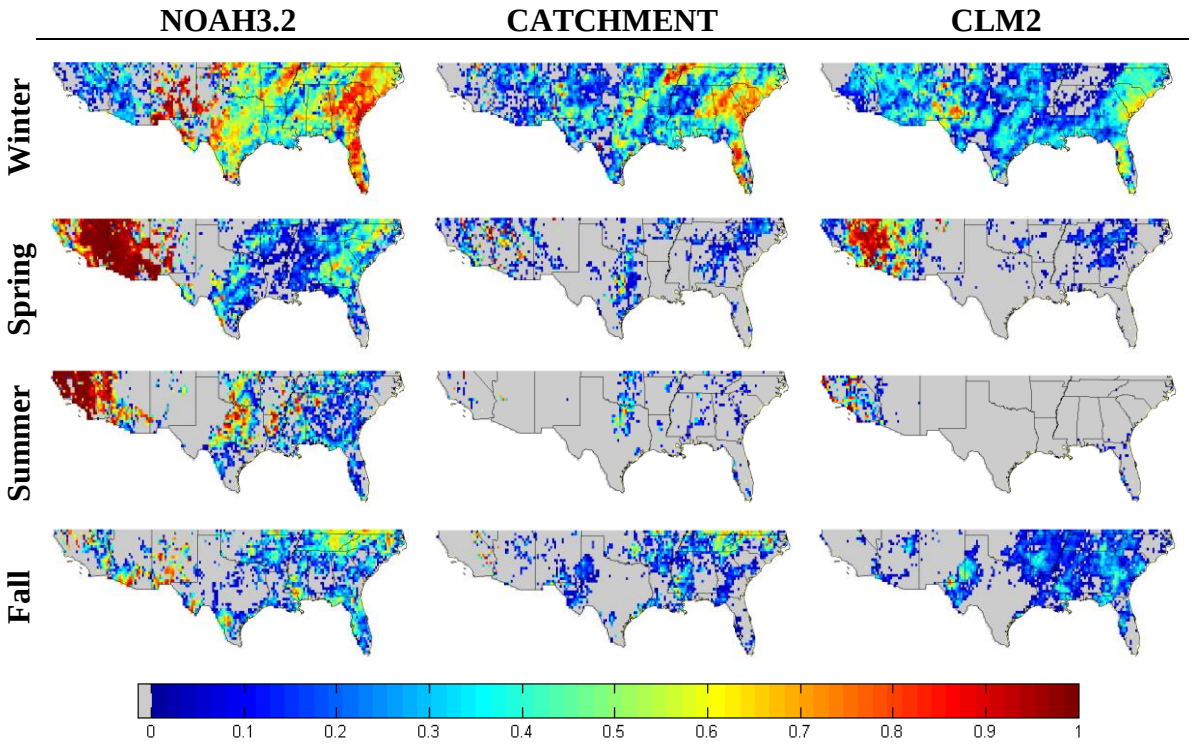

Figure 6. MSSS of seasonal streamflow forecasts (Q^{fcst}). Each column indicates a LSM and each row indicates a specific season. Gray color indicates no improvement in developed streamflow forecasts over climatology.

Figure 6 indicates that streamflow forecasts during winter are more skillful than the other seasons. During this season, the skill over the Southeast and central US are relatively higher than the skill over the West. However, the western US exhibit better skill during the spring and summer seasons. This may be due to the dominant role of initial conditions such as snow storages over this region. Furthermore, the role of initial conditions decreases during fall resulting in relatively lower skill than other seasons. Overall, the skill of NOAH3.2 model is better than other models. In spring, we see higher skill over the western US as opposed to the eastern US which is primarily due to the snowmelt. Further, low skill during the fall season is primarily due to the limited role of initial conditions coming out of summer.

4.2. Decomposition of Sources of Errors in Seasonal Streamflow Forecasting

After obtaining streamflow simulations/forecasts based on different schemes described in section 3.4 under different schemes (Figure 3), we decomposed sources of errors in the form of Root Mean Square Errors (RMSE) using the metrics defined in section 3.5. In order to analyze the results over the Sunbelt, we normalized RMSEs of each LSM based on its own mean seasonal reference streamflow (Q^{ref}) to obtain Relative Root Mean Square Error (R-RMSE) (equation (4.3)).

$$R-RMSE_{m,s} = \frac{RMSE_{m,s}}{\overline{Q_{m,s}^{ref}}} \tag{4.3}$$

4.2.1. Error due to Disaggregation and Downscaling

Figure 7 shows R-RMSE associated with both spatial downscaling and temporal disaggregation and is obtained based on errors between reference streamflow (Q^{ref}) and simulated streamflow under "Spatial downscaling and temporal disaggregation" scheme (Q^{down_dis}).

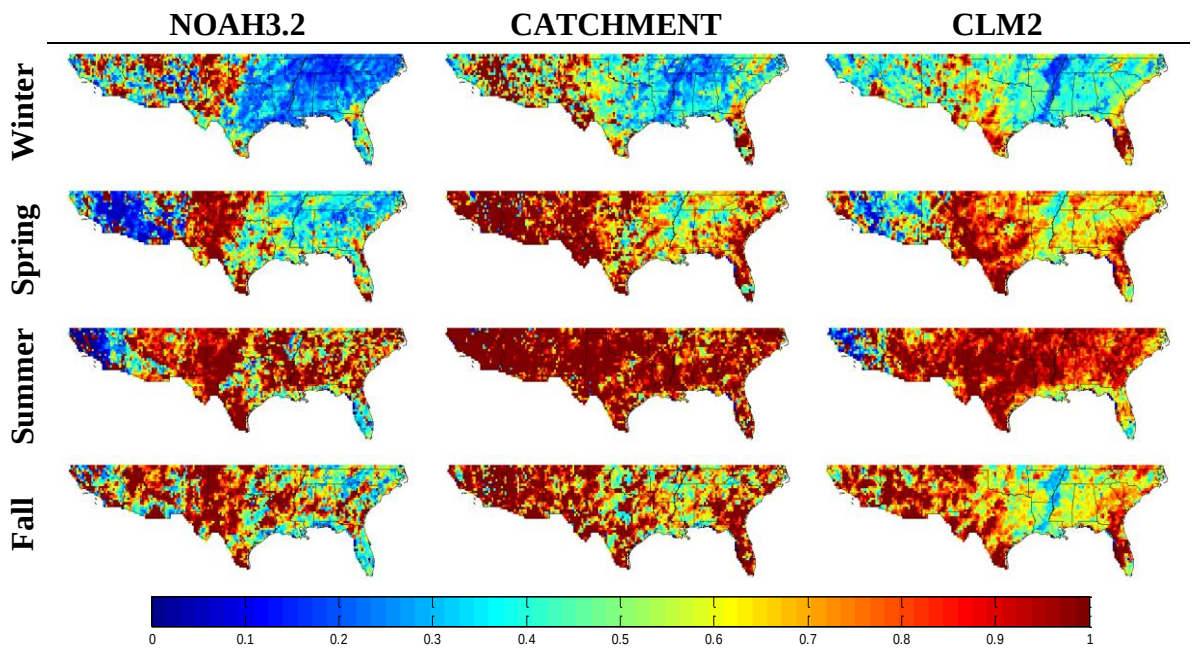

Figure 7. R-RMSE of streamflow simulations due to both downscaling and disaggregation

Based on Figure 7, the combined errors due to downscaling and disaggregation are highest during the summer and lowest during winter. In the winter season, regardless of which LSM was used, we see more errors in the West in comparison to the East. In contrast, the errors are lesser in the West in comparison to the East during spring and summer seasons due to dominant role of initial conditions. The errors are highest during the summer season over the entire study region, since local convection patterns are not well captured by both spatial downscaling and temporal disaggregation schemes.

