Potential Predictability of Streamflow and Soil Moisture in a Humid Alabama-Coosa-Tallapoosa River Basin using the National Water Model

by

Yanan Duan

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
December 14, 2019

Keywords: WRF-Hydro, soil moisture initial states, climate forcing, streamflow predictability

Approved by

Sanjiv Kumar, Assistant Professor of Forestry and Wildlife Sciences
Jasmeet Lamba, Assistant Professor of Biosystems Engineering
Latif Kalin, Professor of Forestry and Wildlife Sciences
Puneet Srivastava, Professor of Biosystems Engineering
William Batchelor, Professor of Biosystems Engineering

Abstract

This study investigates predictability of daily streamflow and soil moisture in the Alabama-Coosa-Tallapoosa (ACT) river basin in southeastern United States. The study employs the state-of-the-art National Water Model (NWM), and compares and contrasts effects of initial soil moisture condition with that of seasonal climate forecast on streamflow and the soil moisture forecasts skill. We had designed and implemented seasonal streamflow forecast ensemble experiments following the methodology suggested by Dirmeyer et al. (2013). The study also compares the soil moisture variability in the NWM with the in-situ observations and remote sensing data from the Soil Moisture Active and Passive (SMAP) satellite.

The NWM skillfully simulates the observed streamflow in the ACT basin from 1990 to 2018. The soil moisture variability is 46% smaller in the NWM compared with the SMAP data, mainly due to a weaker amplitude of seasonal cycle in the NWM. This study finds that initial soil moisture condition is a major source of predictability for the seasonal streamflow forecast. Contribution of the initial soil moisture condition is comparable or even higher than that of seasonal climate anomalies effects in the dry seasons. Specifically, in the boreal summer season, the initial soil moisture condition contributes to 65%, and 48% improvements in the seasonal streamflow, and soil moisture forecast skill, respectively. This study attributes a higher improvement in the streamflow forecast skill to the lag effects between the soil moisture and streamflow anomalies. Results of this study can inform development and improvement of the operational streamflow forecasting system.

Acknowledgments

My special gratitude to my major advisor, Dr. Sanjiv Kumar, for providing support and guidance that helped me conduct the research presented in this thesis. I would like to thank the members of M.S. committee, Dr. William Batchelor, Dr. Latif Kalin, Dr. Puneet Srivastava, and Dr. Jasmeet Lamba, for their comments during the course of this research. I would also like to thank Auburn University Hopper Cluster (hpcadmin@auburn.edu) for their help in installing WRF-Hydro v3 and v5. I am thankful to WRF-Hydro support team (wrfhydro@ucar.edu) for support of this work.

Auburn University Intramural Grant funding (IGP VPR Project 180286 - 2018) supported this research.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACT	Alabama-Coosa-Tallapoosa
CF	Climate Forcing
CRN	Climate Reference Network
CTRL	Control
ESP	Ensemble Streamflow Prediction
ISMS	Initial Soil Moisture States
NCAR	National Center for Atmospheric Research
NLDAS-2	North American Land Data Assimilation System version-2
NHD Plus	National Hydrography Dataset Plus
NWM	National Water Model
Noah-MP	Noah Multi-parameterization
LSM	Land Surface Model
NSE	Nash Sutcliff Efficiency
Rev-ESP	Vice-versa Ensemble Streamflow Prediction
RMSD	Root mean square difference
SCAN	Soil Climate Analysis Network
SMAP	Soil Moisture Active and Passive
USGS	United States Geological Survey
WRF	Weather Research and Forecasting

WPS	Weather Research and Forecasting Preprocessing System
SWAT	Soil & Water Assessment Tool

CHAPTER 1 INTRODUCTION

1.1 Background and Objectives

Water flowing through creeks, streams, and rivers are called streamflow. It represents an integrated response of the climate, land use, topography, soil type, ground water, and human intervention in the hydrologic system (Gochis et al., 2018; Kam & Sheffield, 2016; Kumar, Merwade, Kam, & Thurner, 2009; Kumar et al., 2016; Milly, Kam, & Dunne, 2018; Guo-Yue Niu, Yang, Dickinson, & Gulden, 2005; Schilling, 2016; Xiao, Udall, & Lettenmaier, 2018; Zhang & Schilling, 2006). Precipitation is one of the major factors contributing to streamflow. Other factors that contribute to streamflow include snow storage and melt (Berghuijs, Woods, & Hrachowitz, 2014; D. Li, Lettenmaier, Margulis, & Andreadis, 2019), soil moisture memory (Koster, Mahanama, Livneh, Lettenmaier, & Reichle, 2010; S. Mahanama, Livneh, Koster, Lettenmaier, & Reichle, 2012), and ground water contribution to the baseflow in the streams (Larocque, Fortin, Pharand, & Rivard, 2010; Priest, 2004). The soil moisture memory measures degree of persistence or dissipation of soil moisture anomalies through time (Koster & Suarez, 2001). Precipitation has a short memory that ranges from a few days to less than four weeks. Whereas, a common characteristic of the later processes is their long-memory that ranges from several months to a year or even longer and can potentially contribute to improving streamflow forecast skill on seasonal to inter-annual time scale (Kumar, Newman, Wang, & Livneh, 2019). We investigated this question by using the state-of-the-art National Water Model's (NWM) (https://water.noaa.gov/about/nwm) forecasting system. This is an important question because in the current system the forecast is limited to one month only. National Water Model is WRF-Hydro, are used interchangeably in the thesis.

The National Water Model (NWM) is an unprecedented effort by the NOAA to provide high-resolution soil moisture (1-km) and streamflow forecasts for 2.7 million river reaches in the continental United States. The NWM is comprised of the Noah Multi-parameterization land surface model (Noah-MP LSM), subsurface and terrain routing module, gridded diffusive wave channel and reservoir routing modules (Gochis et al., 2018; Hooper, Nearing, & Condon, 2017). A number of previous studies have investigated effects of the soil moisture memory, snow, and climate forcing using land model only experiments and found improved skill in the seasonal streamflow forecast (S. Mahanama et al., 2012; Shukla, Sheffield, Wood, & Lettenmaier, 2013; Wood et al., 2016). However, an integrated response of the coupled hydrologic prediction system, especially two way coupling between the land model and the subsurface and terrain routing module, in the NWM has not been investigated.

The seasonal streamflow forecast can be helpful for water resource planning and management including mitigation planning for drought impacts. Two main contributors of the seasonal streamflow forecast skill are: (1) climate forcing (CF), and (2) initial soil moisture states (ISMS) (S. P. Mahanama, Koster, Reichle, & Zubair, 2008). Since skill in the precipitation forecast (beyond two weeks) is limited, seasonal streamflow forecasts rely on initial hydrologic conditions. Among the initial hydrologic conditions, soil moisture is a dominant factor for improving forecast skill (Wanders et al., 2019). In Europe, Greuell et al (2019) found that ISMS contributed to 0-2 months lead forecast skill. Numerous studies have found that soil moisture memory is about 2-4 months (Amenu, Kumar, & Liang, 2005; Paul A Dirmeyer et al., 2016; Entin et al., 2000; Nicolai-Shaw, Gudmundsson, Hirschi, & Seneviratne, 2016; Orth & Seneviratne, 2012; Vinnikov, Robock, Speranskaya, & Schlosser, 1996; Wu, Geller, & Dickinson, 2002), thus ISMS can contribute to a more accurate forecast (Maurer & Lettenmaier, 2003).

The ensemble streamflow prediction (ESP) that generates streamflow forecast using observed initial conditions and climatological or randomized climate forcing. And contrary, i.e. random initial condition and realistic climate forcing (the so called vice-versa or Rev-ESP). They have been used to assess the role of initial hydrologic conditions on streamflow forecast skill (Wood et al., 2016). Li et al. (2009) found skill in the 2-month lead forecast due to initial conditions. The ESP and Rev-ESP methods have been used to assess the streamflow forecast skill globally, and found that ISMS influenced 1-month lead forecast, especially in the dry season (Shukla et al., 2013). They also found that climate forcing dominated the predictability in the wet equatorial regions. Yossef et al. (2013) found 1-2 months lead forecast skills during the wet season in the eastern US.

A second objective of this study is to assess the soil moisture states in the NWM using SMAP and in-situ observations. A knowledge gap exists in the area of calibrating the streamflow forecast model with the observed streamflow data, and assessing the fidelity of initial hydrologic condition by observations (Wanders et al., 2019). So this project, NWM initial soil moisture are compared with SMAP and SMERGE soil moisture data. SMERGE is reanalysis soil moisture merged Noah land surface model and surface soil moisture retrievals acquired from multi-sensor remote sensing products (Han, Crow, Holmes, & Bolten, 2014). Particularly, incorporation of human water management processes such as dams and reservoir operations are challenging. New developments in the National Water Model allow us to parametrize dam and reservoir operations in the model (Gochis et al., 2018). Further, real-time availability of the remotely sensed soil moisture observations, also called the SMAP (Soil Moisture Active and Passive) (Entekhabi et al., 2014), makes it feasible to incorporate observed hydrological condition in the forecast model (Felfelani, Pokhrel, Guan, & Lawrence, 2018).

Studies have investigated the relative importance of snow and soil moisture initial conditions on streamflow forecast skill, especially in the snow dominated regions of the US and Europe (Koster et al., 2010; Rene Orth & Sonia I Seneviratne, 2013). Koster et al. (2010) found that the early-season snow-water storage (snow initialization) contributes most to streamflow forecast skill in the mountain area basins in the Northwest US. For example, the snow water storage is the dominant source of predictability in the Columbia River basin (S. Mahanama et al., 2012). Similarly, Orth and Seneviratne (2013) found that the initial snow condition is a major contributor to streamflow forecast skill in high-altitude catchment (>2000m) in Switzerland. However, there are a limited number of studies that have investigated streamflow predictability in the humid southeast U.S., which is a rain-dominated region.

The long-term goal of such research is understanding the changing water dynamics in the Southeast. The southeastern U.S. is a hotspot of climate variability, land cover change, and irrigation expansion, especially in the recent decades. The southeastern U.S. has shown an anomalous longer temperature trend than the rest of the globe (Kumar et al. 2013; Meehl et al. 2015; Pan et al. 2013). From 1973-2011, the Southeastern plains had one of the highest rates of land use and land cover change of any ecoregion nationally (Napton et al. 2009). There is renewed pressure to increase production of the warm weather crops in the Southeast due to long-term drought in the west, particularly in California (Alston et al. 2010; McNider and Christy 2007). Since 1998, irrigation has increased by 60% in Alabama and 56% in Georgia, the largest increases among the 31 Eastern States (Schaible and Aillery 2012). Competing needs of water use has led to a tri-state conflict involving Alabama, Georgia and Florida (Center 2015).

This study investigates streamflow and soil moisture predictability in the Alabama-Coosa-Tallapoosa (ACT) River Basin, which is one of the two river basins involved in the tri-state water

conflict. The other basin, Apalachicola–Chattahoochee–Flint, is heavily irrigated (Mitra, Srivastava, & Singh, 2016; Singh, Mitra, Srivastava, Abebe, & Torak, 2017), and hence difficult to simulate. We employed the WRF-Hydro model version 5, which is the research version of the National Water Model (Gochis et al 2018). We investigated the predictability using random forcing and observed climate forcing which are both generated from the reanalysis dataset. Section 2 gives the description of the data, WRF-Hydro model configuration, and its parameterization, and predictability experiment design. Section 2 also describes model calibration and validation results. Section 3 presents findings of this study. Summary and discussions are presented in Section 4.

1.2 Thesis Organization

The thesis is organized in five chapters. Chapter One provides a literature review of the research work related to soil moisture and streamflow predictability. Chapter Two presents description of study area, and procedure to prepare the model. Chapter Three presents methodology. Chapter Four is result. Chapter Five is summary and conclusion.

CHAPTER 2 STUDY AREA, DATA AND MODEL DESCRIPTION

2.1 Study Area

The ACT River Basin is located in Georgia and Alabama in the Southeastern US. The ACT basin stretches between 31°- 35.2°N and is characterized by the humid subtropical climate using the Koppen-Geiger climate classification map (Peel et al. 2007). The highest elevation in the Northeastern basin is 1284 meters, and the lowest elevation touches to sea level in the southwest corner. The average slope of the basin is 5%. There are 15 active United States Geological Survey (USGS) streamflow gages in the basin (Table 1 and Figure 1a).

These gages were used as the sub-watershed outlets for the watershed delineation. The observed streamflow at the watershed outlet: Alabama River at Claiborne L&D near Monroeville (USGS gage # 02428400) is used for model calibration and validation. In addition, there are 15 reservoirs those are treated as one-orifice one-weir reservoirs for the model set up. It is a static model without reservoir management simulation.

Majority land cover type is forest (57%), followed by Grassland/Shrub (24%), and urban area (9%). Agriculture only occupies 3%, (Figure 1b and Figure 2). Irrigation water consumption is not significant in the ACT basin (Schaible and Aillery 2017), hence the model simulates streamflow very well (shown later). Two of the main surface soil textures are silt loam and sandy loam (Figure 1c). These soil types have good infiltration, the soil infiltration parameter is the most sensitive parameter in the model calibration (shown later).

Annual average temperature decreases from south to north (Figure 1d), but the temperature remain well above the freezing point all around the year and snow is not a major factor in the ACT basin. The mountainous range in the northeast corner receive the highest precipitation (>1500

mm/year). The annual precipitation decreases toward the center of the basin to ~1300 mm/year and then increases again to the 1500mm/year toward to Gulf coast (Figure 1d).

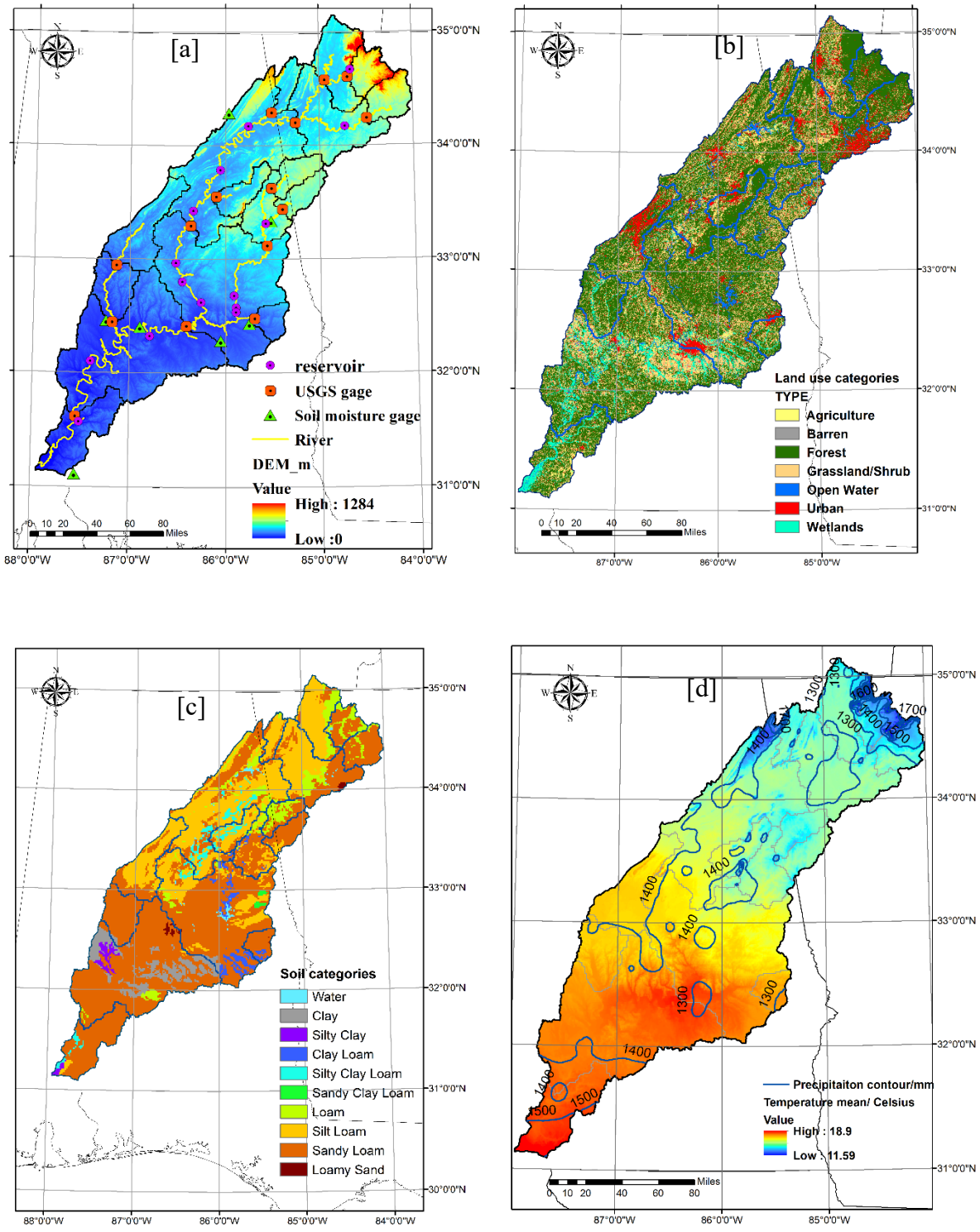


Figure 1: The ACT basin topography (a), land use texture (b), reclassified from NLCD 2011 land cover data, surface soil categories (c), and 30 years normal precipitation and temperature (d). Figure (a) also shows 15 sub-watersheds, 15 USGS gages, 7 SCAN and CRN soil moisture gages, and 15 reservoirs in the ACT basins.

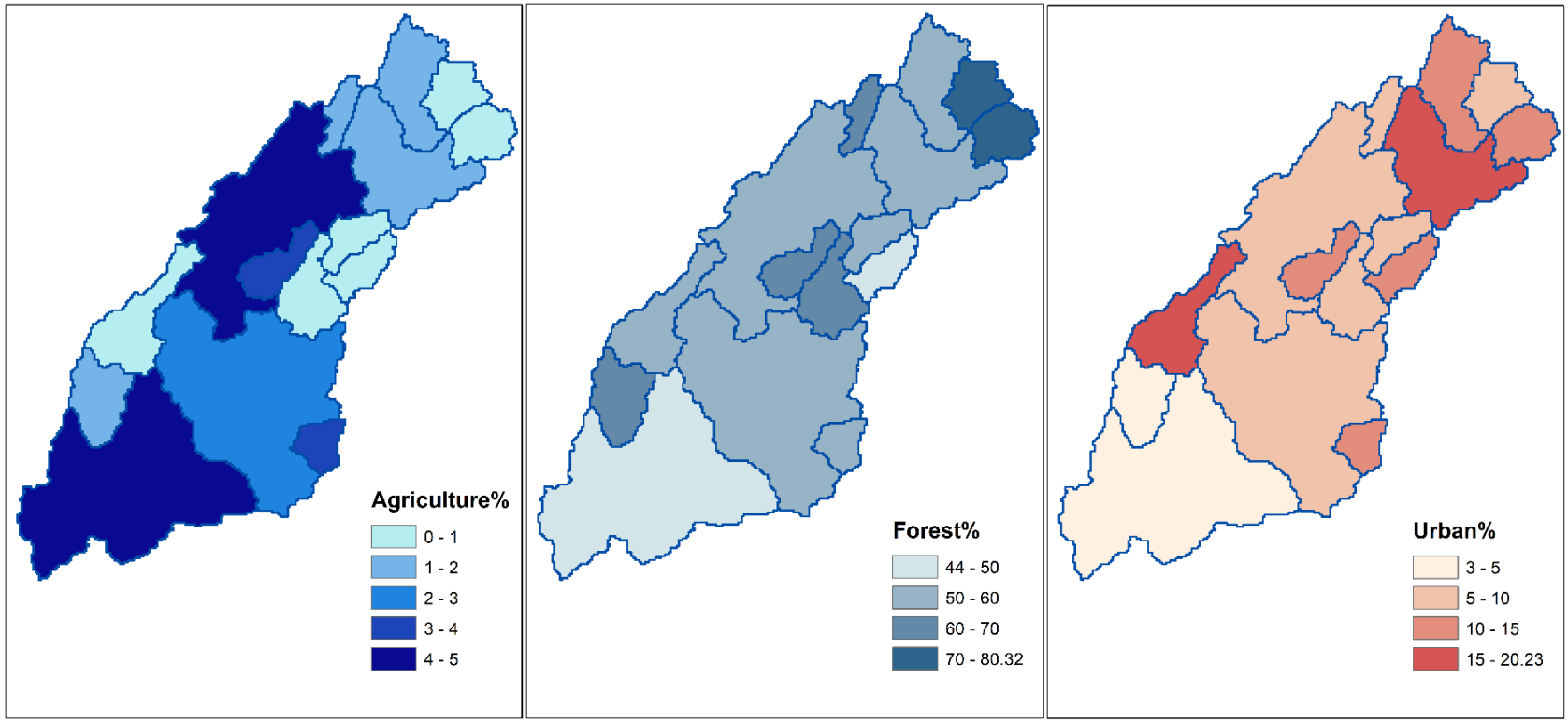


Figure 2: Agriculture, forest, and urban land cover percentiles in each sub-watershed.