To investigate the errors due to spatial downscaling and temporal disaggregation individually, we plotted Figure 8 and Figure 9. Comparing these two figures, we see that the dominant source of error is the temporal disaggregation procedure. Figure 8 illustrates that the Catchment LSM is very sensitive to downscaling procedure over the West. Further,

regardless of the LSMs, the errors increase from winter to summer and then decrease in the fall. Figure 9 shows that the highest error due to temporal disaggregation occurs in the summer. The disaggregation scheme dampens the peak rainfall events during summer convective storms and spread them over larger number of days resulting in higher errors. Particularly, high magnitude of errors can be found over the semiarid and the desert regions such as Texas and New Mexico, where higher variability in summer precipitation exists. In these regions, it is common to receive extreme precipitation events such as heavy downpours due to local convections, and such events are not well captured in the temporal disaggregation scheme.

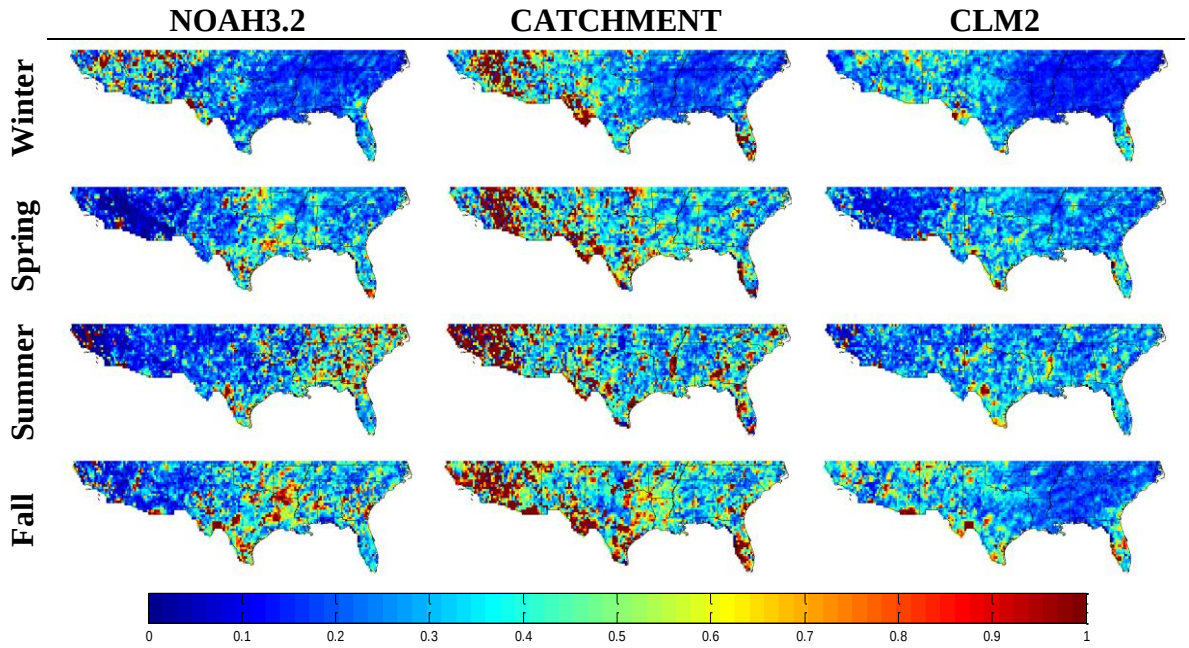

Figure 8. R-RMSE of streamflow simulations due to downscaling process alone (equation 3.4)

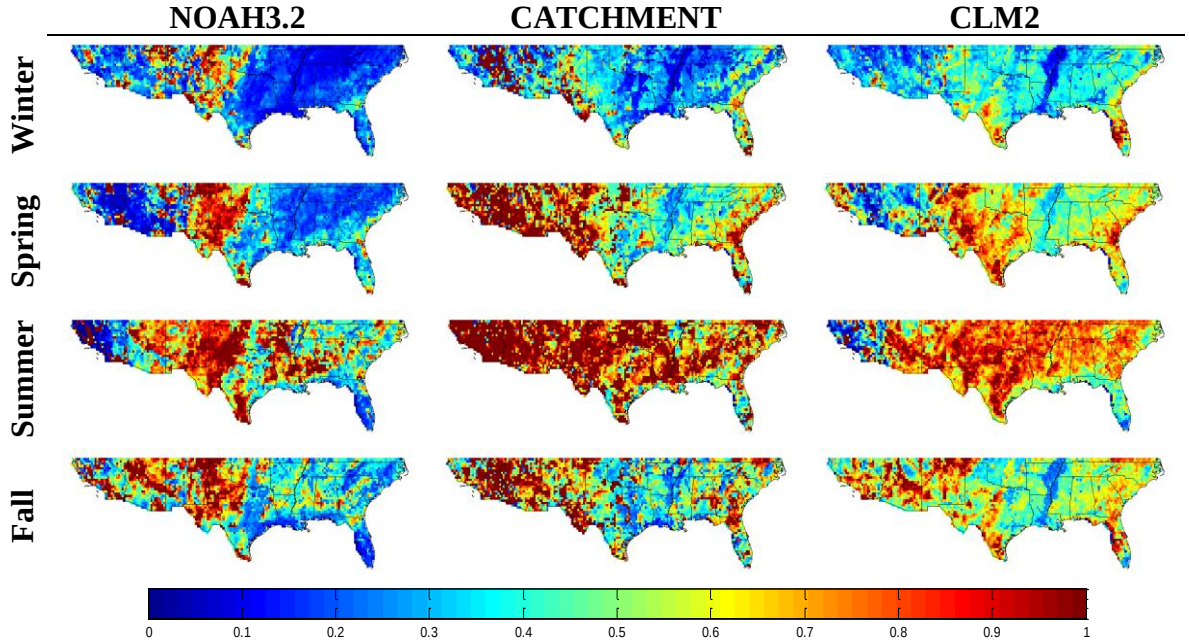

Figure 9. R-RMSE of streamflow simulations due to disaggregation process alone (equation 3.2)

4.2.2. Errors due to Climatological Forcings

Figure 10 shows R-RMSE of forecasted streamflows when the LSMs are implemented with climatological forcings. As expected, the errors are high in most of the study region. Surprisingly, during the spring and summer seasons we see low errors over the West, indicating the skill arising from the updated initial conditions.

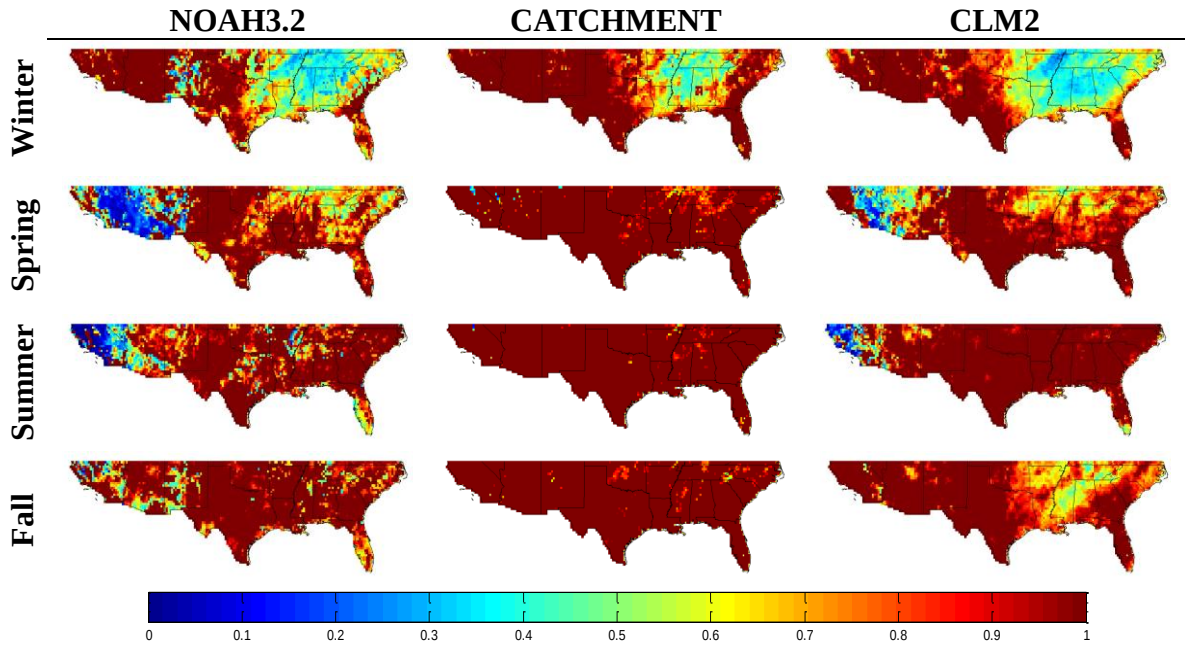

Figure 10. R-RMSE of streamflow simulations obtained using climatological forcings of all variables (equation 3.7)

As discussed in section 3.5.4, we assumed that the errors due to climatological forcings consist of two different sources of errors: 1) *errors due to climatological precipitation forcings* (Figure 11), which means that the only source of error is precipitation and there are no errors associated with other variables (all the variables excluding precipitation obtained

from NLDAS-2 dataset and precipitation forcings obtained from climatology) and 2) *errors due to climatological forcings excluding precipitation* (Figure 12), which means that errors are caused only by the climatological variables excluding precipitation, and there are no errors due to precipitation (observed precipitation were used in this scheme). Comparing between figures Figure 11 and Figure 12, we see that using climatological precipitation forcings and NLDAS-2 forcings other than precipitation, add higher errors in the simulated streamflow in comparison to the streamflow obtained by using observed precipitation and climatological NLDAS-2 forcings other than precipitation. Since precipitation is the primary driver of streamflow, we see limited skill in predicting streamflow particularly when climatological precipitation is used. The forecasting skills over the regions (Figure 11) such as Mississippi River basin during the winter season as well as the West during spring and summer, basically arise from the updated initial hydrological conditions.

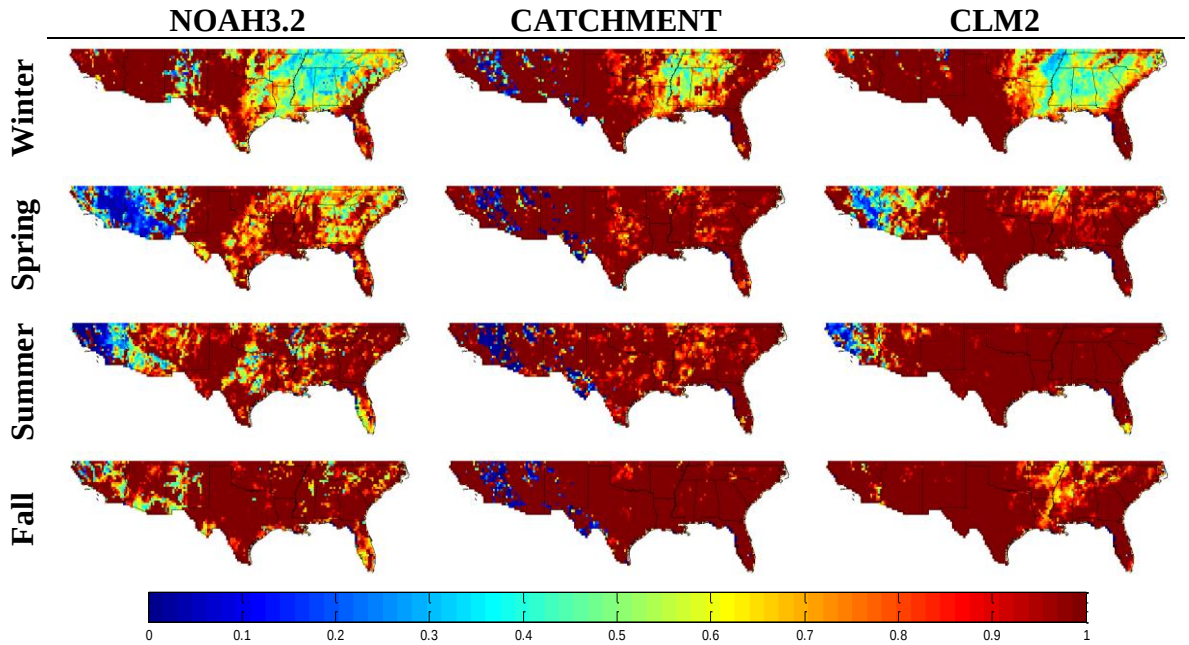

Figure 11. R-RMSE of streamflow simulations due to using climatological precipitation forcings (equation 3.11)