Table 1: List of USGS streamflow gages in the ACT basin

FID	LON	LAT	Gage ID	Drainage area (square miles)	Observed flow average (cfs)	Data duration	runoff in mm/year
1	-84.696	34.604	02382500	521	962	1980-2016	637
2	-84.942	34.577	02387500	1602	2652	1980-2016	571
3	-85.509	34.290	02398300	366	594	1984-2016	560
4	-84.495	34.240	02392000	613	1105	1980-2016	621
5	-85.257	34.201	02397000	4040	6194	1980-2016	529
6	-85.513	33.623	02412000	448	585	1980-2016	451
7	-86.097	33.548	02404400	481	702	1984-2016	503
8	-85.399	33.437	02413300	406	544	1980-2016	462
9	-86.364	33.292	02407000	8392	10462	2011-2016	430
10	-85.561	33.117	02414500	1675	2470	1980-2016	509
11	-87.140	32.945	02424000	1027	1522	1980-2016	511
12	-85.695	32.477	02419000	333	384	1980-2016	397
13	-87.180	32.444	02425000	1766	2621	1980-2016	512
14	-86.408	32.411	02420000	15087	21348	1980-2016	488
15	-87.551	31.615	02428400	21473	29292	1980-2016	470

2.2 Modeling Tools

We had employed the WRF-Hydro Modeling System developed at the National Center for Atmospheric Research (NCAR), Boulder CO. We downloaded the latest version of the model: WRF-Hydro v5, WRF-Hydro ARC-GIS tools, create_wrfinput tool, and NLDAS regridding tool from NCAR. NCAR provides the Weather Research and Forecasting (WRF) Model v3.9.1 and its Preprocessing System (WPS). We use WPS to prepare some model set up files. In addition, we also used Soil & Water Assessment Tool (Arc-SWAT v10.4.1) to set up the model (Neitsch, Arnold, Kiniry, & Williams, 2011).

2.3 Model Input Data

The geospatial data used to set up the model include five data sources. (1) Digital elevation data from the National Elevation Dataset at 30-m resolution. (2) Soil-type and land-use data are provided by the WPS that use USGS 24-type land cover classification. For research area, there are only 8 land use categories. (3) Watershed Boundary data is from the National Hydrography Dataset Plus (NHDPlus). It was used to delineate watershed boundary using WRF-Hydro ARC-GIS tools. The river network and waterbody data are also from NHDPlus High Resolution and NHD, respectively. (4) Reservoirs' information, like broad-crested weir length, is from the water control manuals (US Army Corps of Engineers, 2014). (5) Measured discharge and gauge location were obtained from the USGS National Water Information System.

Other data required for the WRF-Hydro model simulations and its evaluations are the meteorological forcing and the soil moisture. The meteorological forcing data is from the North American Land Data Assimilation System version-2 (NLDAS2) hourly forcing data (Xia et al., 2012). Its spatial resolution is 0.125 degrees, which was interpolated into the WRF-hydro model

grid using the area conserve interpolation method. The in-situ daily soil moisture dataset is from the Soil Climate Analysis Network (SCAN) and Climate Reference Network (CRN) Soil Moisture data (Bell et al., 2013; Schaefer, Cosh, & Jackson, 2007; Schaefer & Paetzold, 2001). Surface layer soil moisture data were downloaded from the Soil Moisture Active and Passive (SMAP) L4 Global 3-hourly 9 km EASE-Grid Surface, and Root Zone Soil Moisture Geophysical Data V004 that has a 9-km grid resolution.

2.4 Model Description

The three main components of the WRF-Hydro model are the Noah Multi-parameterization land surface model (Noah-MP LSM), subsurface and terrain routing module, and gridded diffusive wave channel and reservoir routing modules (Gochis et al., 2018) (Figure 3). The WRF-Hydro model for the study area is configured using two meshes: the coarse mesh of 13667 m × 13667 m (1/8° latitude × 1/8° longitude) for the land surface model grid, and fine mesh of 1367 m × 1367 m (1/80° latitude × 1/80° longitude) for the terrain routing module. There are 36 × 36 coarse mesh grids (Figure 4), and 360 × 360 fine mesh grids in the study area. The model time step is 90 seconds to avoid numerical instability.

The Noah-MP LSM runs at coarse grid resolution, and it has four vertical soil layers with thicknesses of 0.1 m, 0.3 m, 0.6 m, and 1.0 m from surface to the bottom where it is connected to an unconfined aquifer. Total soil depth is 2m, and water table depth is dynamic that can be below or within the soil layers. The Noah-MP model uses a two-stream radiative transfer scheme, a vegetation canopy layer, a Ball-Berry type stomatal resistance scheme, a TOPMODEL-based runoff scheme, a three-layer snow model, a frozen soil scheme, and a simple groundwater model (G. Y. Niu et al., 2011). The incident energy at the surface is balanced by the net long wave radiation, latent and sensible heat fluxes, and the surface heating by iteratively solving the energy

balance equation for the ground temperature. The Noah-MP model uses majority vegetation type to differentiate between vegetated fraction and bare ground fraction in the grid cell using the 'semi-tile' approach (G. Y. Niu et al., 2011).

The Penman-Monteith equation is used to compute latent heat flux or evapotranspiration, which is connecting the link between water and energy balance equations (Cai, Yang, David, Niu, & Rodell, 2014; Kumar & Merwade, 2011). Water flow in four soil layers is modeled using the diffusive form of Richard's equation, which includes soil water potential driven diffusive flow, gravity driven flow, and the sink term due to ET and/or drainage from the bottom soil layer (Zheng et al., 2015). Thus, compared to the traditional hydrological models which only solve the water balance equations, e.g. Soil and Water Assessment Tool (SWAT), the National Water Model solves full sets of water and energy balance equations, hence provides better constrain on water predictions. Infiltration excess, ponded water depths, and soil moisture are disaggregated to a finer resolution for the subsurface and terrain routing module using two-way couplings.

The terrain and subsurface routing model operates at a finer spatial resolution. Fully saturated grid cells generate lateral subsurface flow, which is added to the infiltration excess from the LSM (Gochis et al., 2018). Next, the model determines the water table depths which are the depths to the top-most saturated layer and a key parameter in infiltration and saturation excess calculation (Guo‐Yue Niu et al., 2005). Baseflow contribution to the streams is modeled using the conceptual catchment storage-discharge bucket model (Gochis et al. 2018). The Bucket model calculates baseflow into streams using an empirically-derived model as a function of the water depths in the bucket or overflow when it exceeds maximum bucket depth. Once water is discharged into the streams, the soil moisture states are updated and aggregated back to the coarse resolution

for the LSM calculation in the next time step. Thus, the WRF-Hydro model enables two-way coupling between the LSM and the terrain and subsurface routing module.

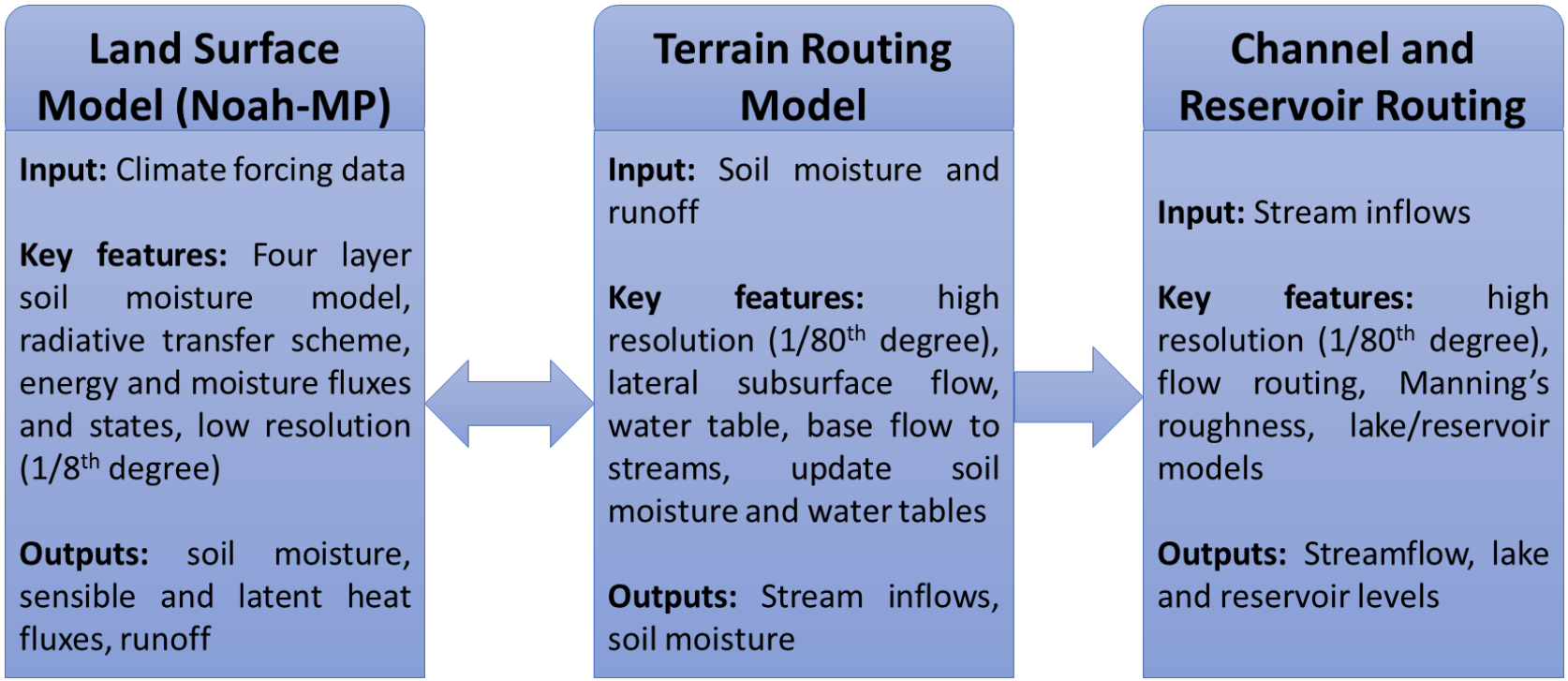


Figure 3: A text schematic of the WRF-Hydro model, and coupling among its components. Compared to the traditional land surface model, the WRF-Hydro model has added strengths in the high-resolution (10 times) terrain routing module and channel routine module. Hence, the WRF-Hydro model is more suitable for simulating daily streamflow in operational conditions.

There is one-way coupling between the terrain routing module and the channel and reservoir routing module, i.e. water cannot move from the channel to the land. The overland flow is routed through the channel using fully unsteady, spatially explicit, diffusive wave formulation and Manning's equation and Manning's roughness coefficients that are the functions of land cover types (Gochis et al., 2018). Water passing into/through lakes and reservoirs is routed using a simple level pool routing scheme. This final section generates streamflow that has been compared against USGS observations in this study. National Water Model takes fully physics-based approaches (solving water and energy balance equations) and also accounts for the hydrology and hydraulics characteristics of the terrain and the channel to generate streamflow forecast. It also using two-way couple to connect land surface model with terrain routing model.

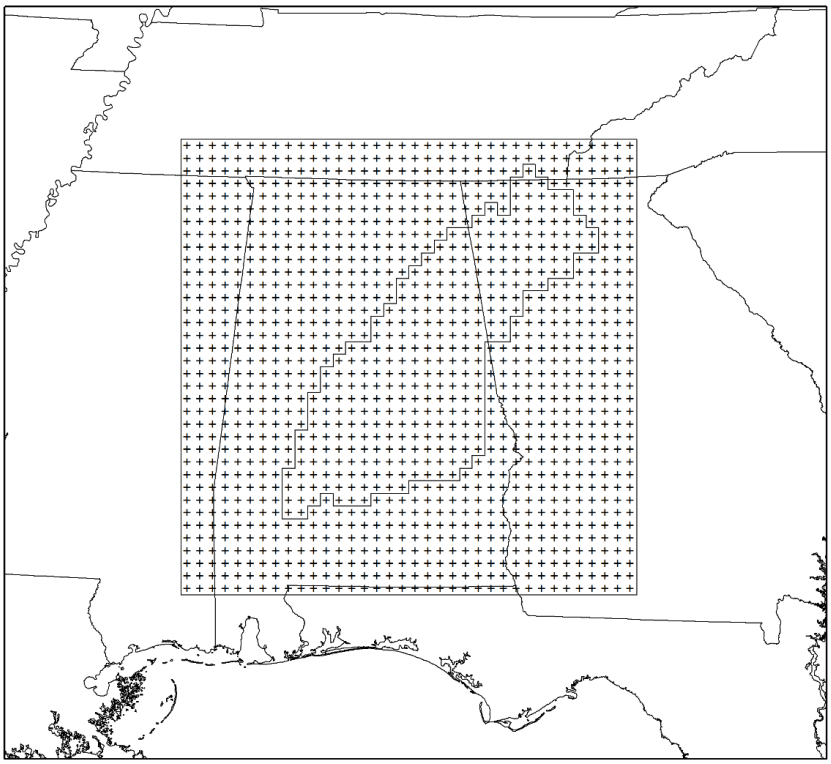


Figure 4: WRF-Hydro terrain routing module grid

2.5 Model Calibration

We performed manual calibration using sensitive parameter lists from the literature (Givati, Gochis, Rummler, & Kunstmann, 2016; Kerandi et al., 2018; Naabil, Lamptey, Arnault, Olufayo, & Kunstmann, 2017; Senatore et al., 2015; Yucel, Onen, Yilmaz, & Gochis, 2015). The baseflow and total flow volume are sensitive to saturated soil conductivity (SATDK). Text in the parenthesis refers to parameter name in WRF-Hydro model. We also adjusted the parameter of the groundwater bucket model to increase the baseflow. A final set of calibrated parameters are given in Table 2. The calibration period was from 1980 to 1990 (Figure 5), and validation period was from 1990 to 2000. The model performance was evaluated using the Nash Sutcliff Efficiency (NSE) coefficient (Nash & Sutcliffe, 1970).

The calibrated model improved the streamflow simulation performance, especially for the low-flow conditions that can be clearly seen in the flow exceedance probability curve shown in the bottom panel (Figure 5). Model performance was generally poor from 1980 to 1986 (calibrated monthly NSE: 0.48) and improved considerately for 1987 to 1990 (calibrated monthly NSE: 0.79) during calibration period, and also during validation period (monthly NSE: 0.69). A pooper performance in the earlier period can be related to more recent land use data (around 2013) used in the model (not investigated). The NSE for monthly streamflow improved from 0.64 for the uncalibrated model to 0.69 for the calibrated model during the validation period (Figure 6). Since NSE is more sensitive to high flows, the NSE improvement is small (Kumar & Merwade, 2009). The performance of the calibrated model for daily streamflow outputs is also satisfactory (daily NSE=0.55 for 1990 to 2016 period, (figure not shown). We performed the predictability experiments (described next) using the calibrated model and during 1990 to 2016.

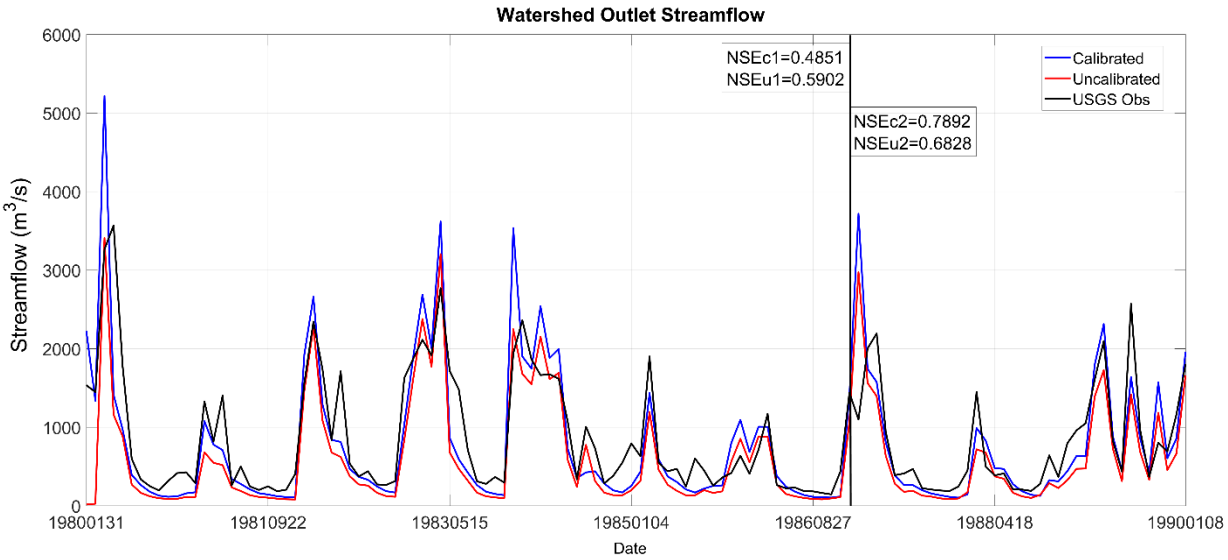
Figure 5 and 6 are model calibration results after model set up. Appendix 1 is model set up method. Figure 7 is calibration results for each sub watershed. The calibration is warm start. Although, we calibrated model only at the watershed outlets. The NSE is generally satisfactory (>0.4) for 11 out of 15 sub-watersheds (Figure 7).

Table 2: Calibrated parameters. Parameter files are HYDRO.TBL and SOILPARM.TBL. First table SATDK is saturated soil hydraulic conductivity [m/s]. Parentheses values are default parameters. Second table is Channel routing parameter tables. Parameter file is CHANPARM.TBL. Bw is channel bottom width [m]. HLINK is Initial depth of water in the channel [m]. ChSSlp is Channel side slope [m/m]. MannN is Manning's roughness coefficient (Chow, 1959). Bucket parameter Coefficient is three. Its default value is one.

16-category Soil Categories	SATDK(m/sec)
Sand	9.32E-5(4.66E-5)
Loamy Sand	2.82E-5(1.41E-5)
Sandy Loam	1.05E-5(5.23E-6)
Silt Loam	5.62E-6(2.81E-6)
Loam	6.76E-6(3.38E-6)
Sandy Clay Loam	8.90E-6(4.45E-6)
Silty Clay Loam	4.06E-6(2.03E-6)
Clay Loam	4.90E-6(2.45E-6)
Silty Clay	2.68E-06(1.34E-6)
Clay	1.95E-06(9.74E-7)

Stream Order	Bw	HLINK	ChSSlp	MannN
1	3(1.5)	0.5(0.02)	0.1(3.00)	0.03(0.55)
2	3(3.0)	0.5(0.02)	0.09(1.00)	0.03(0.35)
3	20(5.0)	0.5(0.02)	0.07(0.50)	0.03(0.15)
4	40(10.0)	0.5(0.02)	0.02(0.18)	0.03(0.10)
5	100(20.0)	0.5(0.02)	0.02(0.05)	0.03(0.07)

(a) Watershed outlet monthly streamflow



(b) Daily flow exceedance probability

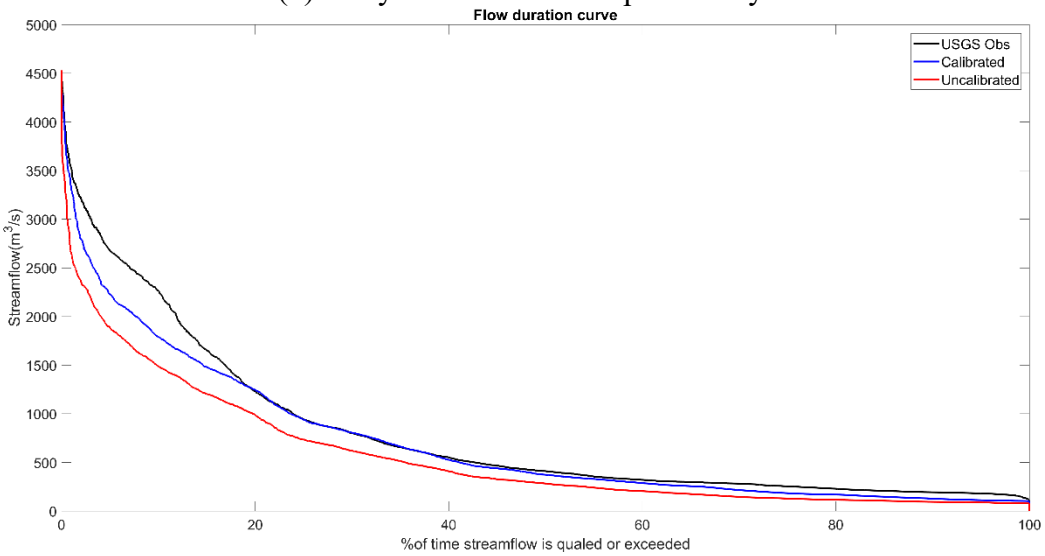
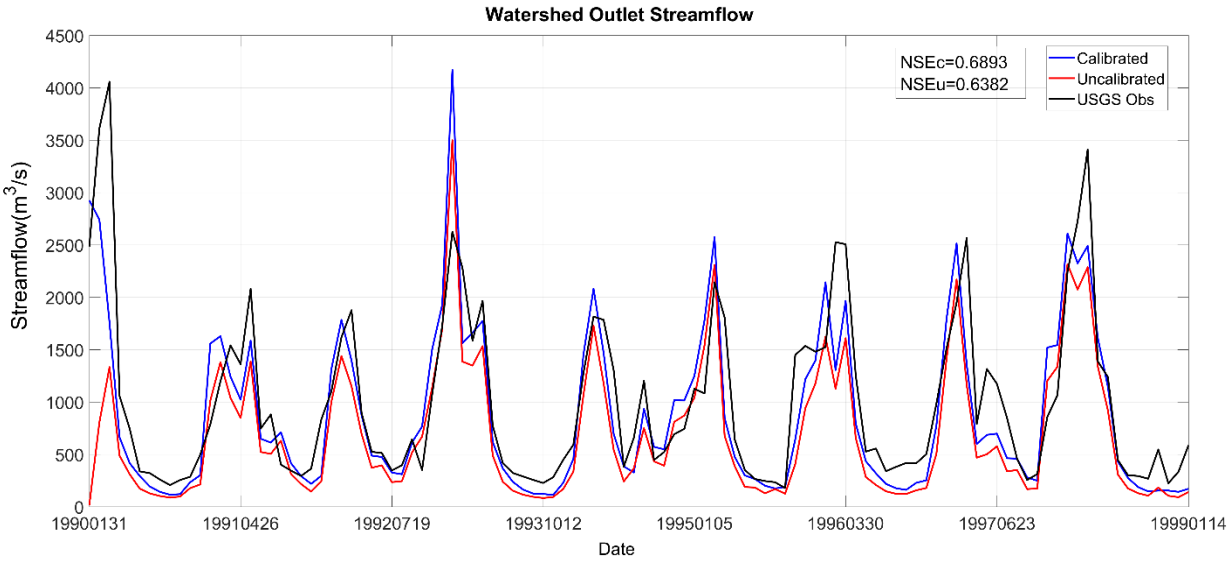


Figure 5: The NWM model performance during the calibration period (1980-1989). The NSE1 period is 19800602-19870124. The NSE2 period is 19870125-19900108. Model performance was generally poor from 1980 to 1986 (calibrated monthly NSE: 0.48) and improved considerately for 1987 to 1990 (calibrated monthly NSE: 0.79). A poor performance in the earlier period (1980 to 1986) may be related to more recent land use data (2011) or issue with the NLADS2 climate forcing (Kumar, Dirmeyer, & Kinter, 2014) (not investigated).

(a) Watershed outlet monthly streamflow



(b) Daily flow exceedance probability

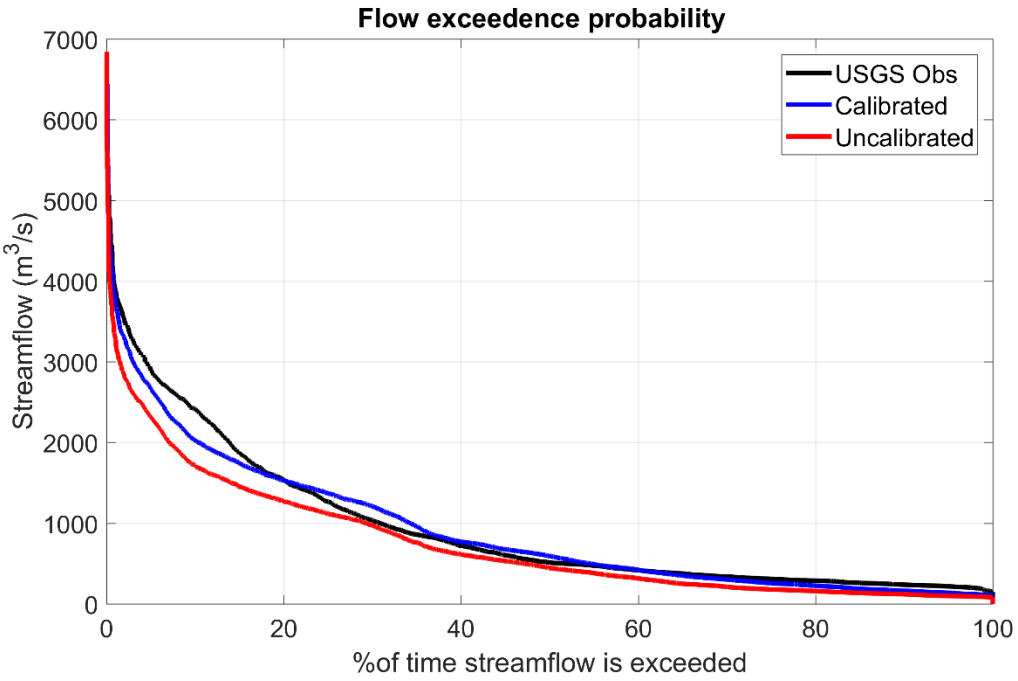


Figure 6: The same as Figure 5 but for validation period.

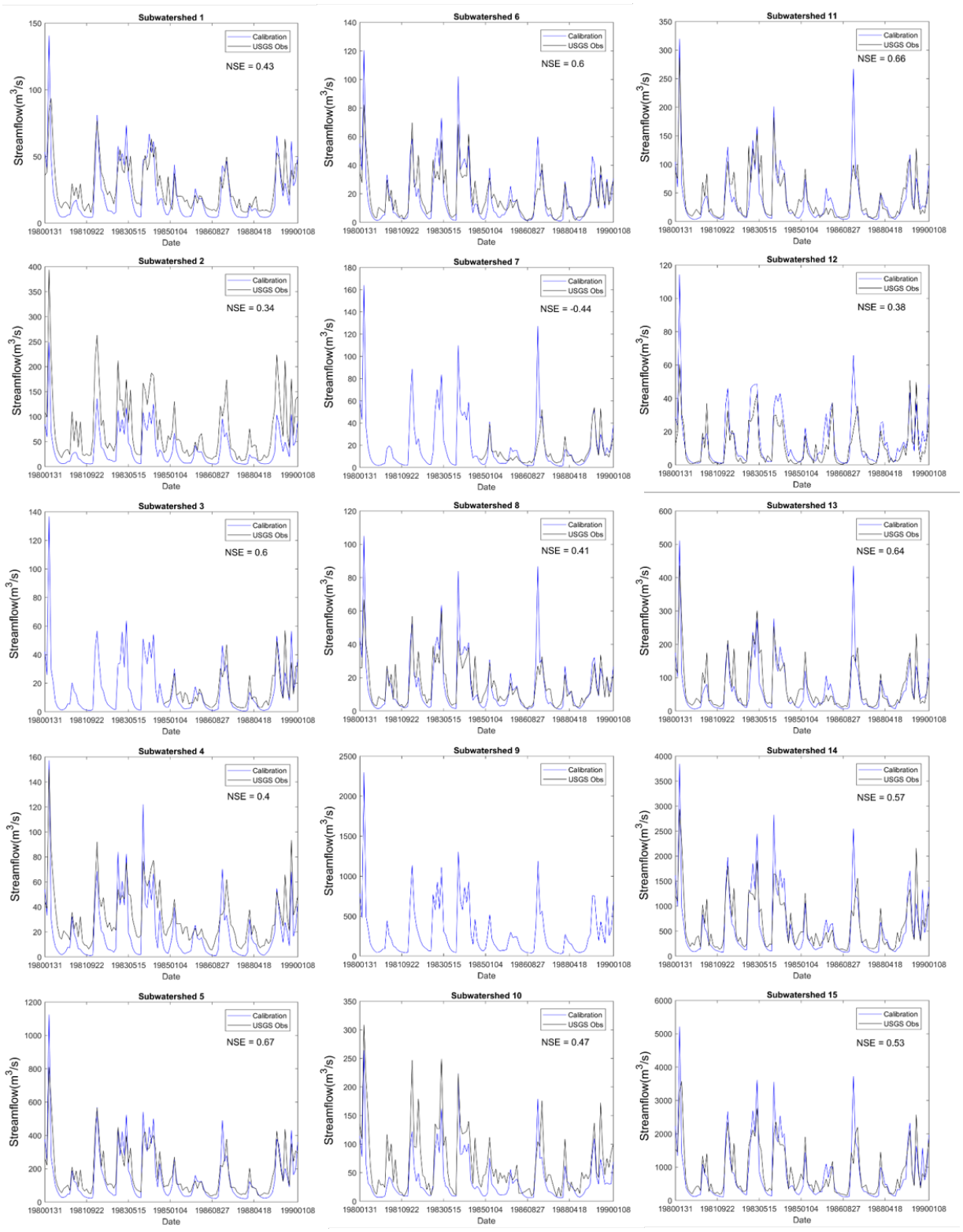


Figure 7: Calibrated monthly average streamflow for 15 subwatersheds. Observation is not available for Subwatershed 9.

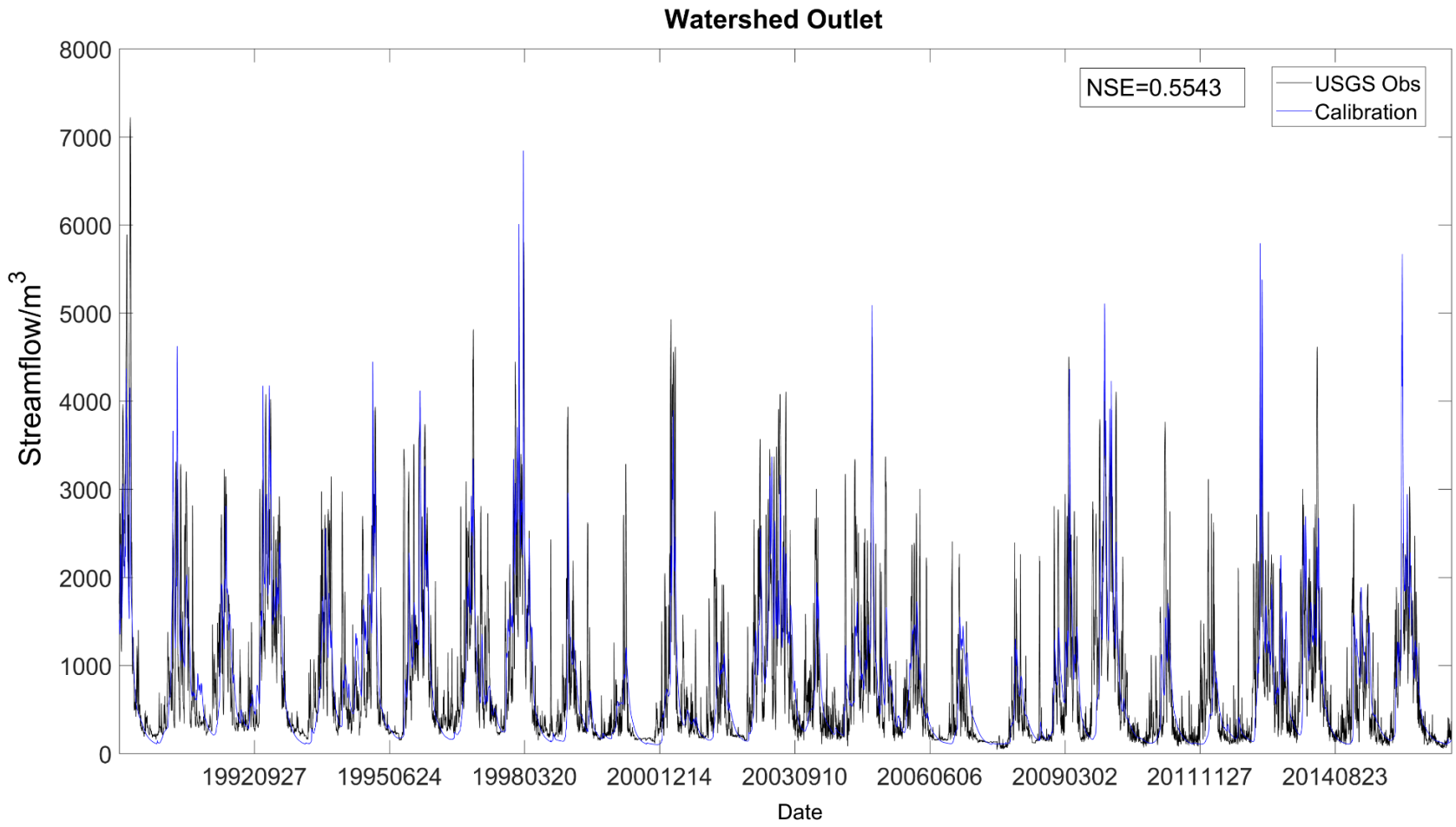


Figure 8: Daily streamflow from 1990.01 to 2016.12. Blue line is calibrated streamflow, black line is observed streamflow.

CHAPTER 3 RESEARCH METHOD AND STATISTICAL METHODS

3.1 Research Method

This study investigates effects of the realistic initial soil moisture state (ISMS) on potential predictability of the soil moisture and streamflow. We had followed the seasonal climate predictability experiment design suggested by P. A. Dirmeyer et al (2013). Here, we briefly describe the experiment design: a long control simulation from year's 1990 to 2016 treated as the synthetic observations or control run (27-yr CTRL) for the forecast evaluations. To set aside the forecast biases due to model structure, the model experiment results are compared with control run but not observation. So the project conclusions are the ideal predictabilities. The real world case, for example, soil moisture data assimilation improved SWAT inherent vertical coupling of soil layers, but its improvement for streamflow simulation is limited (Patil, 2017). In the following discussion, the realistic initial soil moisture state means model generated synthetic soil moisture, it is not observed.

The model is spun-up by repeating the simulations from year 1980 to 1989 three times and using the two restart files generated at the end of the simulation in the subsequent run, thereby equilibrating the model's dynamical state as the warm start for the 27-yr CTRL simulation (Rodell, Houser, Berg, & Famiglietti, 2005). The two restart files are HYDRO_STR and RESTART files. They include all states needed to launch the model. All experiments were driven by the synchronous and random combination of reanalysis climate forcing (NLDAS2) and were initialized by the ensemble of realistic and random ISMS. A noteworthy difference from the Dirmeyer et al. (2013) study is that climate forcing does not respond to soil moisture states in the model. Our experiment design provides an estimate of the potential streamflow forecast skill under realistic and random soil moisture initial condition and climate forcing (Wood et al., 2016).

We selected 15 years for seasonal streamflow forecast experiments: 5 El Niño years (1991, 1992, 1997, 1998, 2015), 5 La Niña years (2000, 2007, 2008, 2010, 2012), and 5 normal years (1993, 1996, 2004, 2006, 2014). The seasonal streamflow forecast experiments were launched in each of the 15 selected years on January 1st, April 1st, July 1st, and October 1st, and the forecast simulations run for 91 days. Thereby, we assessed the streamflow forecast skills in four seasons: winter (JFM), spring (AMJ), summer (JAS), and fall (OND) seasons.

On each forecast experiment start date, for example April 1st, 2006 (Figure 9), we generated 2 sets each, having 14-members ensemble the streamflow forecast: the first set with realistic ISMS, and the second set with random ISMS. Following Dirmeyer et al. (2013), we selected 14 realistic ISMS from the last seven days by saving the HRLDAS and hydro restart file every 12 hours. Table 3 shows the list of the states we edited for the forecast experiment. Except the listed variables, other initial states in restart files remain the same.

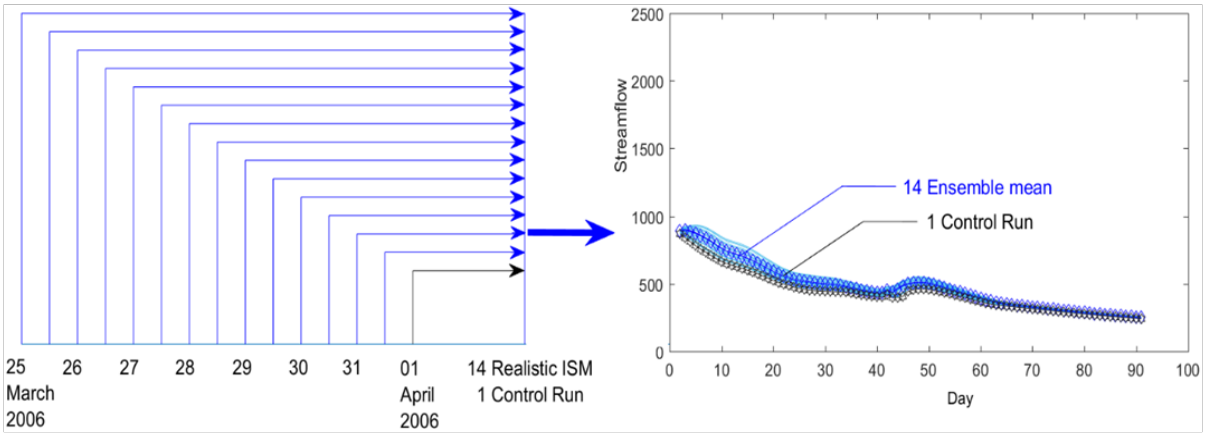
The methodology to generate the ensemble is as follows: extract all soil moisture variables in HRLDAS and hydro restart files, write them into the restart file for the given start date, and thereby generate the 14-member ensemble streamflow forecast corresponding to each restart file but using the same climate forcing (Experiment ID: real ISMS + obs. CF in Table 4). Due to the past seven days of ISMS, the 14-member ensemble soil moisture states are slightly different compared with each other. For the second set of the experiment, the random soil moisture states were selected from the same date but from the remaining 14 years (not including the year for which forecast is generated, i.e. 2006 in Figure 9). The reaming procedure is the same as the first set of the experiment (Experiment ID: random ISMS + obs. CF). The complete set is repeated for each of the 15 selected years and for 4 seasons in each year. Thus, we had generated 2 sets × 14 members

× 15 years × 4 seasons = 1680 seasonal streamflow forecast simulations of 91 days each in the ACT basin. We tested our realistic and random ISMS ensemble methodology by comparing them with the forecast error on the first day of the forecast outputs. We found that the random initialization has significantly higher RMSD than the realistic initialization method (Figure 10). Hence, we concluded that our ensemble methodology works as designed.

Table 4 lists all the experiments and their significance. To separate the contribution of the ISMS and the climate forcing, we re-arranged the forecast ensemble so that it receives realistic initial conditions and the random climate forcing from each of the 14 different years except the year for which forecast is issued (Experiment ID: real ISMS + random CF). Our methodology is similar to the ESP (real ISMS + random CF) and Rev-ESP (random ISMS + obs. CF) methodology (H. Li et al., 2009; Wood et al., 2016). The difference is that instead of testing the two methods separately, we combined the two methods together to assess forecast skill due to initial conditions and climate forcing effects. Additionally, we investigated the potential predictability skill that can be derived using the best available climate forecast (real ISMS + obs. CF). Since, the actual seasonal climate forecast skill would be in between the best forecast (obs. CF) and the random climate forcing. Hence, we covered the range of the forecast skill that can be potentially derived using the operational forecast system (Wood et al., 2016).

We had also differentiated the soil moisture contributions from the ground water contributions by repeating our predictability experiments for the groundwater parameters only (Table 3).