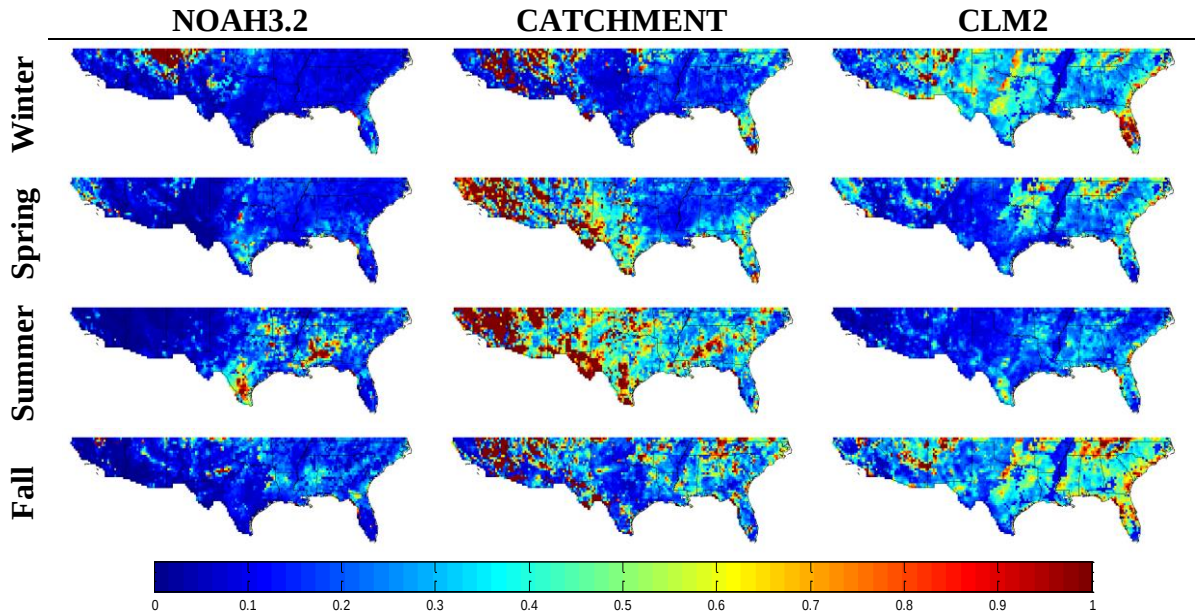

Figure 12. R-RMSE of streamflow simulations due to using climatological forcings for all variables except precipitation variable (equation 3.10)

4.2.3. Errors due to Large-Scale Precipitation Forecasts

Figure 13 illustrates the error due to the large-scale monthly ECHAM4.5 precipitation forecasts alone obtained by eliminating the role of spatial downscaling and temporal disaggregation in defining errors (equation 3.6). This figure is consistent with Figure 6, which shows the skill of streamflow forecasts based on the ECHAM4.5 precipitation forecasts. We can see that in most of the regions where there are low errors due to precipitation forecasts, we have high skill in streamflow forecasts as well. Note that we cannot directly compare the results in Figure 13 with the precipitation forecasts skills (Figure 4 and Figure 5) since the streamflow forecasts are improved by updated initial conditions.

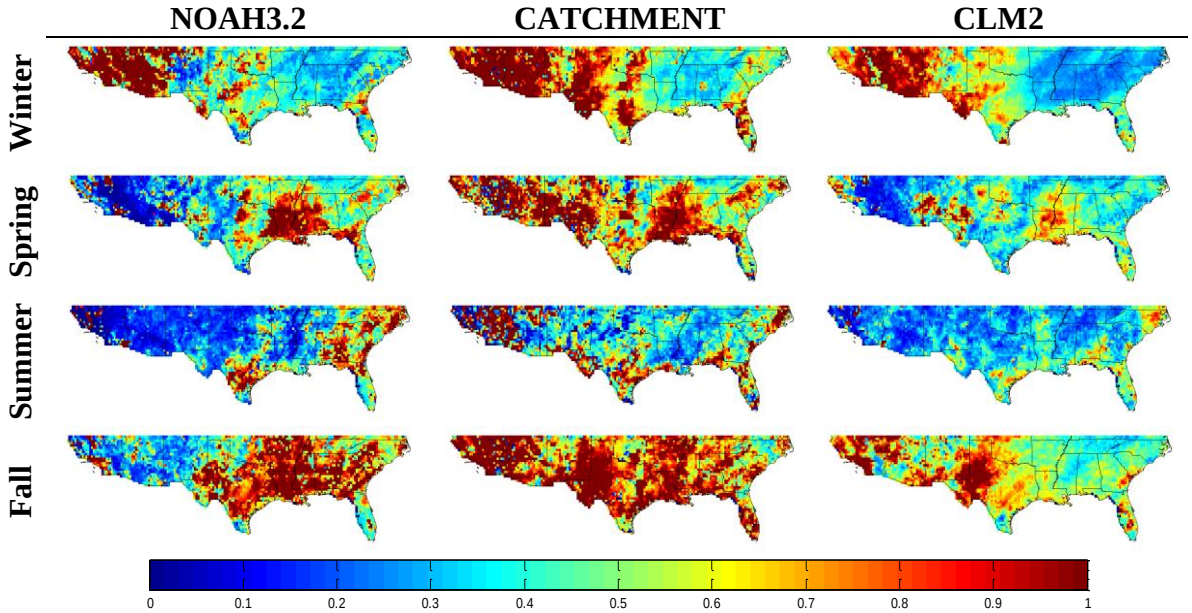

Figure 13. R-RMSE of streamflow simulations due to large-scale monthly ECHAM4.5 precipitation forecasts (equation 3.6)

We also infer that during the winter season, all the LSMs show better skill (lower R-RMSE) over the East in comparison to the West. Basically it is due to high skill in precipitation forecasts over the East brought by ENSO conditions. We also see poor performance of LSMs during the fall season over the East even though we have reliable skill in precipitation forecasts (Figure 4 and Figure 5). Recent studies have shown that LSMs exhibit high bias during the fall season and also due to the limited role of initial conditions in forecasting the flows during that time (Sinha et al. 2014).

4.3. River Basin Analyses

As mentioned in chapter 2, the study also focused on four selected river basins to evaluate the skill in forecasting observed seasonal streamflows. In this regard, retrospective streamflow forecasts (Q^{fcst}) were developed and evaluated based on the observed streamflow records obtained from the USGS data over the period 1991-2010. First, we estimated monthly %bias for each LSM by comparing the model simulated streamflows under the observed forcings scheme (reference streamflows Q^{ref}) to the observed streamflows over the period 1980-1991. These monthly bias corrections were applied to each LSM when forced with the ECHAM4.5 precipitation forecasts (Q^{fcst}) over the analysis period from 1991 to 2010. The time-series of bias-corrected streamflow forecasts as well as observed streamflow over the selected basins during different seasons from the three LSMs is presented in Figure 14. Overall, the streamflow forecasts exhibit similar pattern as the streamflow observations during the winter and spring seasons for all the selected basins.

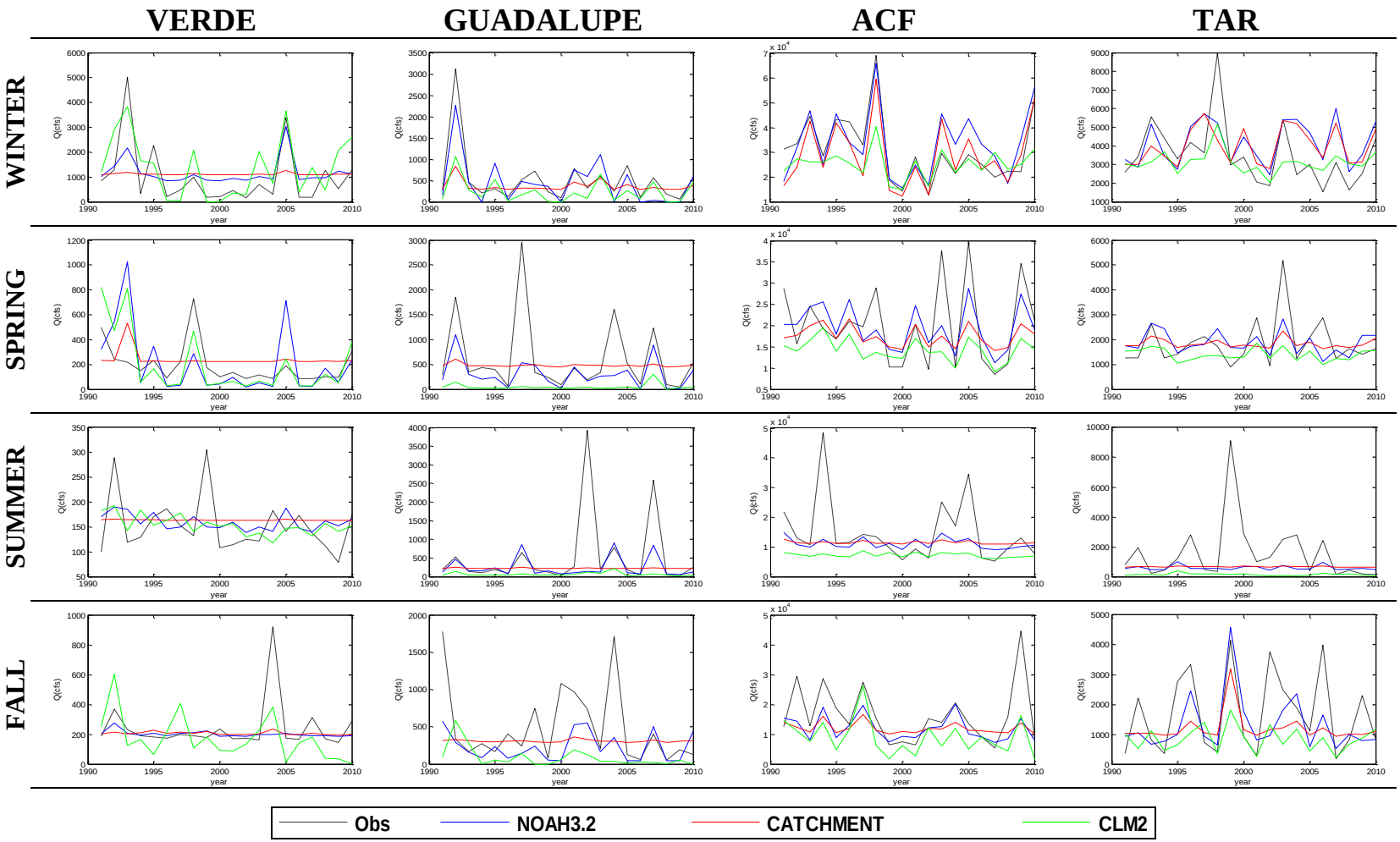

Figure 14. Evaluation of seasonal streamflow forecasts (Q^{fcst}) obtained by three LSMs in predicting observed seasonal streamflow over the four selected basins during different seasons.

In order to quantify the skill in forecasting streamflow, Figure 15 indicates Spearman's rank correlation between bias-corrected seasonal streamflow forecasts and seasonal observed streamflow over the period of 20 years (from 1991 to 2010). The black line in Figure 15 (SL) represents the significant correlation corresponding to 95% confidence level.