(a) Realistic ISMS experiments



(b) Random ISMS experiments

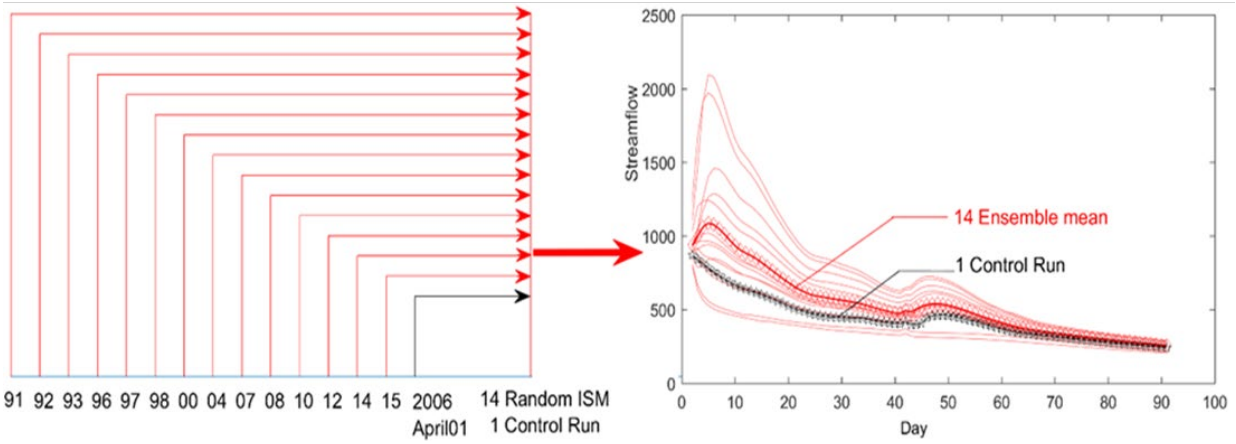


Figure 9: The seasonal streamflow forecast experiment design: the initialization process for the ensemble simulations with (a) realistic ISMS taken from the last seven days (b) randomized land conditions – same date but from different years. Left figure x-axis labels are the last two digital years. Included 2006, they represent five El Niño, five La Niña years, and five normal conditions. April 01 2006 is an example. Control runs (CTRL) in (a) and (b) are the same.

Table 3: List of perturbed soil moisture variables in the NWM for seasonal streamflow forecast experiments

LSM restart variables	**Abbreviation**	**Note**	**Unit**
volumetric soil moisture	SMC	4 layers soil moisture from 0.1 -2.0 m	m3/m3
volumetric liquid soil moisture	SH2O	SH2O=SMC-SICE (soil ice)	m3/m3
volumetric soil moisture	SMOISEQ	equilibrium soil water content	m3/m3
soil moisture below the bottom of the column (water table depth)	SMCWTDXY*	below 4 layers (2 meters)	m3/m3
soil moisture below the bottom of the column (deep reach)	DEEPRECHXY*	below 4 layers (2 meters)	m3/m3
water in aquifer relative to reference level	WA*	below 4 layers (2 meters)	kg/m2 (mm)
water in aquifer and saturated soil	WT*	below 4 layers (2 meters)	kg/m2 (mm)
water table depth	ZWT*	below 4 layers (2 meters)	m3/m3

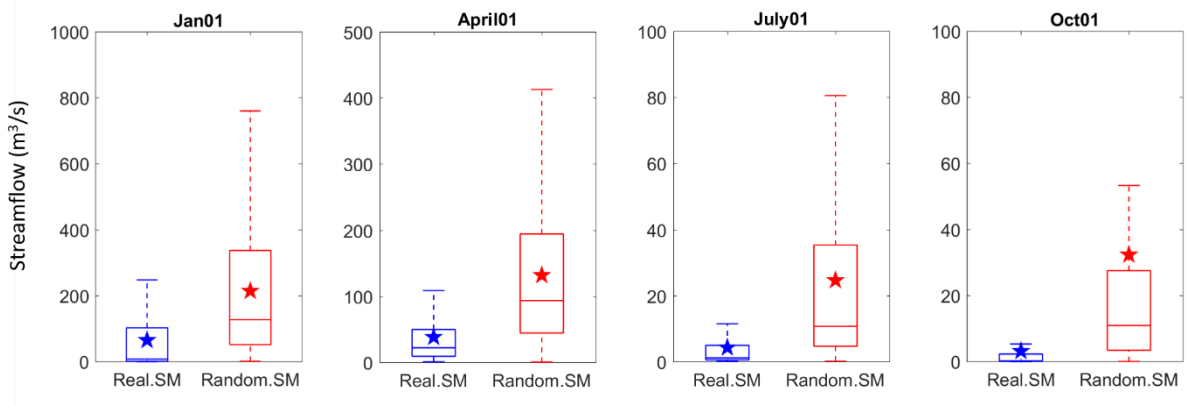
HYDRO restart variables	**Abbreviation**	**Note**	**Unit**
total liq+ice soil moisture	smc	4 layers	m3/m3
liquid soil moisture	sh2ox	4 layers	m3/m3
depth in ground water bucket	z_gwsubbas*	bucket model water depth	m

*Abbreviations marked with star are groundwater initial states. Abbreviations correspond to WRF-Hydro Fortran scripts.

Table 4: List of seasonal streamflow forecast experiments

Exp. No.	Experiment ID	ISMS	Climate Forcing (CF)	Remarks
E1	real ISMS + obs. CF	Realistic	Observed	Most predictable scenario
E2	rand. ISMS + obs. CF	Random	Observed	Effects of ISMS under observed climate forcing (E1-E2)
E3	real ISMS + rand. CF	Realistic	Random	Effect of CF only (E1-E3) Effects of ISMS only (E3-E4)
E4	rand. ISMS + rand. CF	Random	Random	Noise

(a) Streamflow initial error size between the realistic and random ISMS initializations



(b) Soil moisture initial error size between the realistic and random ISMS initializations

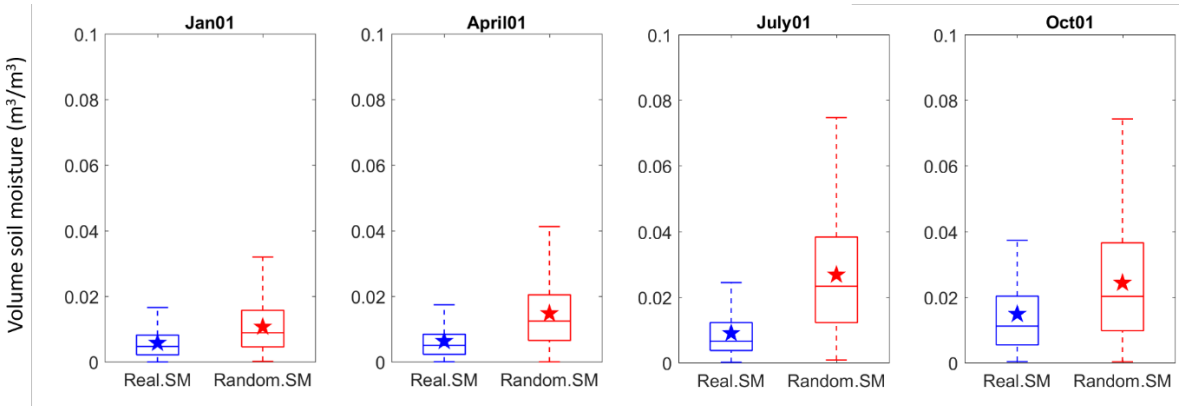


Figure 10: Effects of realistic and random ISMS initialization on streamflow (top row) and soil moisture (bottom row) forecast RMSD on the first day of the forecast period. (a) streamflow initial states (15 years*14 ensemble members) distribution. (b) They are soil moisture initial states (15 years *14 ensemble members) distribution. When the two distributions are significant in t-test with 95% confidence level, their average values are marked with solid stars. The realistic ISMS initialization was generated from the last 7 days before forecast initial date. The 14 initial states are soil moisture with 12 hours intervals in the 7 days. So, realistic ISMS initializations provide reasonable spread in perturbations to generate ensemble forecast experiments.

3.2 Statistical Analysis

We assessed the improvement in the forecast due to realistic ISMS by comparing the root mean square difference (RMSD) (Eq. 1) between realistic and random ISMS experiments:

$$RMSD_i = \sqrt{\frac{1}{N*E}\left[\sum_{y=1}^{N}\sum_{e=1}^{14}\left(f_{y,e,i} - o_{y,i}\right)^2\right]} \qquad (1)$$

Where $f_{y,e,i}$ is the e[th] member seasonal streamflow or soil moisture forecast in the year *y* for the given season and i^{th} day forecast period, $o_{y,i}$ is the corresponding synthetic control run in year *y* and i^{th} day, so $o_{y,i}$ is same for all 14 member ensemble forecast. N is 15 for all years, and N=5 for El-Nino only, La-Nina only, or Normal only years. e is 14 ensemble members. The 95% uncertainty range is computed as two times the standard error of the RMSD across ensemble members and years. Two forecasts (realistic and random ISMS) are statistically indistinguishable if the 95% uncertainty range overlaps.

To assess the improvements over the climatology forecast, we computed the Nash–Sutcliffe efficiency (NSE) (Eq. 2) of the ensemble average forecast in the given season and year, as defined below for the first 10 days of the forecast:

$$NSE = 1 - \frac{\sum_{i=2}^{11}\left(o_{y,i} - \frac{1}{14}\sum_{e=1}^{14} f_{y,e,i}\right)^2}{\sum_{i=2}^{11}\left(o_{y,i} - \bar{o}\right)^2} \qquad (2)$$

where $\bar{o} = \frac{1}{27*10}\sum_{y=1}^{27}\sum_{i=2}^{11} o_{y,i}$ is the climatology forecast for the given forecast period. Here year y is 27 years from 1990-2016, i is the day from the 2[nd] days to the 11[th] days for the first 10 days forecast. $\frac{1}{14}\sum_{e=1}^{14} f_{y,e,i}$ is year y, ensemble average for i[th] day. The year y is the 15 selected years. $o_{y,i}$ is the corresponding y[th] year, i[th] days control run. Similarly, we computed NSE for every 9 non-overlapping forecast period, each having 10 days (2-11, 12-21, 22-31, 32-41, 42-51, 52-61, 62-71, 72-81, and 82-91 days). The NSE can range from -∞ to 1, where values greater than 0 mean

useful forecast skill (better than climatology). We compared NSE score between two experiments (realistic and random ISMS) using student t-test and 15 years of data for each forecast period. We assigned any negative NSE values to zero for the t-test calculations.

Finally, we quantifid the contribution of ISMS, climate forcing, and ground water using signal to noise ratio metric (Guo, Dirmeyer, & DelSole, 2011).

$$SNR = \frac{V_s}{V_n} = \frac{\frac{1}{N}\sum_n (f_{\cdot n} - f_{\cdot\cdot})^2}{\frac{1}{NE}\sum_n \sum_e (f_{en} - f_{\cdot n})^2} \qquad (3)$$

where $f_{\cdot\cdot} = \frac{1}{NE}\sum_{e=1}^{E}\sum_{n=1}^{N} f_{e,n}$ and $f_{\cdot n} = \frac{1}{E}\sum_{e=1}^{E} f_{e,n}$. V_s represents variability of ensemble mean, or the signal, and V_n is variability about the ensemble mean or the noise term. The null hypothesis of no predictability can be rejected at 95% level if $SNR \geq \mathrm{F}^{0.05}_{\mathrm{N-1,N(E-1)}} * \frac{N-1}{N(E-1)}$. Where $\mathrm{F}^{0.05}_{\mathrm{N-1,N(E-1)}}$ is upper 5% threshold for F-distribution with N and N (E-1) degrees of freedom. Signal to total ratio is defined by $STR = \frac{SNR}{SNR+1}$ and varies between 0 to 1.

Chapter 4 RESULTS

We had first compared the soil moisture in the CTRL experiment from 2013 to 2018 with the SCAN and SMAP data, because SMAP data is only available from 2015 to present. All initial soil moisture states, groundwater initial states, and climate forcing effect on streamflow and soil moisture predictability were evaluated in the 15 selected years, consisting of 5 El-Niño, 5 La-Niña, and 5 normal years. The soil moisture states determine the saturated land fraction, and the water table position, and these variables in-turn give surface and sub-surface runoff for the streamflow routing (Cai et al., 2014; Zheng et al., 2015). The soil water diffusivity and hydraulic conductivity are also dependent on the soil moisture states in the model (Campbell, 1974); and hence the streamflow generation and soil moisture states are interlinked in the model and the real world.

4.1 Comparison with the SMAP and SCAN/CRN soil moisture data

By comparing SMAP data with the NWM control experiment, we found that NWM simulates soil moisture climatology (mean) well, but it underestimates the soil moisture variability. Figure 11 shows average and standard deviations of the SMAP 5-cm soil moisture from April 2015 to December 2018. Since the data availability is small, we averaged the data for the entire period. The agricultural sub-watersheds located towards the south show a higher soil moisture than forested sub-watersheds that are generally located in the north within the research area. The NWM generally captures large-scale spatial variations in the study domain. It is worth noting the drier soil moisture condition in the southeast corner of the study domain in both NWM and the SMAP (Figure 11). The NWM considerably underestimates the soil moisture variability in the study domain. The soil moisture daily variability ranges from 0.02 to 0.04 in the NWM compared with the SMAP's range: 0.06 to 0.08 m^3/m^3.

We also compared NWM soil moisture data with in-situ observations and found that the NWM overestimated soil moisture in dry seasons. Figure 12(a) shows a comparison of the NWM and SMAP with the in-situ observations at seven sites from the SCAN and CRN networks. Data are interpolated to soil moisture gage and the concurrent period are used for the analysis. The interpolation method is inverse distance weighted interpolation algorithm. The NWM shows a comparable soil moisture with the observations during winter and spring seasons but the NWM overestimates the soil moisture in the summer and fall seasons. In addition, the NWM underestimates the soil moisture variability generally for all seasons.

The SMAP is wetter than SCAN/CRN soil moisture in winter and spring seasons. The positive bias is generally reduced during the summer and fall seasons. The SMAP soil moisture variability is generally higher than that of the NWM. A comparison at the individual sites shows that SMAP performs better than the NWM data, except gage 5 and gage 6 (Table 5). Figure 14 and Figure 15 are figures for individual gage comparison. Figure 14 corresponds with Table 5, Figure 15 time period is different with Table 5. Compare Figure 14 and Figure 15, both SMAP and NWM overestimate soil moisture for station 3, 4, 6, 7. SMAP are generally better than NWM as discussed above.

Figure 12(b) shows a comparison of the seasonal cycle of basin average soil moisture between SMAP and the NWM. The lack of soil moisture variability in the NWM mainly stems from considerably smaller inter-seasonal soil moisture variability. The NWM shows a generally drier soil moisture condition in the winter and spring and a wetter condition in the summer and fall, compared with the SMAP soil moisture seasonal cycle. A weak seasonal cycle in the NWM is concerning here; nevertheless, we continued with our potential predictability experiments. Investigation of the root cause of the weak seasonal cycle is outside the scope of this study.

The NWM simulates soil moisture variability well, but much greater than SMERGE climatology (mean). Figure 16 is the soil moisture climatology and variability in the NWM and its comparison with the SMERGE data. They are 40-cm layer daily soil moisture from 1990-2015. The NWM soil moisture distribution is similar with SMERGE, the moisture is almost double. The NWM generally captures large-scale spatial variations in the study domain. The NWM considerably overestimates the soil moisture variability in the northeast part. For monthly average, Figure 12 (c) is NWM and SMERGE soil moisture basin average comparison from 1990-2015. NWM is lack of variability, and has greater standard deviation during dry season.

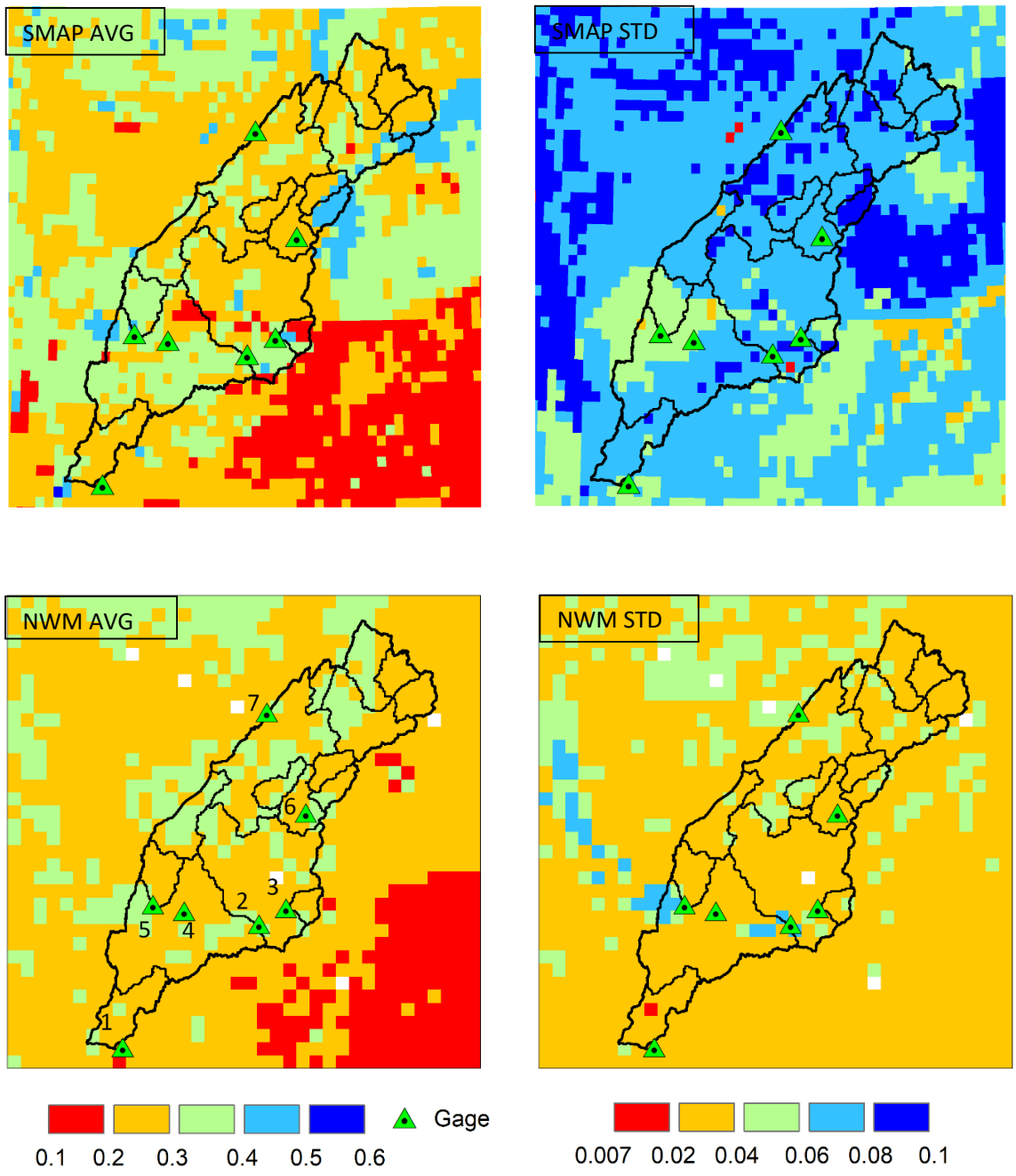


Figure 11: The soil moisture climatology and variability in the NWM and its comparison with the SMAP observations. Top row shows daily soil moisture average and standard deviation for March 2015 to December 2018 period from SMAP data at its native resolution (9km). Bottom row show the same for NWM model at is original resolution (13.7 km). The blank pixel is water body. Please note that SMAP data is available from March 2015 onwards, only. Triangles show ground truth observation sites from SCAN and CRN networks. Units: m^3/m^3. SMAP is 5-cm layer soil moisture and NWM is 10-cm layer soil moisture.