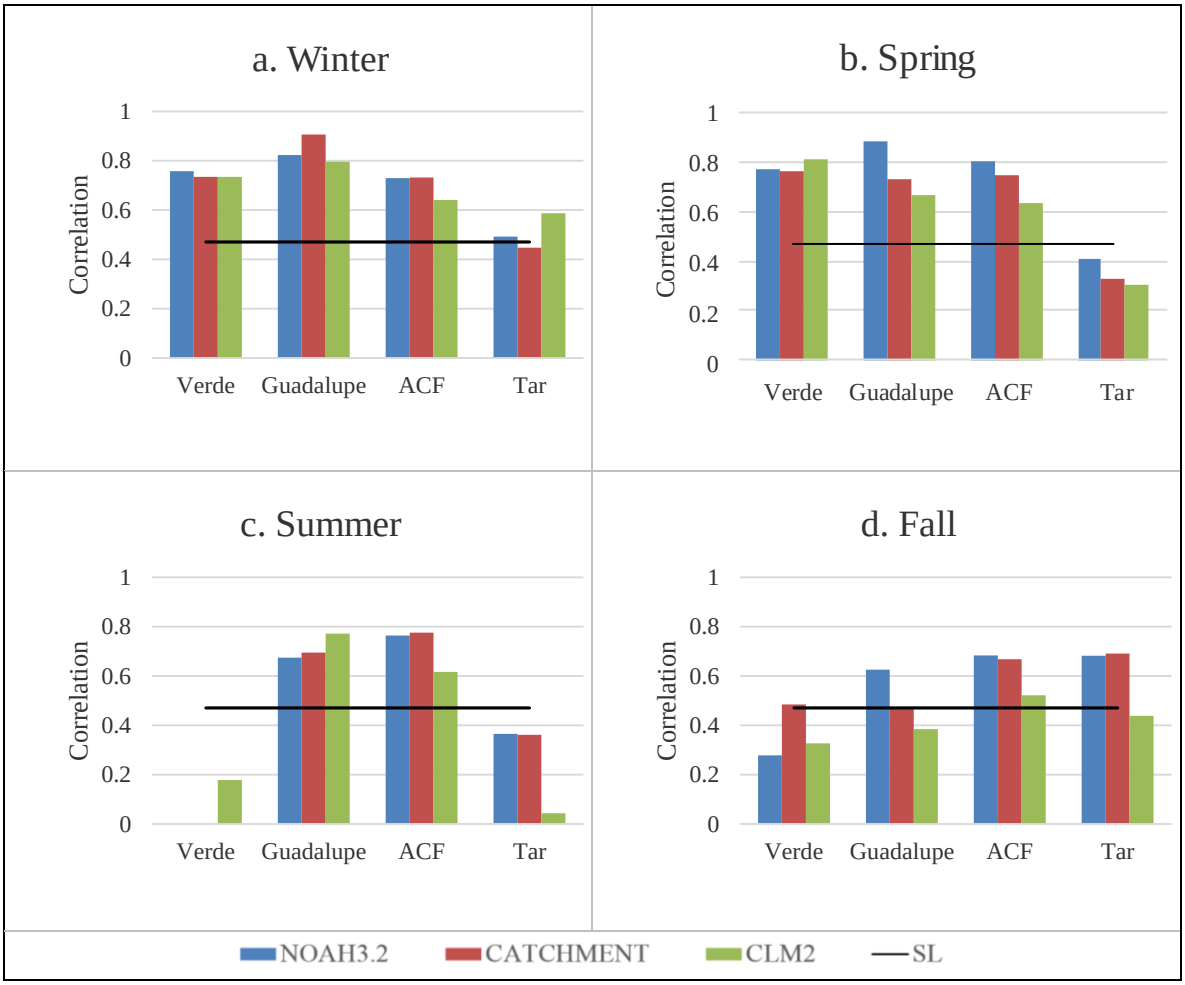

Figure 15. Spearman's rank correlation between forecasted streamflow (after bias correction) and observed streamflow shown over four selected basins during different seasons. The black line indicates significant correlation corresponding to 95% confidence level.

We infer that most of the correlation values are statistically significant, indicating the forecasted flows explain the interannual variability of observed flows.

To evaluate the accuracy of streamflow forecasts by each LSM, we computed MSSS for each river basin during different seasons based on equation (4.4) (Figure 16).

$$MSSS_{m,s} = 1 - \frac{MSE(Q_{t,m,s}^{fcst}, Q_{t,s}^{obs})}{MSE(\overline{Q_{t,s}^{obs}}, Q_{t,s}^{obs})} \tag{4.4}$$

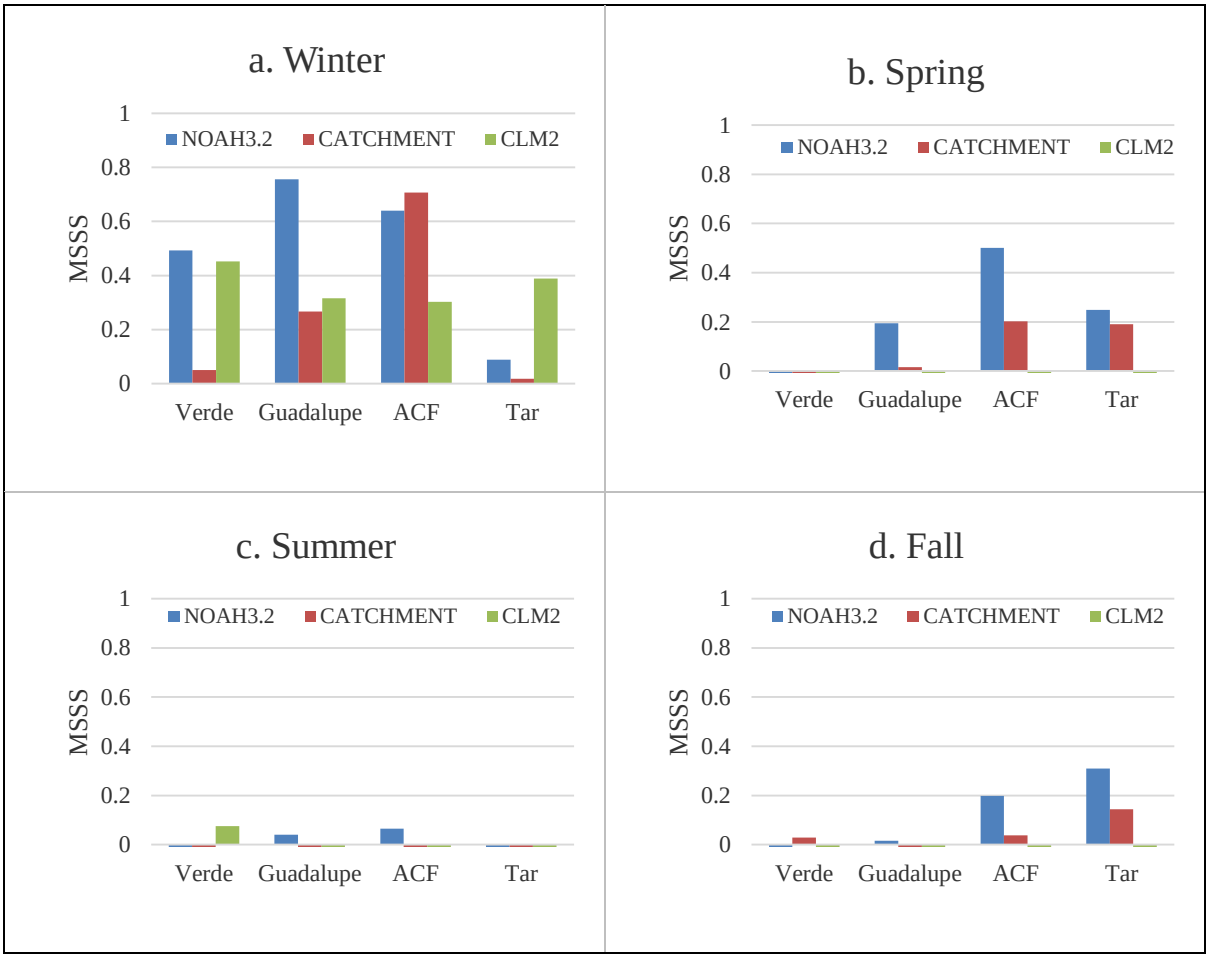

Figure 16. MSSS of forecasted seasonal streamflow obtained using ECHAM4.5 precipitation forecasts over four target basins

Figure 16 indicates that seasonal streamflow forecasts have the highest skill during the winter season, and the lowest skill during the summer season. The reason is that the ECHAM4.5 precipitation forecasts have the best skill during October to March (Figure 4 and Figure 5), which primarily arise from ENSO conditions. The Verde River basin which is the driest basin among all the selected basins shows the lowest skill and the ACF River basin which has the largest drainage area shows the highest skill. Compared across different LSMs we see that the NOAH3.2 model performs better than the other two LSMs.