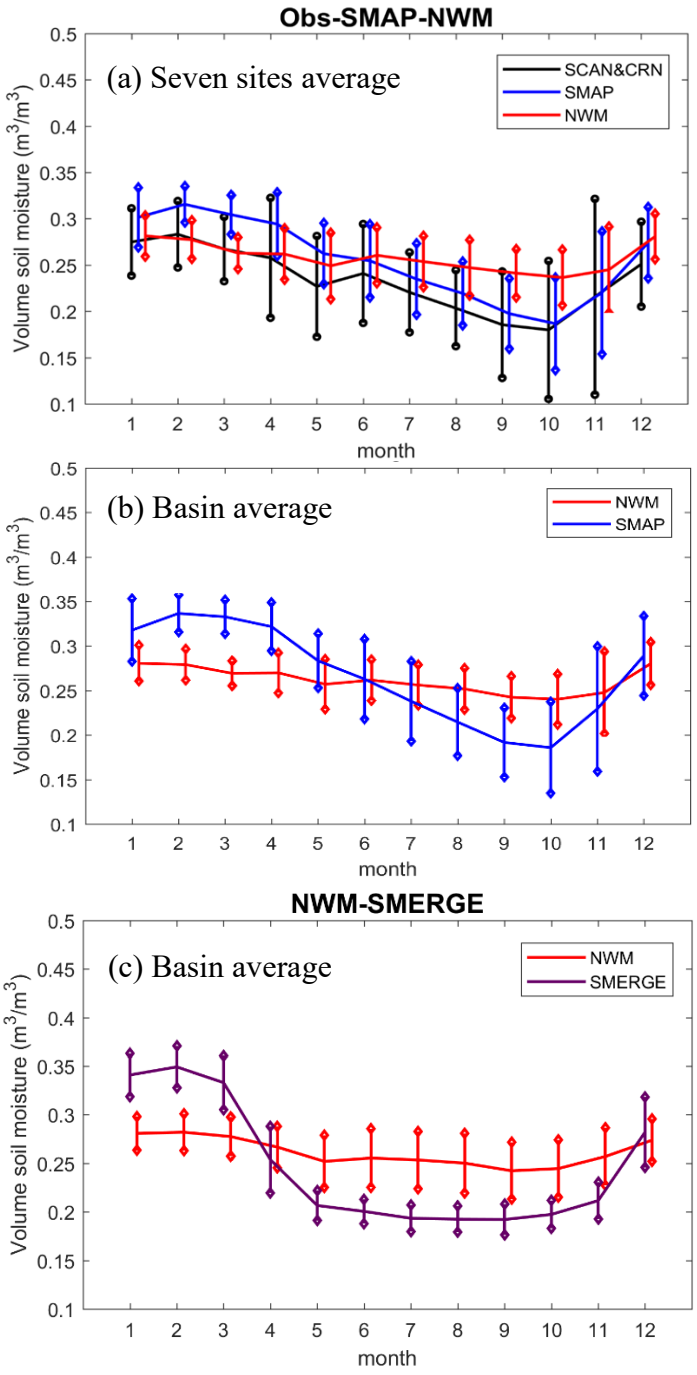


Figure 12: Comparison of the SMAP and the NWM soil moisture data with ground-truth observations from SCAN and CRN Network. (a) Seven site average monthly soil moisture and standard deviation from April 2015 to December 2018, and (b) basin average monthly climatology comparing SMAP and the NWM for the same period. The x-axis represent months January to December and y-axis represents soil moisture (m^3/ m^3). Error bar represents ± 1 standard deviation calculated from daily values in the corresponding month. Solid line is monthly mean. SMAP is 5-cm layer soil moisture and NWM is 10-cm layer soil moisture. Figure (a) SCAN/CRN is 5-cm layer soil moisture. (c) NWM and SMERGE soil moisture comparison from 1990-2015.

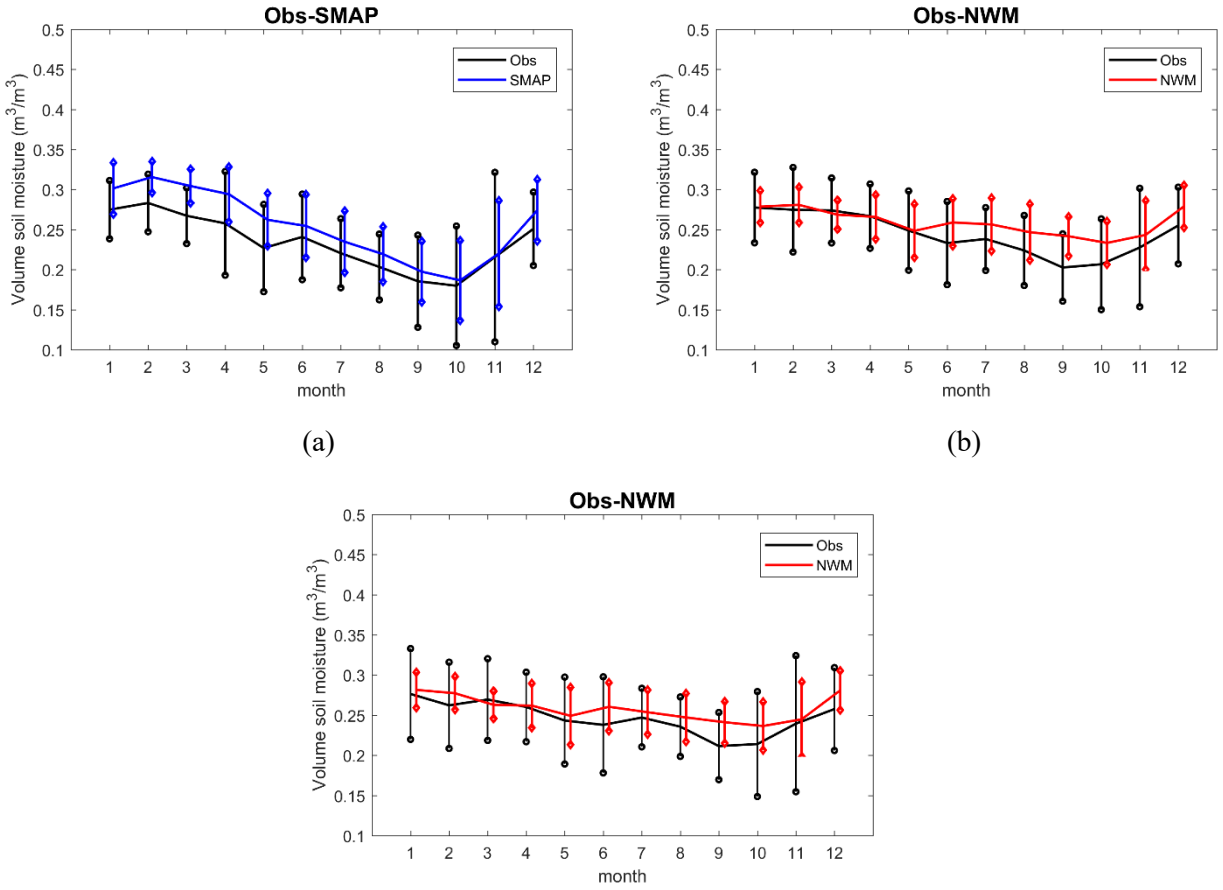


Figure 13. Similar with Figure 12 (a) Period is April.2015 - Dec.2018. Soil layer is 5-cm. (b) Period is Jan.2013 – Dec.2018. Soil layer is 10-cm. (c) Observation and NWM are in the same period (April 2015 – December 2018). Soil layer is 10-cm.

Table 5: Unbiased RMSE (UbRMSE) (Martens et al., 2017) and sample time series correlation in each soil moisture gage. Obs-SMAP and Obs-NWM period is April 2015 - Dec.2018. Obs-SMAP is 5cm soil moisture, Obs-NWM is 10cm soil moisture. Italic and bold represents significant test with 0.05 level.

UbRMSE	1	2	3	4	5	6	7
Obs-SMAP	0.05	0.11	0.05	0.04	0.11	0.07	0.04
Obs-NWM	0.07	0.09	0.08	0.04	0.06	0.06	0.05

R	1	2	3	4	5	6	7
Obs-SMAP	***0.77***	***0.79***	***0.61***	***0.66***	***0.73***	***0.34***	***0.82***
Obs-NWM	***0.38***	0.50	***0.33***	***0.45***	***0.50***	***-0.08***	***0.73***

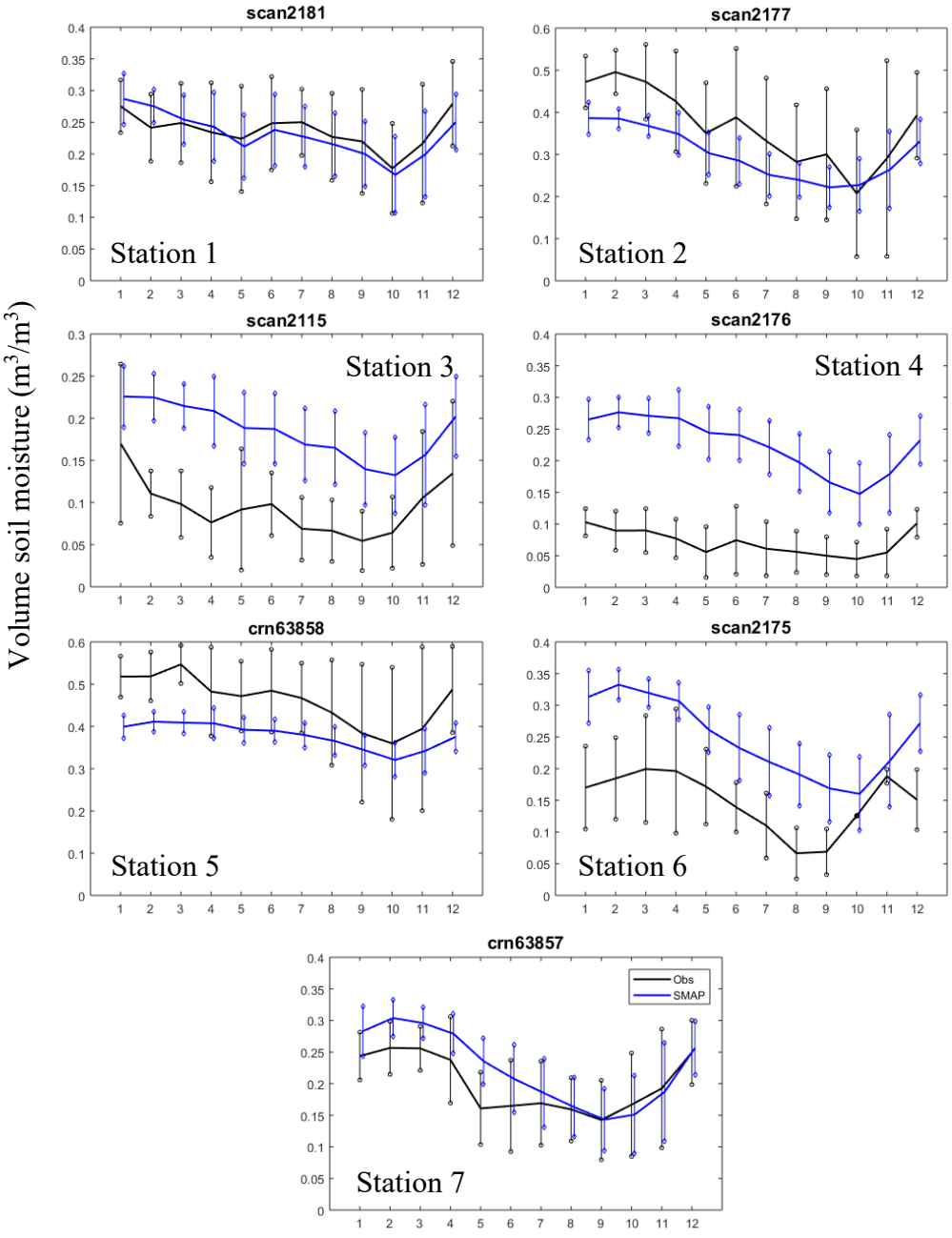


Figure 14: Comparison of the SMAP and the ground-truth observations from SCAN and CRN Network. SMAP soil moisture are interpolated into soil moisture gage. Seven site average monthly soil moisture and standard deviation from April 2015 to December 2018. The x-axis represent months January to December and y-axis represents soil moisture (m^3/ m^3). Error bar represents ± 1 standard deviation calculated from daily values in the corresponding month. Solid line is monthly mean, both the SMAP and observation are 5-cm soil moisture.

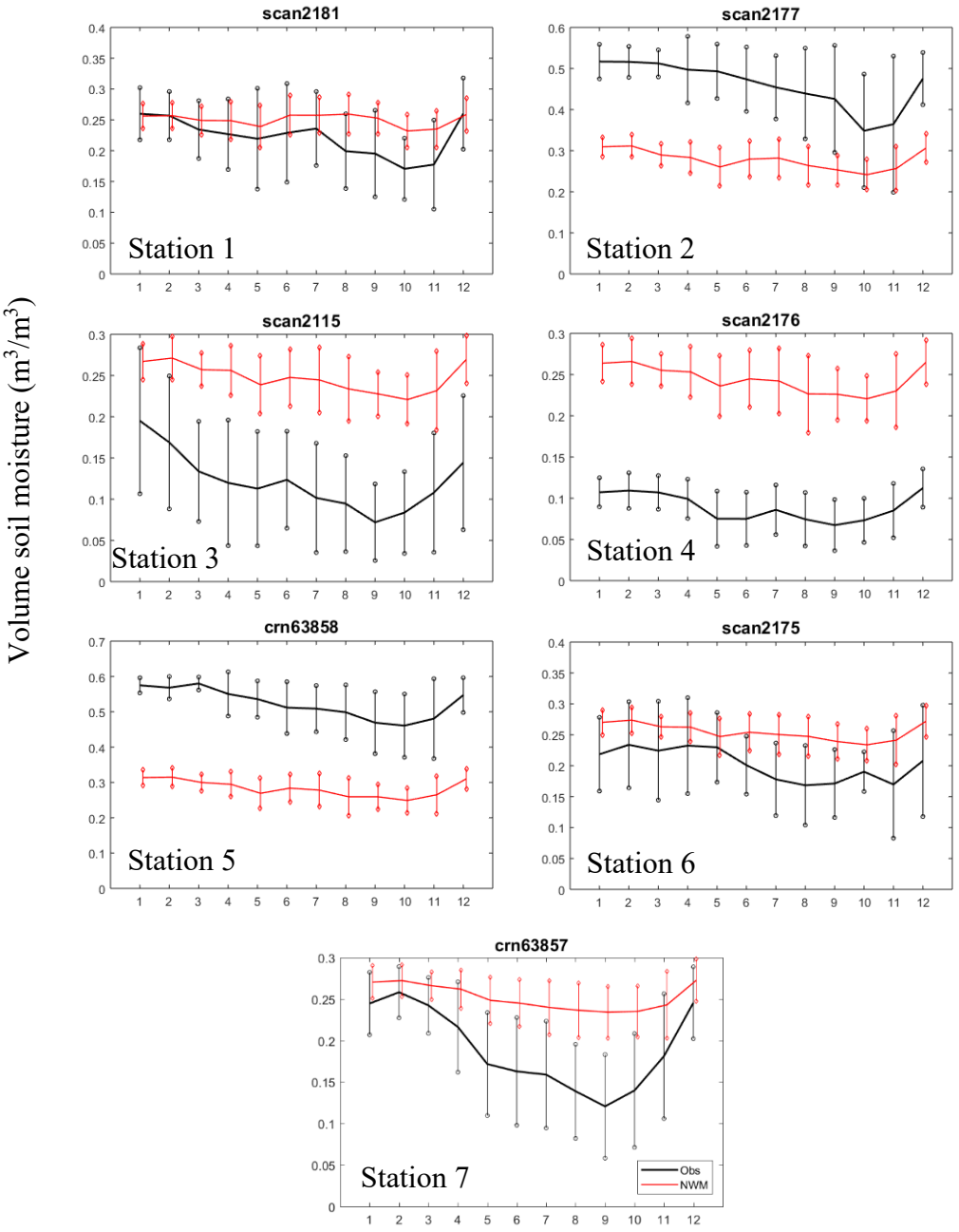


Figure 15: Comparison of the NWM and the ground-truth observations from SCAN and CRN Network. NWM soil moisture are interpolated into soil moisture gage. Seven site average monthly soil moisture and standard deviation from April 2015 to December 2018. The x-axis represent months January to December and y-axis represents soil moisture (m^3/ m^3). Error bar represents ± 1 standard deviation calculated from daily values in the corresponding month. Solid line is monthly mean, both the NWM and observation are 10-cm soil moisture.

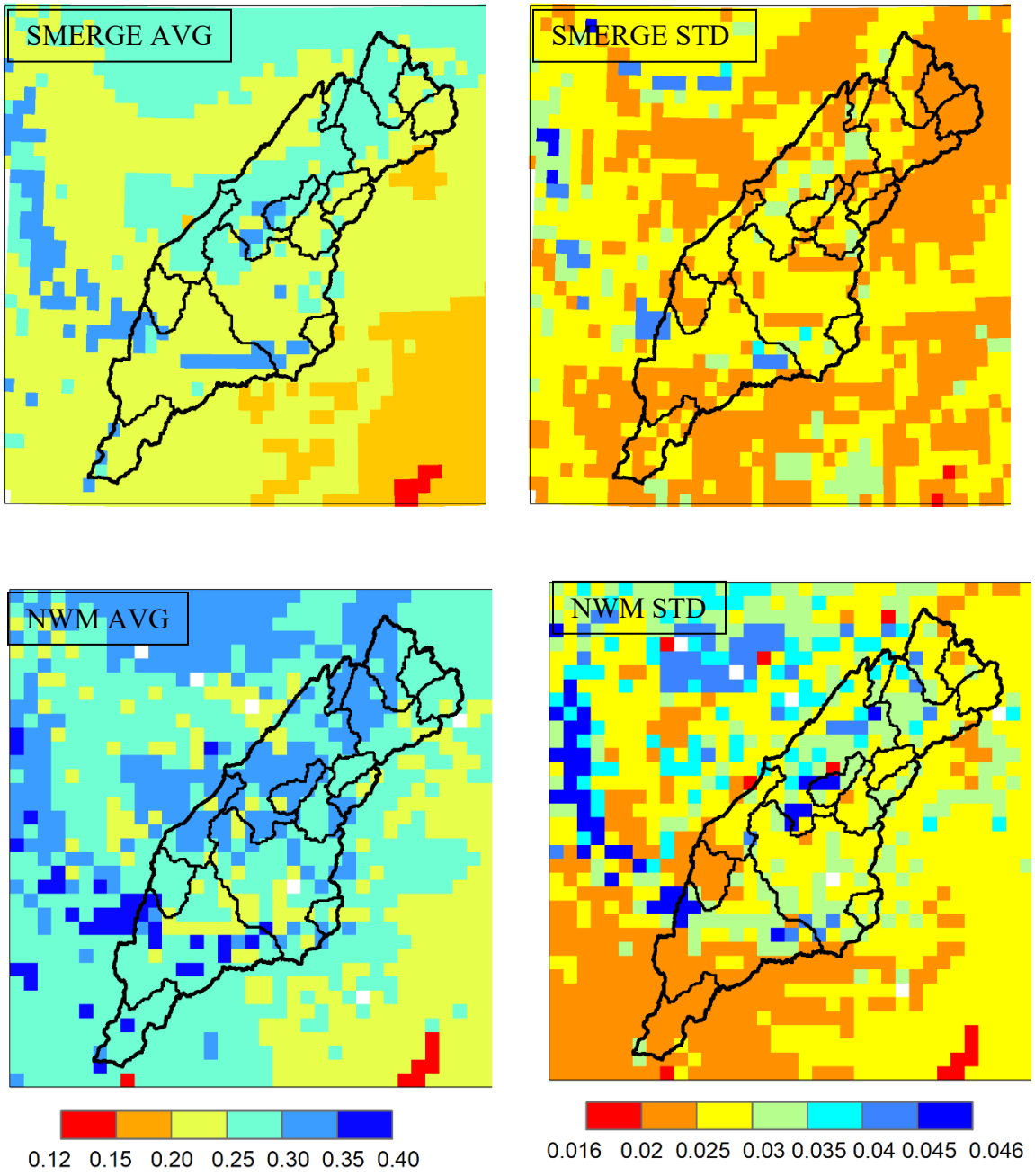


Figure 16: The soil moisture climatology and variability in the NWM and its comparison with the SMERGE observations. Top row shows daily soil moisture average and standard deviation for Jan 1990 to December 2015 period from SMERGE data at 45/58°. Bottom row show the same for NWM model at is original resolution (13.7 km). The blank pixel is water body. Units: m^3/m^3. They are 40-cm layer soil moisture.

4.2 Effects of ISMS on streamflow predictability with the perfect (observed) climate forcing

Figure 17 show improvements in the streamflow forecast due to realistic ISMS effects and the observed climate forcing. We computed the RMSD between the forecast ensemble experiment and the CTRL and then compared it between the E1 (real ISMS + obs. CF) and the E2 (rand. ISMS + obs. CF) experiments (Table 4). The difference between these two experiments can be attributed to the ISMS under the observed climate forcing. A smaller value of the RMSD suggests a reduced forecast error, and thereby improvement in the forecast. Results shown are for all 15 years together, five El Niño years, five La Niña years, and five normal years. The shaded region shows a 95% uncertainty range in the mean RMSD estimate, and in the respective experiments.

The realistic ISMS brings significant improvement in the forecasts as seen by a significantly smaller RMSD in the realistic ISMS experiments compared with the random ISMS effects. The improvement in forecast is found for almost all 90 days. This result suggests that ISMS can improve forecast skill at a seasonal time scale.

A closer look at the RMSD growth trajectory shows that the RMSD grows at a faster rate in the first 10 days of the forecast but then the RMSD declines during the winter and spring seasons (First two columns of Figure 17). Initial growth in the RMSD can be related to the lag effect between the soil moisture and streamflow (discussed later). The difference between the realistic and random ISMS experiments also grows initially, but then declines as the forecast lead increases during the winter and the spring seasons, i.e. the climate forcing effects become predominant with the increasing forecast lead-time.

To investigate the seasonality issue, we analyzed the absolute values of the streamflow forecasts and the CTRL (Figure 18). We found that the declining RMSD with the increasing lead-

time is not attributable to the seasonal variability in streamflow. For example, streamflow generally increases from January to March in the ACT basin (Figure 18) but the RMSD declines(Figure 17).

In the summer and fall seasons, ISMS plays a larger role in predicting the streamflow than the climate forcing. This can be seen from the RMSD growth as the difference between the E1 and the E2 experiments' RMSDs does not decline after the initial growth period. We also found that dry years have a higher predictability than wet years by segregating the streamflow predictability in El Niño (Wet), La Niña (Dry), and Normal years shown in rows two to four in Figure 17 (wet dry year figure not shown). The RMSD differences between the E1 and the E2 experiments are larger for the dry years particularly in the winter and spring seasons.

Reduction in the RMSD due to ISMS effects brings in significant improvements in the NSE for seasonal streamflow forecast. We assessed the NSE metric using ensemble average results for the nine 10-days forecast periods (Eq. 2 and Figure 19). We adopted the 10-days forecast window period from the NWM medium forecast range (https://water.noaa.gov/map). The realistic ISMS brings statistically significant improvements in NSE for all 90 days forecasts in the summer and fall seasons. The NSE score is close to one with minimal spread in the realistic ISMS experiments; whereas the mean NSE score drops to 0.9 or less in the summer and 0.8 or less for the fall season (rand. ISMS + obs. CF). The significant improvement in the NSE is limited to the first 30 days in the winter.

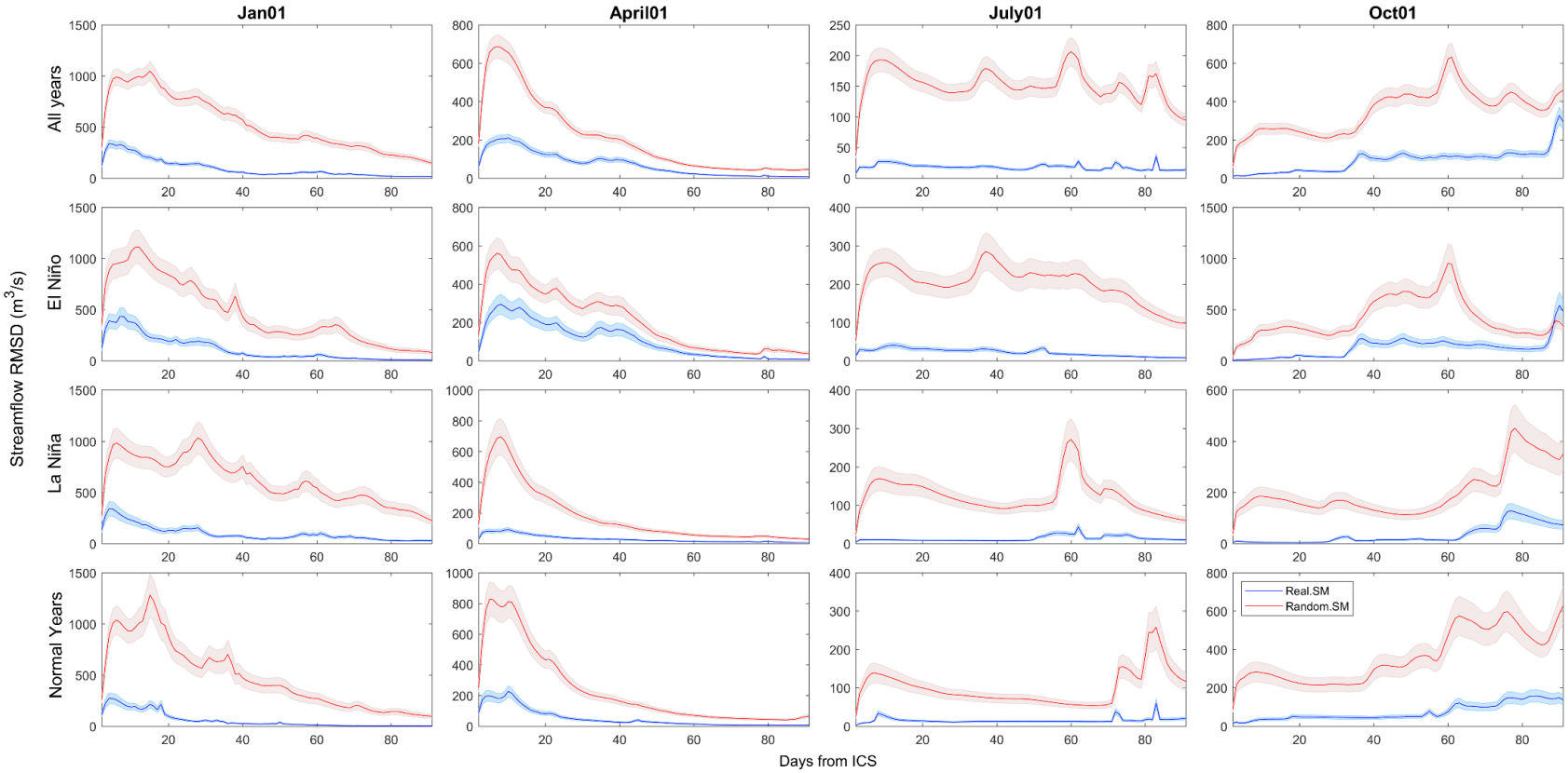


Figure 17: Effects of ISMS and CF on seasonal streamflow predictability. Forcing is observation based reanalysis NLDAS data. The figure shows the RMSD with realistic and randomized ISMS in all 15 years (first row), 5 El Niño years (second row), 5 La Niña years (third row), and 5 normal years (fourth row). Simulations initialization dates are the first day of January, April, July, and October, then averaged over 14 ensemble members and 15 selected years as a function of simulation sustain time (1-91) days. The solid lines are results of equation (1). Shaded regions show 95% confidence interval uncertainty range as equation (2).

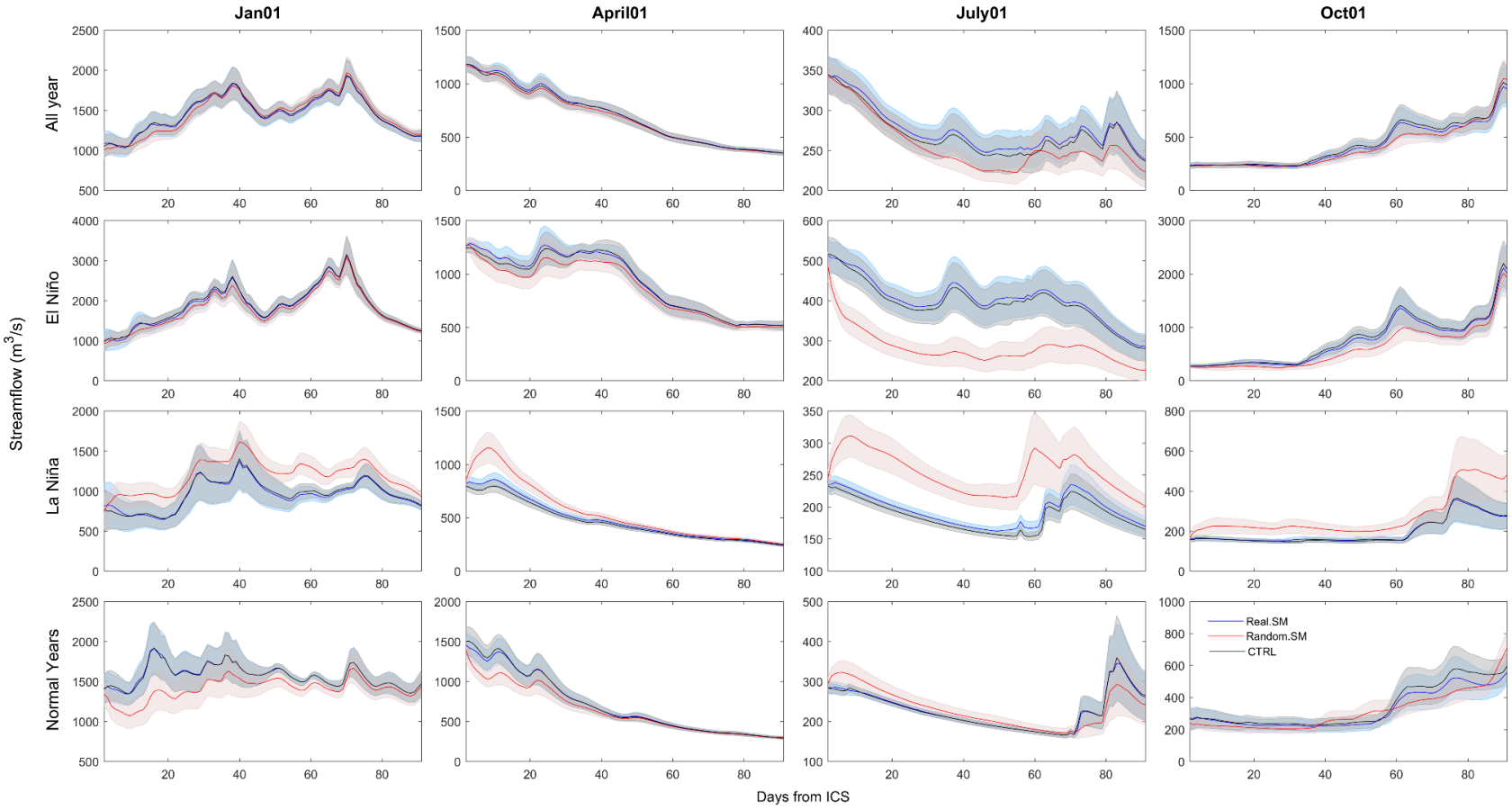


Figure 18: It is the same experiment as Figure 17. The only difference is that they are streamflow values in all years, El Niño, La Niña, and normal years. Shaded regions show 95% confidence interval uncertainty range. (Unit: m^3/s). Forcing is based reanalysis NLDAS data. Blue line is experiment with realistic ISMS, red line is experiment with random ISMS, the black line is model control run.

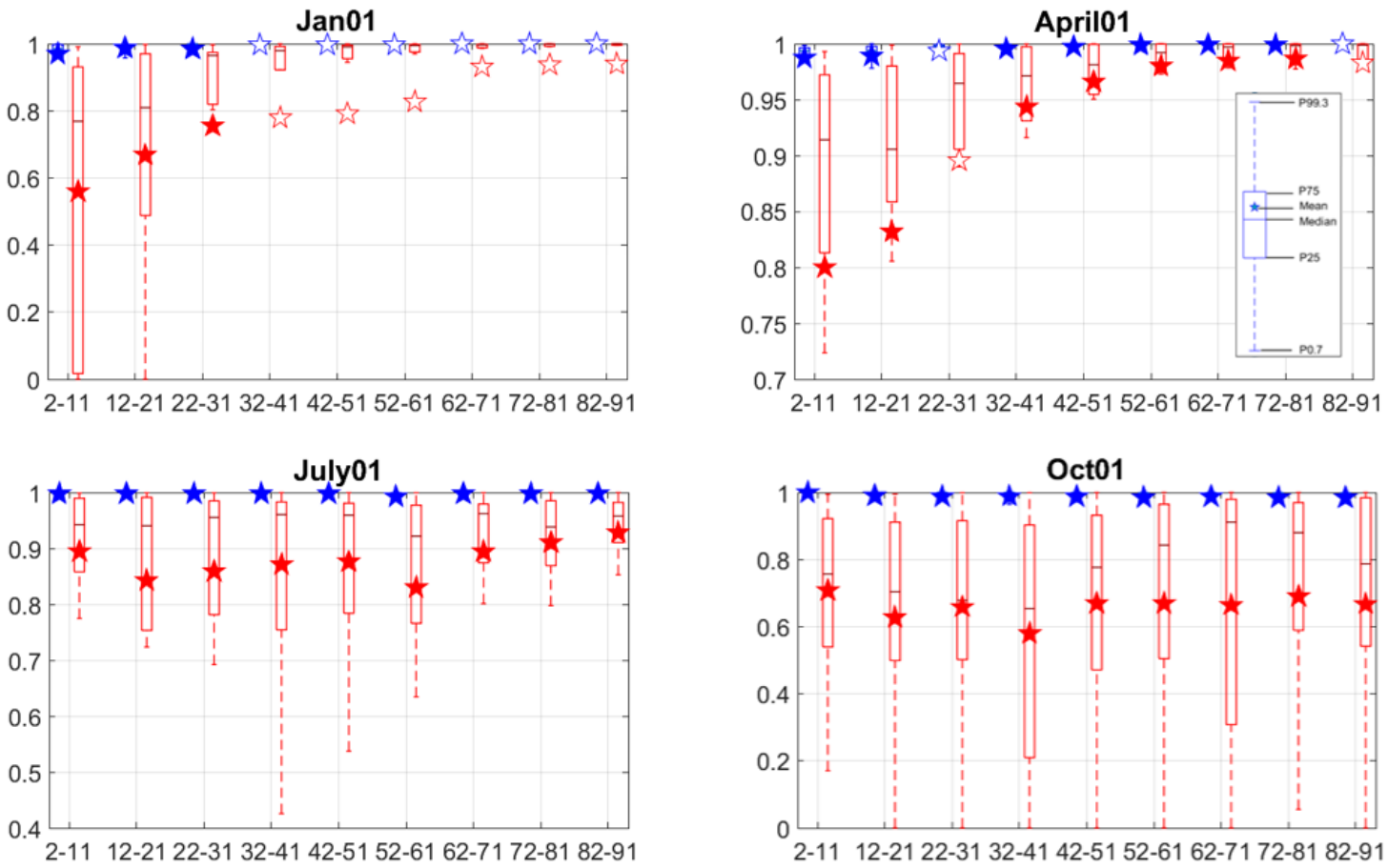


Figure 19: Streamflow in all 15 years NSE in each 10 days. X axis is each of the 10-days. Y axis is NSE. Blue is realistic ISMS, red is random ISMS. NSE less than zero is meaningless, so zero replaces NSE values less than zero. As equation (1), simulations in each 10 days averages over the 14 ensemble members. So it is 15 years NSE distribution in each 10 days interval. When the two distributions are significant in t-test with 95% confidence level, their average values are marked with solid star.

4.3 Effects of ISMS on streamflow predictability with random climate forcing

One advantage of the streamflow predictability experiment design is that we can re-arrange the ensemble to generate realistic and random ISMS experiments but with the random climate forcing. This climate forcing does not correspond to the same years as in the realistic ISMS conditions (i.e., 14 out of 15 years). We repeated the same for all 15-forecast years. Thus, we generated 2×14 ×15 ensemble sets for realistic and random ISMS but with the random climate forcing as shown by the E3 and the E4 experiments in Table 4. The differences between the E3 and E4 experiments are directly attributable to the effect of ISMS, only.

The ISMS brings in significant improvement in the streamflow forecast even with random climate forcing (Figure 20). The RMSD of the realistic ISMS experiment remains significantly smaller than the random ISMS experiment for several days in the initial forecast period and then becomes similar to the random ISMS experiment at the longer lead. We defined the ISMS predictability time scale as the first lead days when the signal becomes statistically indistinguishable between realistic and random ISMS experiment and with random climate forcing, i.e., their 95% uncertainty range as shown by blue and red shaded regions overlap (Figure 20). The predictability due to ISMS persists for ~ 60 days in the summer, and ~ 20 days in the spring. In the winter, the predictability time scale is approximately 30 days. By comparing rows 2 and 3 in Figure 20, we found that the predictability time scale is longer in La-Niña (dry) years than El-Niño (wet) years for all four seasons. This result implies that realistic ISMS can improve the streamflow forecast skill despite limited skill in the climate forecast.

I found minimal spatial variability in predictability time-scale across 15 sub watersheds (Figure 21). I performed the predictability analysis for each sub-watersheds separately, and plotted their predictability time scale in Figure 21. As expected, the predictability time scale is smaller in

JFM and AMJ seasons (wet), and higher in JAS and OND seasons (dry season). The predictability time-sale is mostly spatially uniform except in the spring season. A generally higher predictability time-scale is found in the headwater reaches than the downstream in AMJ seasons. The spatial variability can be topography and climate characteristics (not investigated). The predictability metrics is signal to total ratio. If the ratios for realistic and random ISMS are no different, or the ratio cannot pass significant test, the days before that point is the predictability time scale.

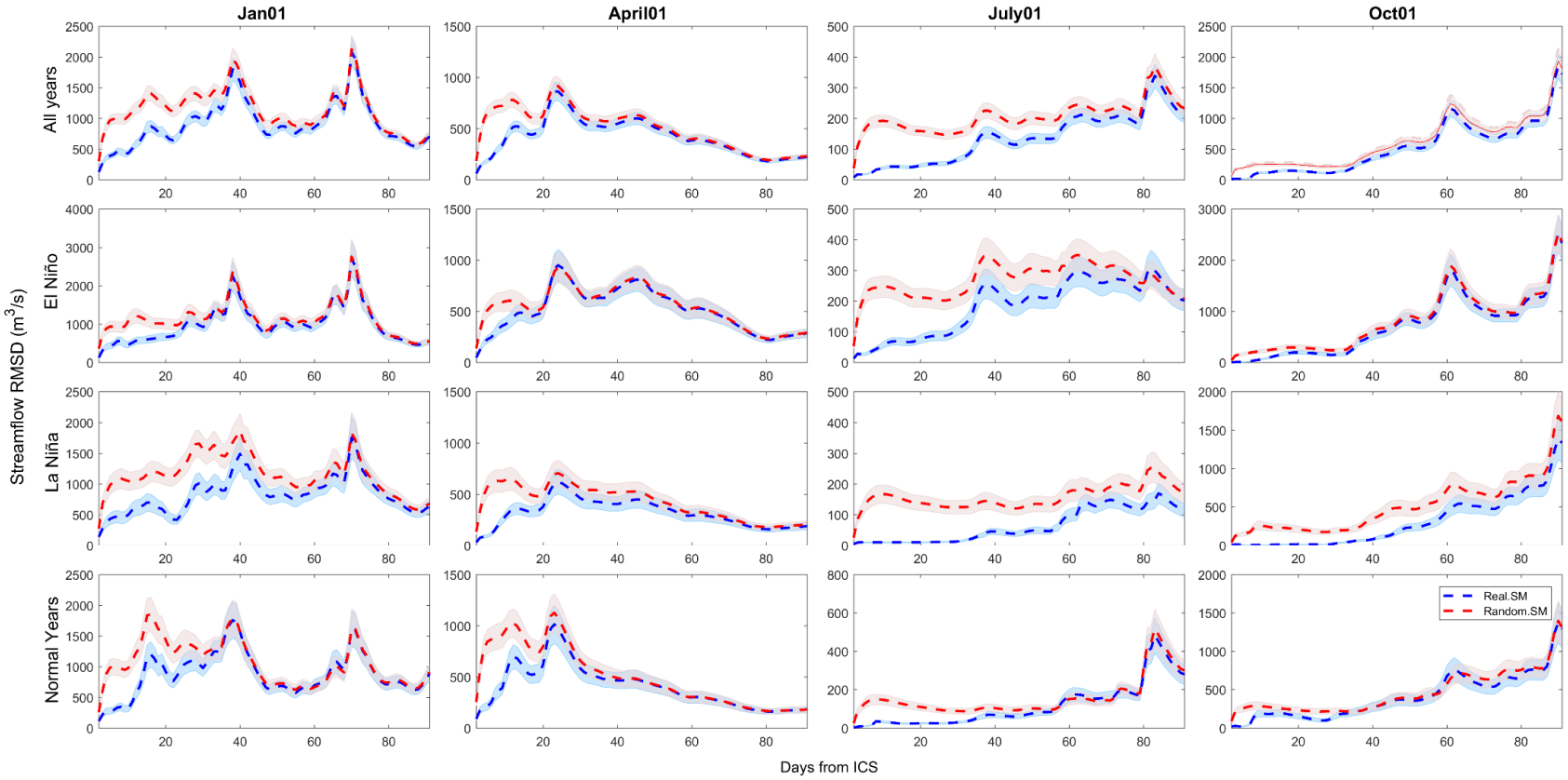


Figure 20: Effects of ISMS only on the seasonal streamflow predictability. Forcing is random climate forcing. The figure shows the RMSD with realistic and randomized ISMS in all 15 years (first row), 5 El Niño years (second row), 5 La Niña years (third row), and 5 normal years (fourth row). Simulations initialization dates are the first day of January, April, July, and October, then averaged over 14 ensemble members and 15 selected years as a function of simulation sustain time (1-91) days. The dash lines are results of equation (6) and (7). Shaded regions show 95% confidence interval uncertainty range as equation (8).