CHAPTER 5
SUMMARY AND CONCLUSIONS

5.1. Summary

By using spatially downscaled and temporally disaggregated precipitation forecasts from ECHAM4.5 GCM along with hourly climatological forcings from the NLDAS-2 dataset, this study developed hindcasts of seasonal streamflows for a period of 20 years (1991-2010) derived from three LSMs. Using these hindcasts, we performed a quantitative assessment of the different sources of errors in developing seasonal streamflow forecasts. To quantify these errors, our study employed Root Mean Square Error (RMSE) due to different processing techniques (i.e., Temporal Disaggregation, and Spatial Downscaling) and due to large-scale ECHAM4.5 precipitation forecasts, and climatological forcings. Our analysis based on the decomposition shows that disaggregation procedure introduces more errors in streamflow forecasting in comparison to spatial downscaling approach. The maximum and minimum errors due to downscaling and disaggregation procedures were found during the summer and the winter respectively. The error due to the disaggregation approach alone was dominant during dry seasons (i.e., summer) and dry regions (i.e., semiarid and desert regions) because of poor performance of the disaggregation method in reproducing the mixed nature (i.e., discrete and continuous) distribution of precipitation. Analysis on errors arises from climatological forcings revealed that using climatological precipitation forcings produce significant errors in the developed streamflows in comparison to the errors arising from the

climatological forcings other than precipitation. In addition, we realized that some regions (e.g., the West) provide good skill in streamflow forecasting purely because of the updated initial hydrological conditions (IHCs) even when the models are forced with climatological precipitation forcings. Errors arise from large-scale precipitation forecasts provide information on skill in streamflow forecasting in a given season over a region. During the winter season, since skill in precipitation forecasts is high due to ENSO conditions, we have a good skill in streamflow forecasting as well over the entire Sunbelt.

5.2. Conclusions

As mentioned earlier, under streamflow forecasting with LSMs, four sources of errors (i.e., temporal disaggregation, spatial downscaling, climatological forcings excluding precipitation, and large scale ECHAM4.5 precipitation forecasts) were considered in quantifying the errors arises in developing streamflow forecasts. To summarize, we computed the relative magnitude of dominant source of error (maximum R-RMSE among other sources) with respect to total errors (divided be sum of R-RMSEs) in streamflow forecasts (Figure 17). Over the entire study area, we infer that the most dominant sources of uncertainties during different seasons are ECHAM4.5 precipitation forecasts and temporal disaggregation.

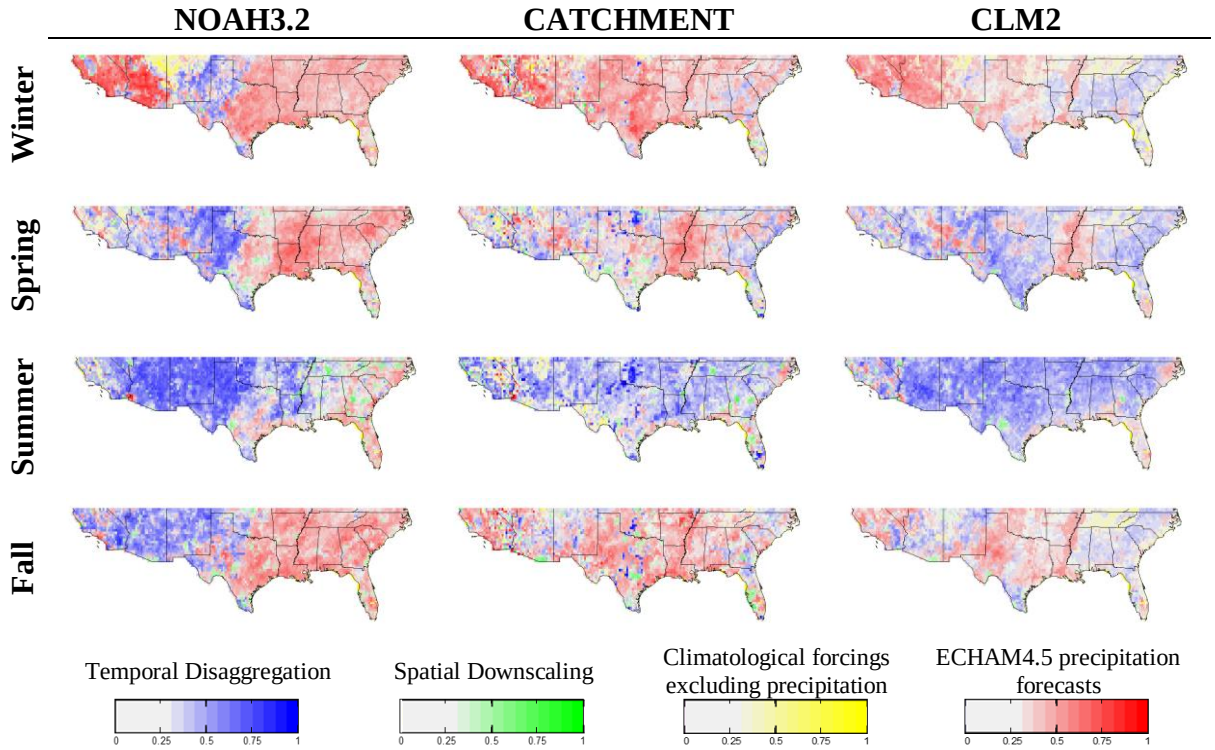

Figure 17. Relative magnitude of dominant source of errors with respect to the total uncertainty in seasonal streamflow forecasts over different seasons (rows) by three LSMs (columns). Darkness of each color denotes the magnitude of dominancy of the source of uncertainty among other sources (white color infers no dominancy which means all the four sources have equal contributions in total uncertainty)

As discussed in previous chapter, there is a low contribution of error sources, such as spatial downscaling and climatological forcings excluding precipitation, in contributing to the errors in streamflow forecasting. During the winter and fall seasons, the dominant source of error is imprecise ECHAM4.5 precipitation forecasts. This shows that in order to improve streamflow forecasting skills during winter and fall, the preference is to improve climate forecasts from GCMs. On the other hand, during the summer season, we see temporal disaggregation as the dominant source of uncertainty in most of the regions. In dry seasons, the inability of the disaggregation model to reproduce the peak rainfall, flatten the variability

of precipitation, thus results in higher errors. Consequently, improving the disaggregation technique skill is needed to improve the skill in summer streamflow forecasting.

5.3. Scope for Future Works

Considering the findings of this study, we outline the scope for future studies as following:

1) One of the steps in developing streamflow forecasts where a huge source of errors are associated with, is the temporal disaggregation. Thus improving temporal disaggregation techniques should be a focal point in future. Improvements in developing disaggregation approaches that minimize the errors in daily rainfall frequency distribution will be of interest.

2) Using ensemble combination of multiple General Circulation Models (GCMs) to improve forecasted forcings is another possibility to improve streamflow forecasts. In our study, we only used one GCM (ECHAM4.5) to obtain precipitation forecasts while the success of multi-model ensemble combination has been demonstrated in many studies (Weigel et al. 2008).