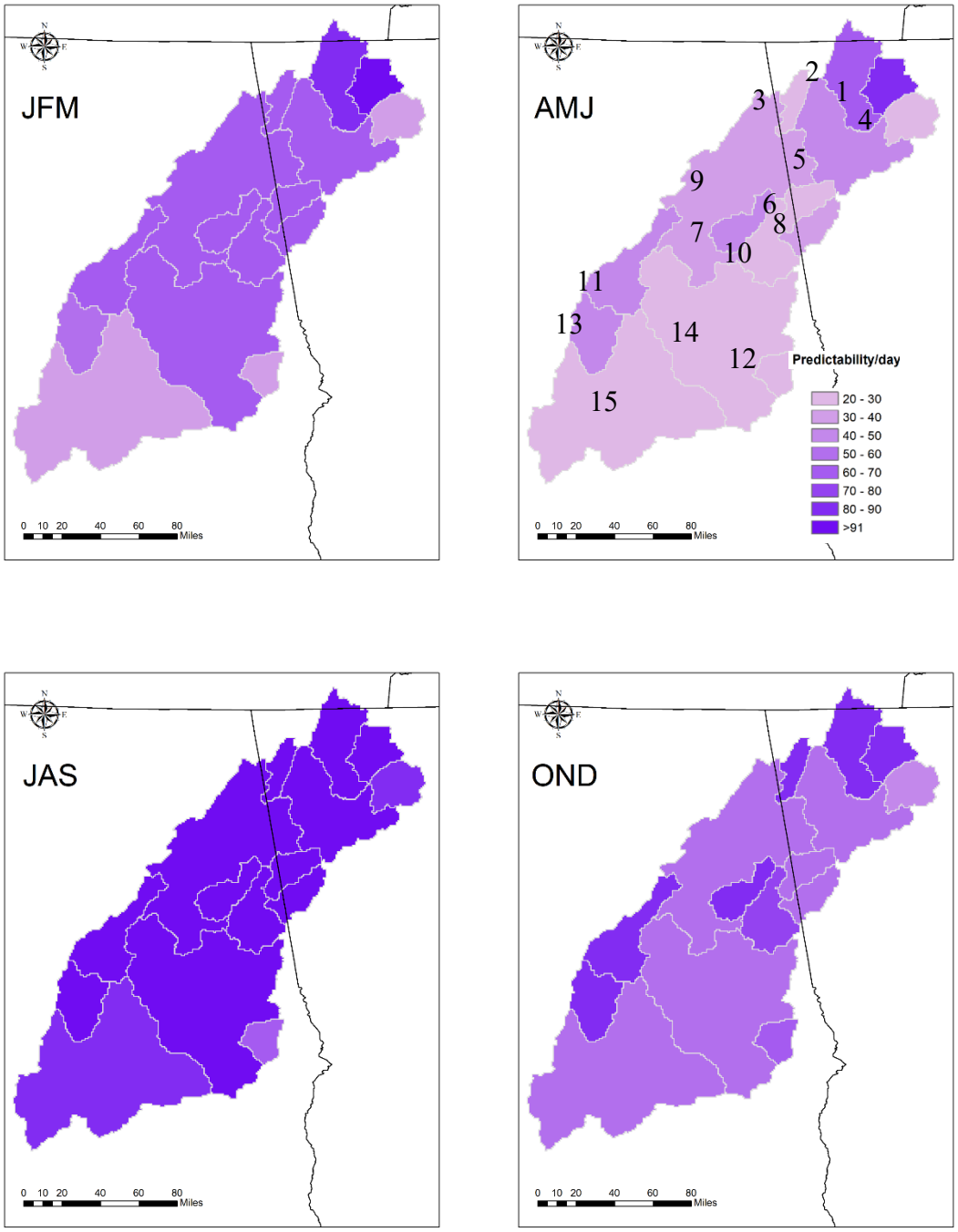


Figure 21. Streamflow predictability time scale with random forcing in each subwatershed. In this figure, predictability was calculated for each subwatershed outlet.

4.4 Role of the ground water in streamflow predictability in the NWM model

We assessed the role of ground water (GW) by randomizing the soil moisture state variables below 2m (Table 3) and comparing it with the realistic ISMS effects and with the observed climate forcing. Since the realistic ISMS experiments also included realistic GW states, and so the difference between the realistic ISMS experiment and randomized GW experiments can be attributed to the effects of the ISMS condition from surface to 2m depth. When the RMSD of the randomized GW experiment becomes equal to or less than the realistic ISMS forecast experiment, then we called it the loss of predictability due to GW.

The GW contribution to the streamflow predictability is generally smaller than the full field ISMS contributions (Figure 22). This can be seen by comparing the area between the blue and yellow curve with the area between blue and red curves in the figure. By comparing winter (Jan 1) results with those of summer (July 1), we found that GW contribution to streamflow predictability is higher in the summer than winter. The summer and fall seasons show maximum groundwater contribution that is limited to the 0-40 days forecast period. Compared to the realistic ISMS experiment, a smaller RMSD in the randomized GW experiment suggests that even a small error in the 0-2m soil moisture state can produce a larger error in the streamflow forecast than the randomized GW states (see the 40-90 days forecast in the fall). In the summer, the GW contribution is higher in the El-Niño and normal years than La-Niña years. The GW contribution to streamflow predictability is minimal in the winter and spring season. Considering the simplistic representation of the GW in the NWM, the GW results may not be robust here because ground water was expected to have a longer predictability time scale (Sutton 2019).

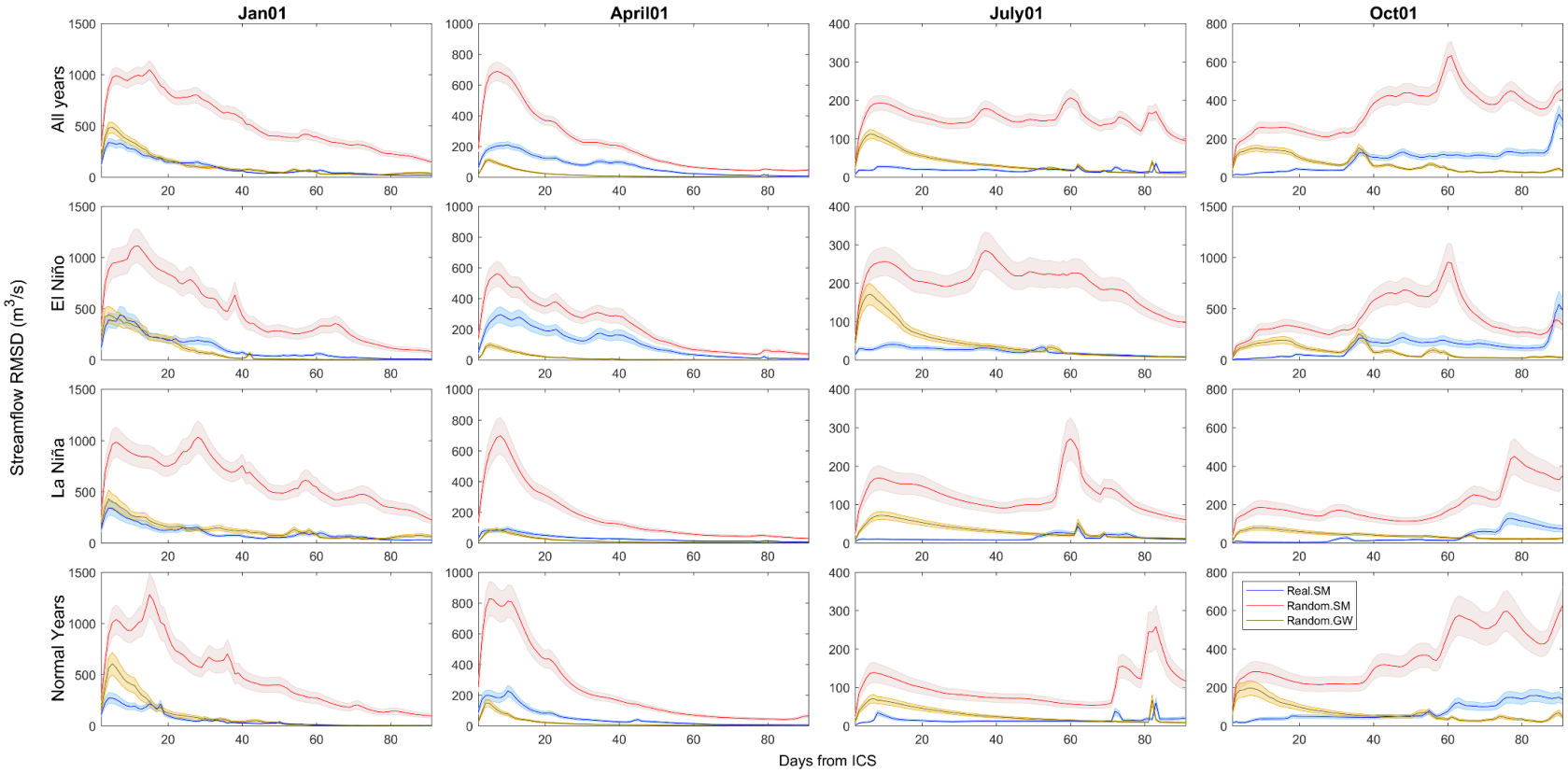


Figure 22: Comparing the GW contribution to that of ISMS experiments, effects of ISMS and CF on seasonal streamflow predictability. Same as Figure 17 with added data from randomized GW experiments with observed climate forcing. Forcing is observation based reanalysis NLDAS data. The figure shows the RMSD with realistic and randomized ISMS in all 15 years (first row), 5 El Niño years (second row), 5 La Niña years (third row), and 5 normal years (fourth row). Simulations initialization dates are the first day of January, April, July, and October, then averaged over 14 ensemble members and 15 selected years as a function of simulation sustain time (1-91) days. The solid lines are results of equation (6) and (7). Shaded regions show 95% confidence interval uncertainty range as equation (8).

4.5 Effects of ISMS on soil moisture predictability using perfect (observed) and random forcing

We investigated the effects of ISMS on the soil moisture predictability itself by comparing the 0-2m soil moisture RMSD between the E1 and E2 experiments (Table 4, and Figure 23). As shown in the last section, 0-2m soil moisture condition contributes most to the streamflow predictability; and hence, 0-2m choice is justified for water resource application. We computed the basin average soil moisture RMSD using 13.6 km LSM gridded data. The realistic ISMS experiments have significantly smaller RMSD than the randomized ISMS experiments for all through the 90 days and in all four seasons; i.e. ISMS can bring in significant improvements in seasonal soil moisture forecasts in the ACT basin (first row in Figure 23). Comparison of the 2nd and 3rd rows, show that the soil moisture predictability is higher during La-Nina (dry) years than El-Nino (wet) years and during the wet season (JFM and AMJ). During the dry seasons (JAS, and OND), the differences between El-Nino and La-Nina years are minimal or slightly higher for the El-Nino and in the fall season. In fact, normal years (fourth row) show a higher difference than both El-Nino and La-Nina years in the fall season.

The characteristics of the soil moisture predictability are considerably different from that of streamflow predictability as can be seen by comparing Figures 22 and Figure 16. A major difference is that the RMSD of the soil moisture forecast is highest at the start of the forecast and then continuously declines for all through the 90 days of the forecast. This means that the effects of the ISMS perturbations is highest at 0th day and then decreases as both the experiments receive the same climate forcing. To explain the characteristic difference between the soil moisture and streamflow predictability, we performed lag-correlation analysis between soil moisture and soil moisture (auto-correlation) and soil moisture and streamflow (cross-correlation) in the CTRL experiment.

The soil moisture lag-autocorrelation decays with time consistent with the previous study (Sanjiv Kumar et al., 2019), but the soil moisture-streamflow lag-correlation first increases to reach a peak value several days after the soil moisture anomalies put into the system and then decays (Figure 25). This result makes physical sense because there can be several days of lag effects for initial soil moisture perturbations to propagate through the soil system, update the water table, and drive surface and sub-surface runoff to the streams. The cross-correlation between the soil moisture and the streamflow increases in the first several days, e.g. 10 days for the summer, then either remain high in the summer and fall seasons, or decreases in the winter and spring seasons. The RMSD growth and decay in Figure 17 and Figure 23 generally follow the corresponding lag correlation structure shown in Figure 25. We limited our lag-correlation to 20-days lag because longer lags will have significant overlaps from one season to the next season.

We also computed lag-correlation between NWM soil moisture and observed streamflow. Since long-term soil moisture observations are not available; we did not compute lag-correlation between observed soil moisture and streamflow. When NWM streamflow is replaced by USGS observed streamflow (Figure 26), the correlations between soil moisture and flow decrease but remain statistically significant. A drop in lag-correlation can be attributed to uncertainty in soil moisture parametrization in the NWM compared with the observed as discussed earlier.

Effects of the ISMS on the soil moisture forecast is generally smaller under the randomized climate forcing (Figure 27). However, when it matter most, i.e. during dry years and dry season the ISMS can bring in statistically significant improvement in the soil moisture forecast skill despite limited skill or no skill in the climate forecast (see La-Nina years and summer and fall RMSD difference in Figure 27). By comparing Figure 27, and Figure 20, we hypothesized that the climate forcing plays a greater role for improving the soil moisture forecast skill than that of

streamflow, whereas ISMS play a greater role for the streamflow forecast. We made a quantitative assessment of this hypothesis in the next sub-section.

Figure 24 and Figure 28 are the same as Figure 23 and Figure 27, they are for root zone layers (0-0.4m). Both of them have shorter predictability compared with 0-2m layers, since random ISMS for 4 layers will sustain longer than 2 layers only.

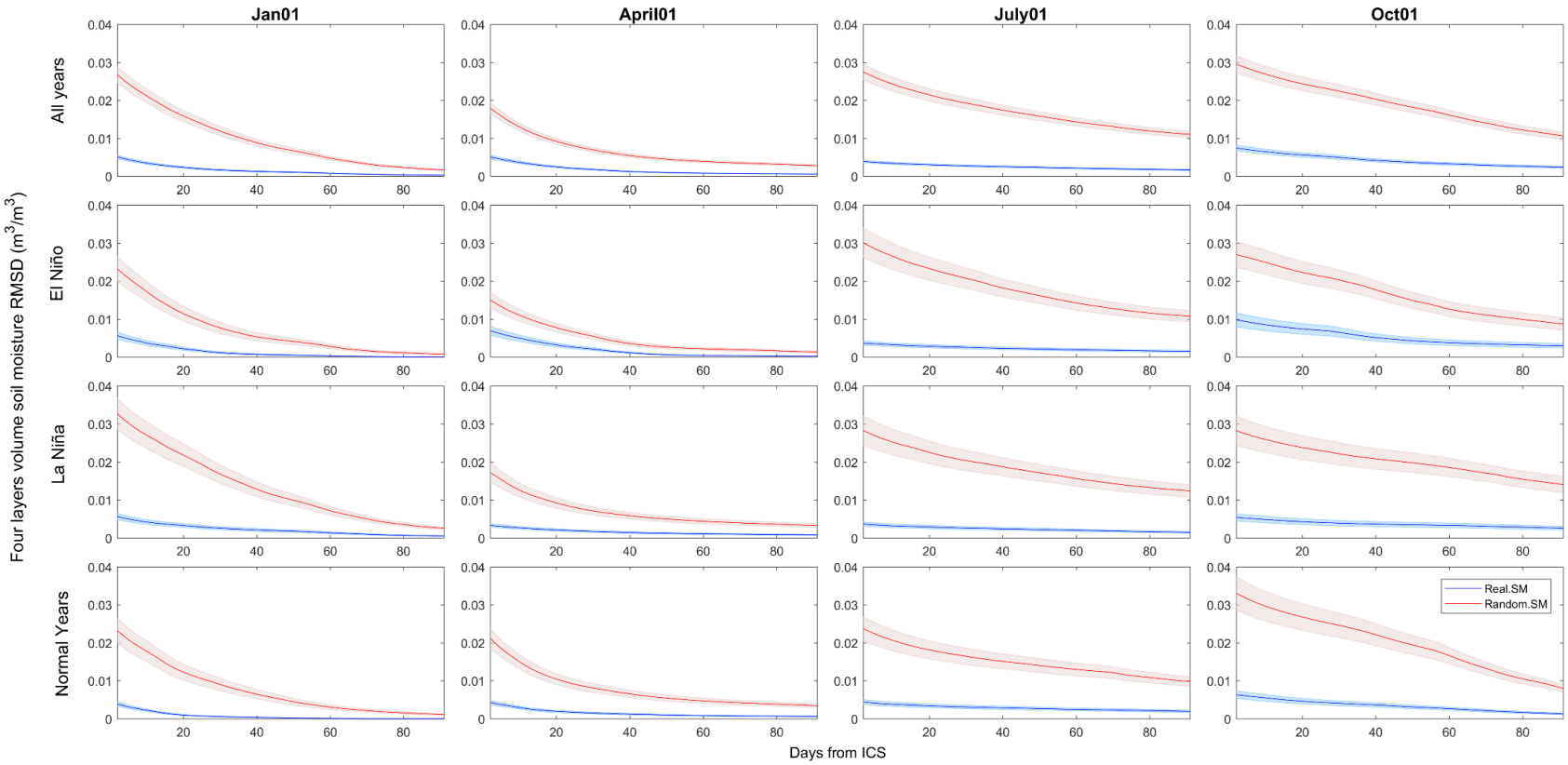


Figure 23: Effects of ISMS and CF on soil moisture predictability using perfect observed forcing. It is 0-2 m soil moisture averaged over the basin. Forcing is observation based reanalysis NLDAS data. The figure shows the RMSD with realistic and randomized ISMS in all 15 years (first row), 5 El Niño years (second row), 5 La Niña years (third row), and 5 normal years (fourth row). Simulations initialization dates are the first day of January, April, July, and October, then averaged over 14 ensemble members and 15 selected years as a function of simulation sustain time (1-91) days. The solid lines are results of equation (6) and (7). Shaded regions show 95% confidence interval uncertainty range as equation (8).