3) By expanding this study to the entire continental US, we can develop more insights on regions with snow-melt regimes.

4) This study focused on errors in seasonal streamflow forecasting. Our suggestion is to consider monthly updated climate forecasts for further research as this will provide information on the lead times in seasonal forecasting.

REFERENCES

Betts, R.A. et al., 1997. Contrasting physiological and structural vegetation feedbacks in climate change simulations. *Nature*, 387, pp.796–799.

Beven, K., 1997. TOPMODEL: a critique. *Hydrological processes*, 11(9), pp.1069–1085.

Bonan, G.B. et al., 2002. The land surface climatology of the Community Land Model coupled to the NCAR Community Climate Model. *Journal of Climate*, 15(22).

Chen, F. et al., 1996. Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 101(D3), pp.7251–7268.

Devineni, N., Sankarasubramanian, A. & Ghosh, S., 2008. Multimodel ensembles of streamflow forecasts: Role of predictor state in developing optimal combinations. *Water resources research*, 44(9).

Ek, MB et al., 2003. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 108(D22).

Famiglietti, JS & Wood, E., 1991. Evapotranspiration and runoff from large land areas: Land surface hydrology for atmospheric general circulation models. In *Land Surface—Atmosphere Interactions for Climate Modeling*. Springer, pp. 179–204.

Famiglietti, JS & Wood, E., 1994. Multiscale modeling of spatially variable water and energy balance processes. *Water Resources Research*, 30(11), pp.3061–3078.

Hamlet, A.F., Huppert, D. & Lettenmaier, D.P., 2002. Economic value of long-lead streamflow forecasts for Columbia River hydropower. *Journal of Water Resources Planning and Management*, 128(2), pp.91–101.

Hartmann, H.C. et al., 2002. Confidence builders: Evaluating seasonal climate forecasts from user perspectives. *Bulletin of the American Meteorological Society*, 83(5).

Hayhoe, K. et al., 2004. Emissions pathways, climate change, and impacts on California. *Proceedings of the National Academy of Sciences of the United States of America*, 101(34), pp.12422–12427.

Hoffman, F.M. et al., 2005. Vectorizing the community land model. *International Journal of High Performance Computing Applications*, 19(3), pp.247–260.

Koster, R.D. et al., 2000. A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 105(D20), pp.24809–24822.

Koster, R.D. & Suarez, M.J., 1995. Relative contributions of land and ocean processes to precipitation variability. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 100(D7), pp.13775–13790.

Kumar, S.V. et al., 2006. Land information system: An interoperable framework for high resolution land surface modeling. *Environmental modelling \& software*, 21(10), pp.1402–1415.

Lall, U. & Sharma, A., 1996. A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resources Research*, 32(3), pp.679–693.

Leung, L Ruby et al., 2004. Mid-century ensemble regional climate change scenarios for the western United States. *Climatic Change*, 62(1-3), pp.75–113.

Li, H. et al., 2009. The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 114(D4).

Li, W. & Sankarasubramanian, A., 2012. Reducing hydrologic model uncertainty in monthly streamflow predictions using multimodel combination. *Water Resources Research*, 48(12).

Di Luca, A., de El'\ia, R. & Laprise, R., 2012. Potential for added value in precipitation simulated by high-resolution nested Regional Climate Models and observations. *Climate dynamics*, 38(5-6), pp.1229–1247.

Luo, L. & Wood, E.F., 2008. Use of Bayesian merging techniques in a multimodel seasonal hydrologic ensemble prediction system for the eastern United States. *Journal of Hydrometeorology*, 9(5).

Luo, L., Wood, E.F. & Pan, M., 2007. Bayesian merging of multiple climate model forecasts for seasonal hydrological predictions. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 112(D10).

Mahanama, S. et al., 2012. Soil moisture, snow, and seasonal streamflow forecasts in the United States. *Journal of Hydrometeorology*, 13(1), pp.189–203.

Maurer, E. et al., 2002. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *Journal of Climate*, 15(22).

Maurer, E. & Hidalgo, H., 2008. Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods. *Hydrology \& Earth System Sciences*, 12(2).

Maurer, E.P. & Lettenmaier, D.P., 2003. Predictability of seasonal runoff in the Mississippi River basin. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 108(D16).

Maurer, E.P., Lettenmaier, D.P. & Mantua, N.J., 2004. Variability and potential sources of predictability of North American runoff. *Water Resources Research*, 40(9).

Mitchell, K.E. et al., 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 109(D7).

Murphy, J., 1999. An evaluation of statistical and dynamical techniques for downscaling local climate. *Journal of Climate*, 12(8).

Niu, G.-Y. et al., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 116(D12).

Prairie, J. et al., 2007. A stochastic nonparametric technique for space-time disaggregation of streamflows. *Water Resources Research*, 43(3).

Sankarasubramanian, A., Lall, U. & Espinueva, S., 2008. Role of Retrospective Forecasts of GCMs Forced with Persisted SST Anomalies in Operational Streamflow Forecasts Development. *Journal of Hydrometeorology*, 9(2).

Shukla, S. & Lettenmaier, DP, 2011. Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill. *Hydrology \& Earth System Sciences Discussions*, 8(4).

Sinha, T. & Sankarasubramanian, A., 2013. Role of climate forecasts and initial conditions in developing streamflow and soil moisture forecasts in a rainfall-runoff regime. *Hydrology and Earth System Sciences*, 17(2), pp.721–733.

Sinha, T., Sankarasubramanian, A. & Mazrooei, A., 2014. Decomposition of Sources of Errors in Monthly to Seasonal Streamflow Forecasts in a Rainfall-Runoff Regime. *Journal of Hydrometeorology (under review)*.

Slack, J., Lumb, A. & Landwehr, J., 1993. Hydro-Climate Data Network (HCDN)—Steamflow Data Set, 1874-1988: US Geological Survey Water-Resources Investigations Report 93-4076.

Stieglitz, M. et al., 1997. An efficient approach to modeling the topographic control of surface hydrology for regional and global climate modeling. *Journal of Climate*, 10(1), pp.118–137.

Weigel, A., Liniger, M. & Appenzeller, C., 2008. Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Quarterly Journal of the Royal Meteorological Society*, 134(630), pp.241–260.

Wilby, R.L. et al., 2000. Hydrological responses to dynamically and statistically downscaled climate model output. *Geophysical Research Letters*, 27(8), pp.1199–1202.

Wood, A.W. et al., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic change*, 62(1-3), pp.189–216.

Wood, A.W. et al., 2002. Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research: Atmospheres (1984-2012)*, 107(D20), p.ACL–6.

Wood, A.W. & Lettenmaier, D.P., 2006. A test bed for new seasonal hydrologic forecasting approaches in the western United States. *Bulletin of the American Meteorological Society*, 87(12), pp.1699–1712.

Yuan, X. et al., 2011. A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophysical Research Letters*, 38(13).