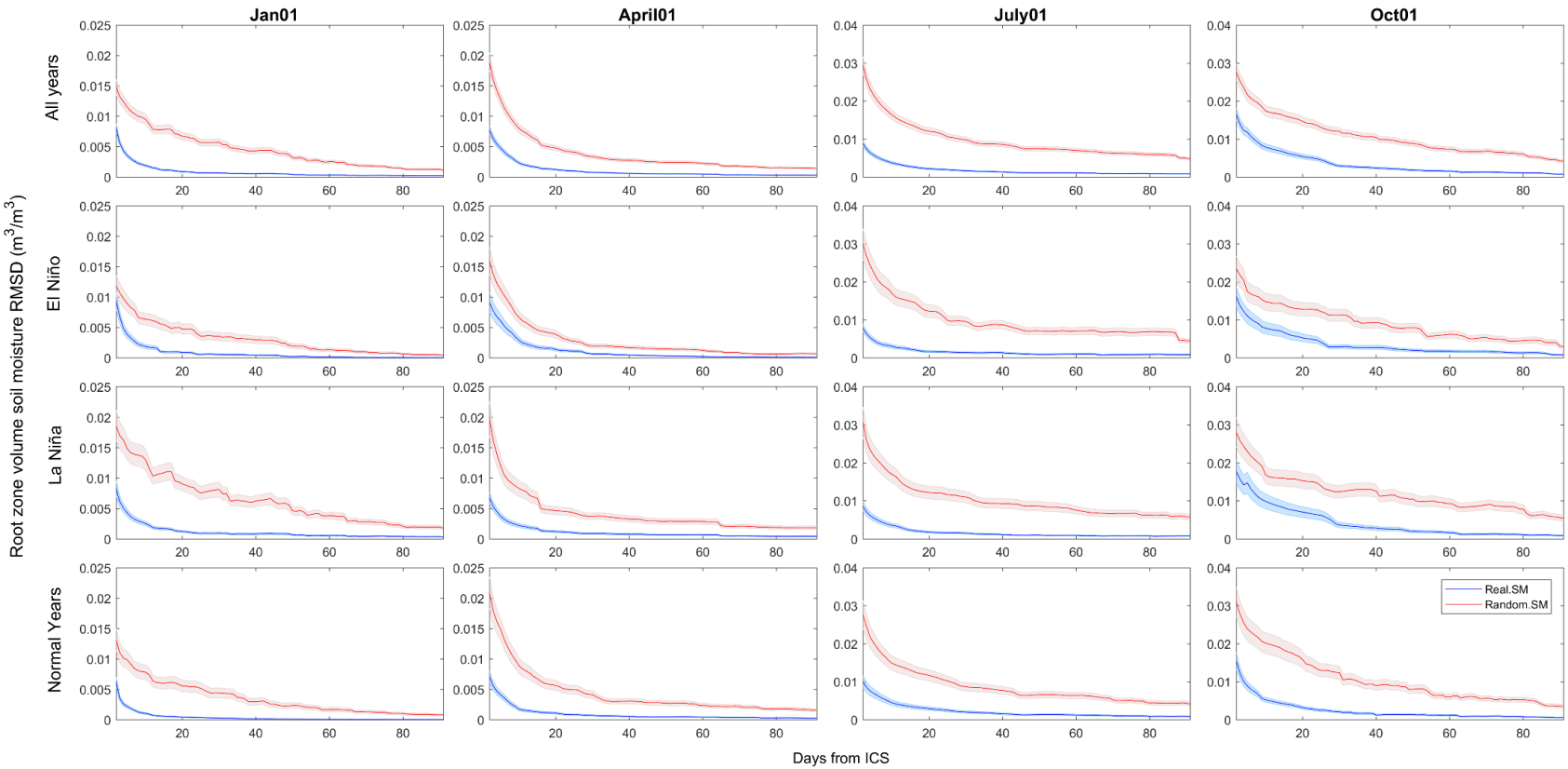


Figure 24: The same as Figure 23, but for 0-40 cm soil moisture averaged over the basin.

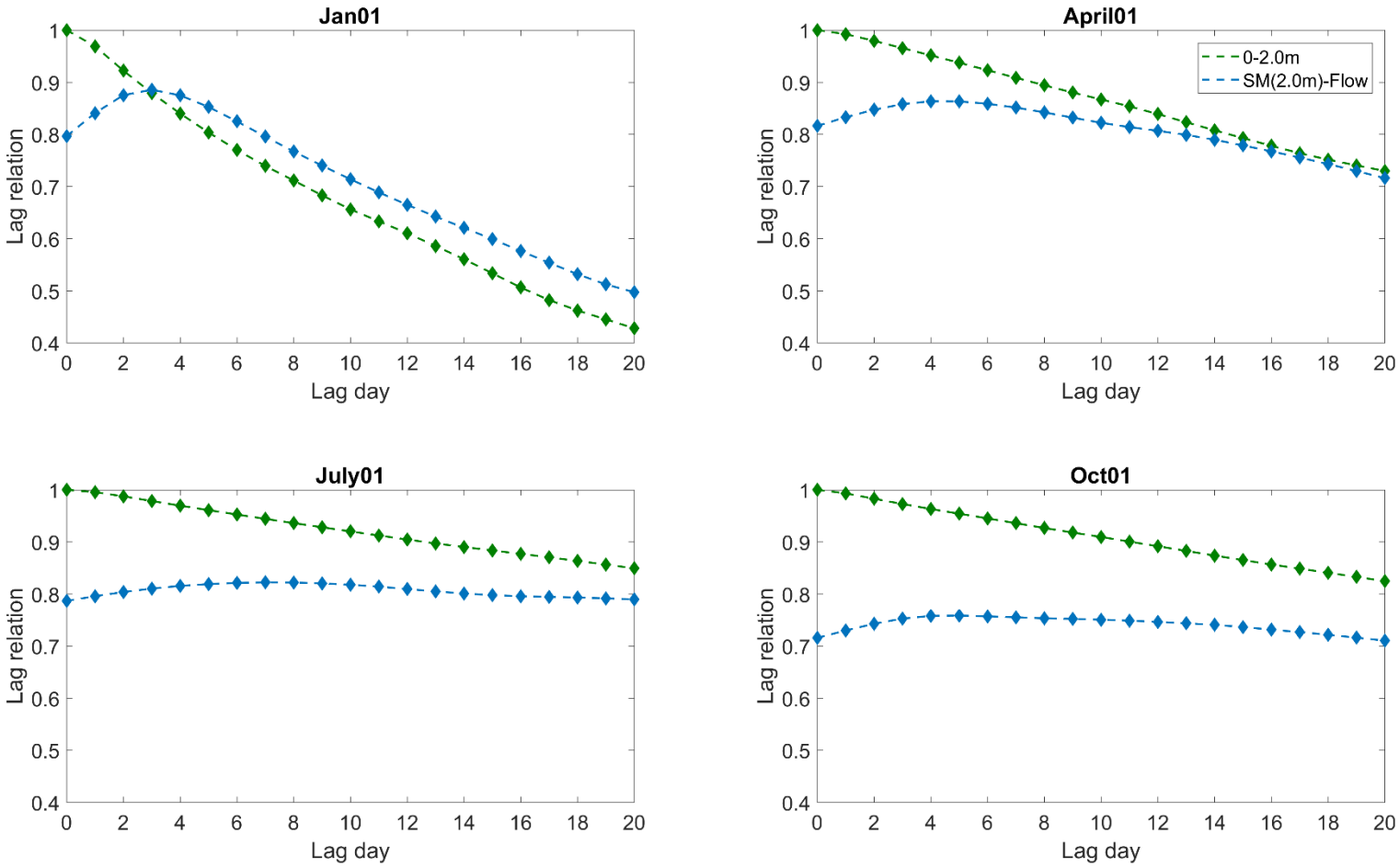


Figure 25: Pearson sample linear cross-correlations between soil moisture (0-2m) and streamflow with soil moisture as the lead (blue dotted lines), and the auto-correlation for the soil moisture (green dotted lines). We computed the correlation using CTRL experiment data from 1980-2016. X-axis is lag days from zero to twenty. Soil moisture is basin weighted average. Streamflow is at watershed outlet streamflow.

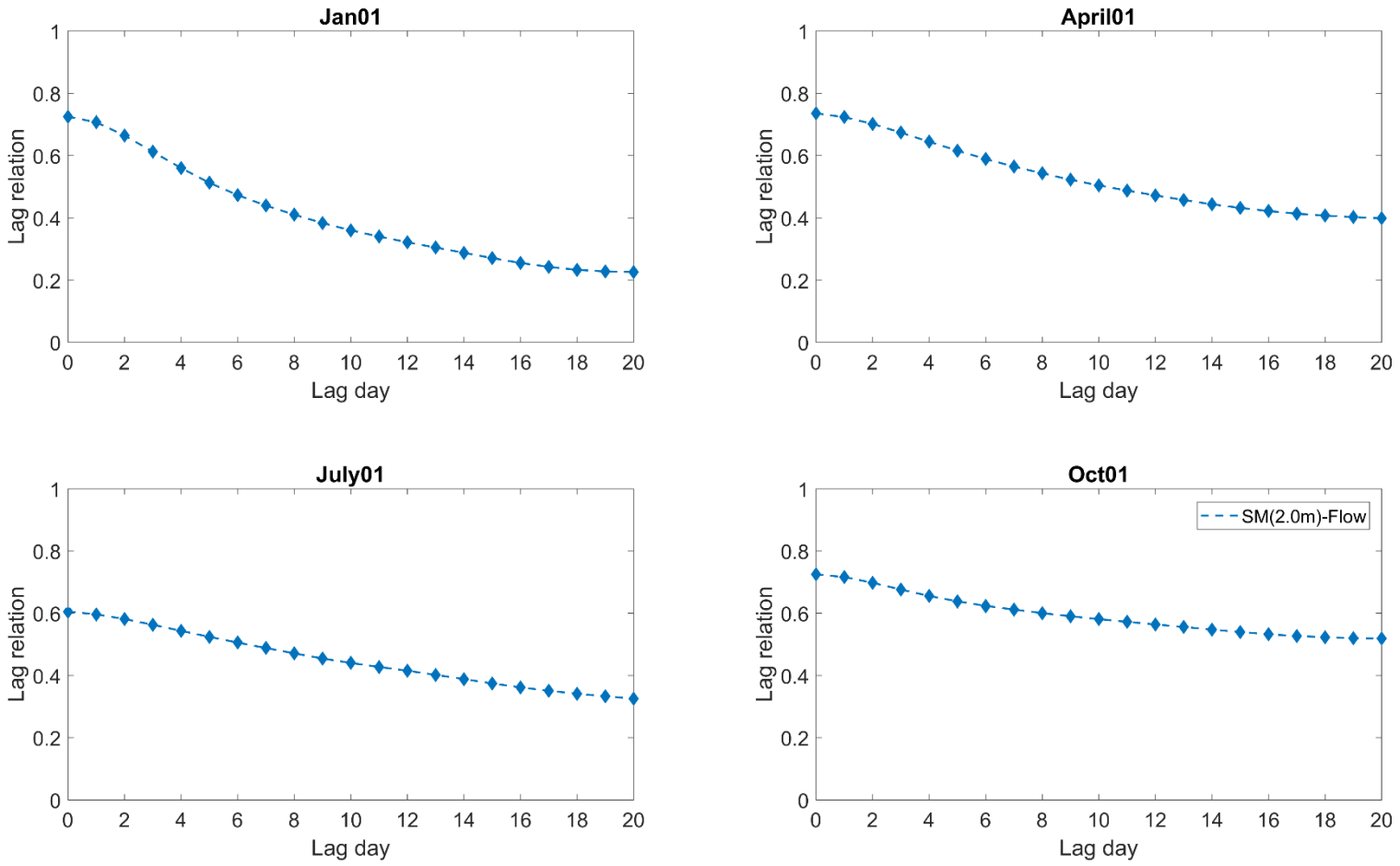


Figure 26: The same as Figure 25, the only difference is the Streamflow. It is USGS observed streamflow in the watershed outlet.

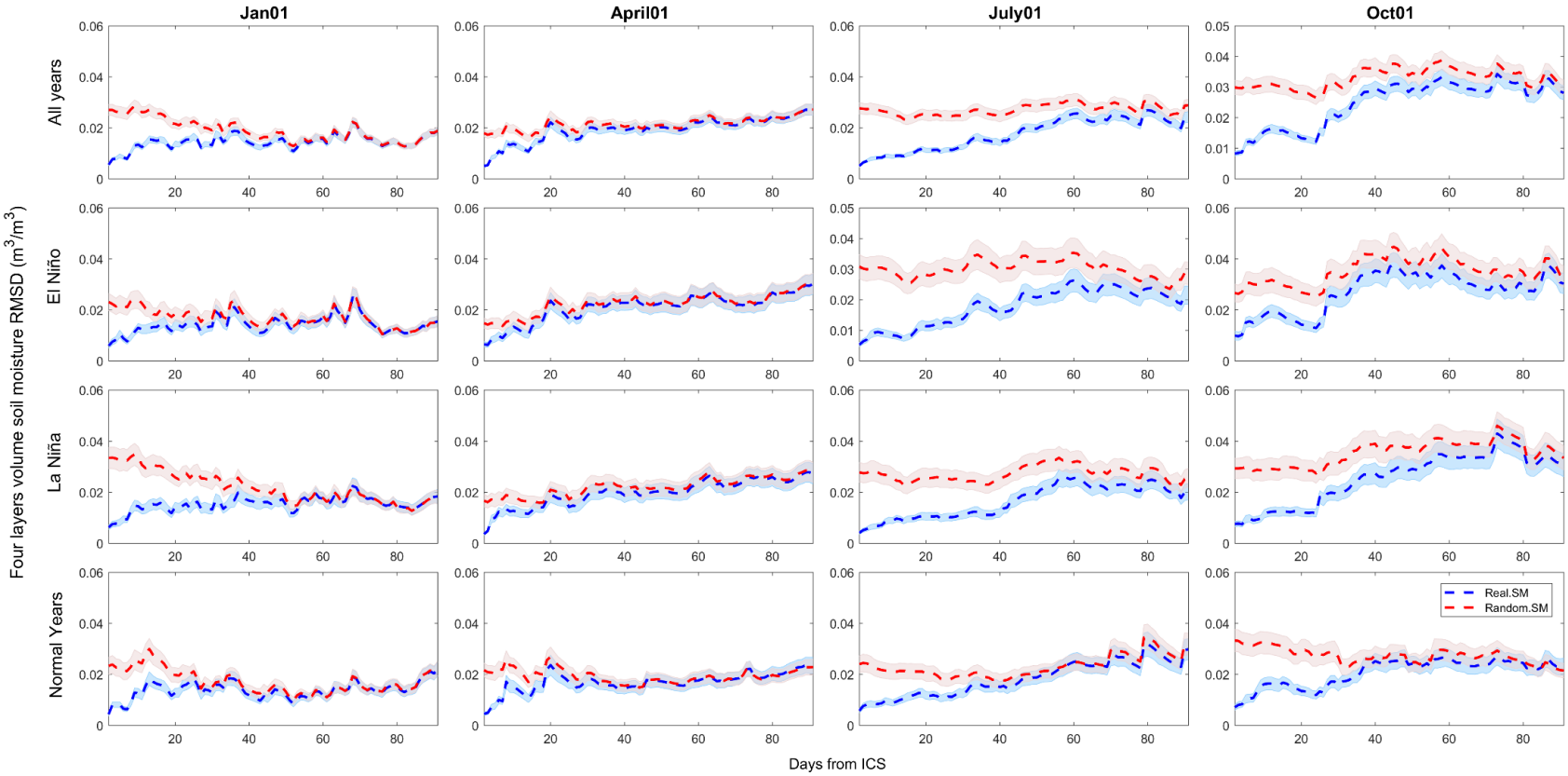


Figure 27: Effects of ISMS only on soil moisture predictability using random forcing. It is 0-2 m soil moisture averaged over the basin. The figure shows the RMSD with realistic and randomized ISMS in all 15 years (first row), 5 El Niño years (second row), 5 La Niña years (third row), and 5 normal years (fourth row). Simulations initialization dates are the first day of January, April, July, and October, then averaged over 14 ensemble members and 15 selected years as a function of simulation sustain time (1-91) days. The dash lines are results of equation (6) and (7). Shaded regions show 95% confidence interval uncertainty range as equation (8).

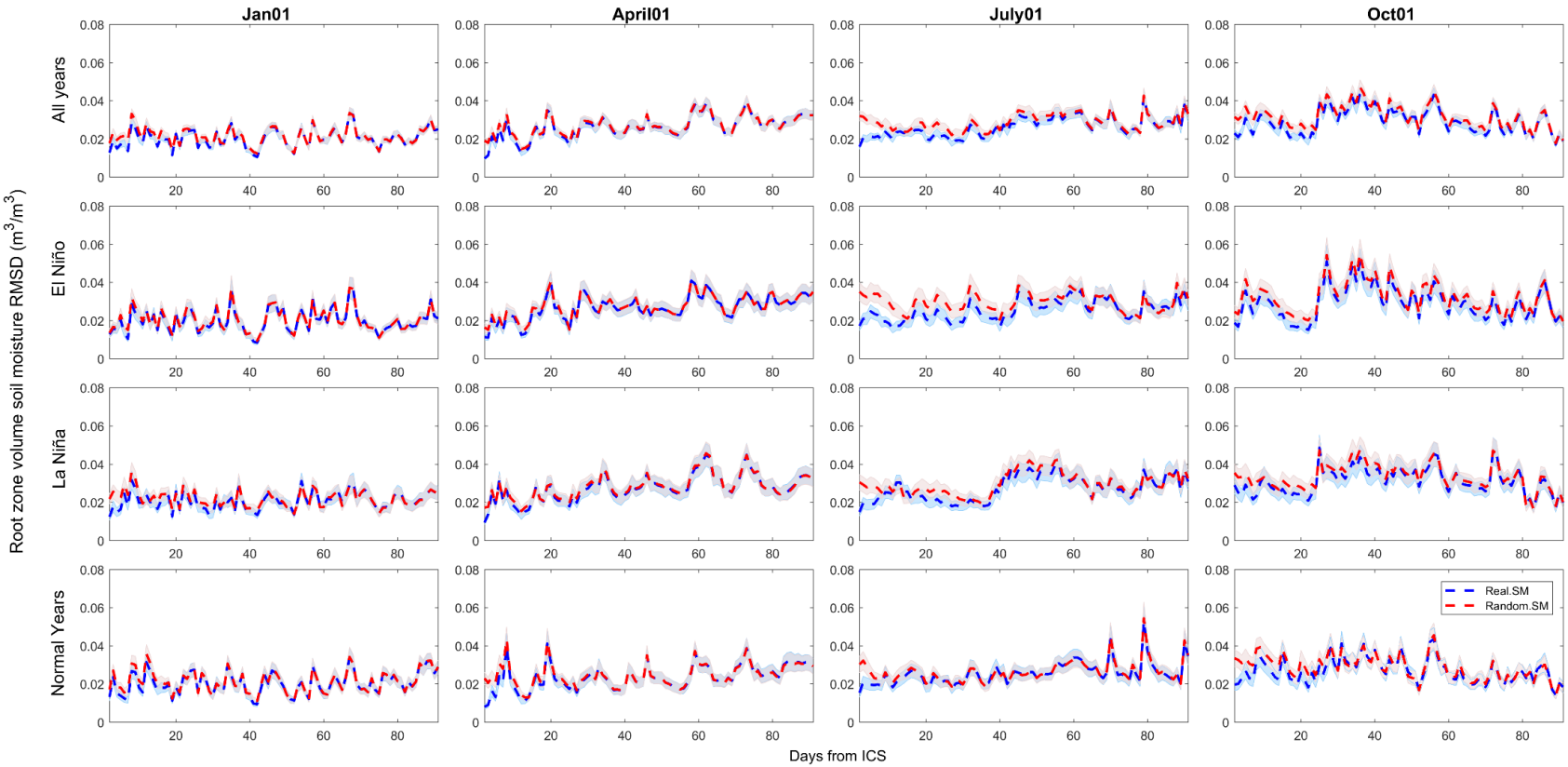


Figure 28: Same as Figure 27 for the 0-0.4 m soil moisture predictability using random climate forcing.

4.6 Relative importance of factors contributing to streamflow and soil moisture predictability

The ISMS plays a larger role for improving seasonal streamflow forecast than the soil moisture forecast. Figure 29 compares relative importance of the four sources of predictability: A – effects of the ISMS conditions only, B – effects of ISMS with observed climate forcing, C-effects of the climate-forcing anomaly only, and D- effects of climate forcing anomaly with realistic ISMS using signal to total ratio metric as discussed in the method section. For quantitative assessment, we aggregated the results for 0-90 days as the fraction of total area (1 × 90) contributed by the each sources of the predictability as shown in Table 6.

Figure 29 and Table 6; clearly show that ISMS contributes significantly to the seasonal streamflow forecast. During the summer, and fall seasons, ISMS contributes majority of the seasonal streamflow forecast skill. For example, the ISMS contributes 65% of the forecast skill compared with the 30% contribution by the climate forcing. Contributions of the ISMS only effects are also generally higher than that of ISMS with CF in the JAS season. Hence, we concluded that ISMS is a major source of seasonal streamflow forecast skill in the NWM, and contribution of the ISMS generally comparable or even higher than that of climate forcing at seasonal time scale.

The climate forcing plays an equally important role for the soil moisture forecast as the ISMS (Figure 29 and Table 6). For example, the ISMS contributes 48% of the forecast skill and climate forcing contributes to 49% in the summer season (JAS). It is worth noting that the ACT is a wet basin, and contributions of the ISMS can be higher in dry basins.

Table 6: Fraction of total seasonal forecast skill contributed by the ISMS and CF effects. These two sources are further sub-divided into A- ISMS only, B- ISMS with observed CF, C- CF only, and D – CF with realistic ISMS conditions.

	ISMS (ISMS only/ISMS with obs CF)	Climate Forcing (CF) (CF only/CF with realistic ISMS)
Season	Seasonal Streamflow Forecast	
JFM	30(13/17)	66(51/15)
AMJ	19(07/12)	76(67/09)
JAS	65(36/29)	30(15/14)
OND	53(27/26)	40(29/11)
	Seasonal Soil Moisture (0 – 2.0 m) Forecast	
JFM	27(10/17)	71(53/18)
AMJ	14(03/11)	83(75/08)
JAS	48(21/27)	49(26/23)
OND	39(16/23)	58(39/19)

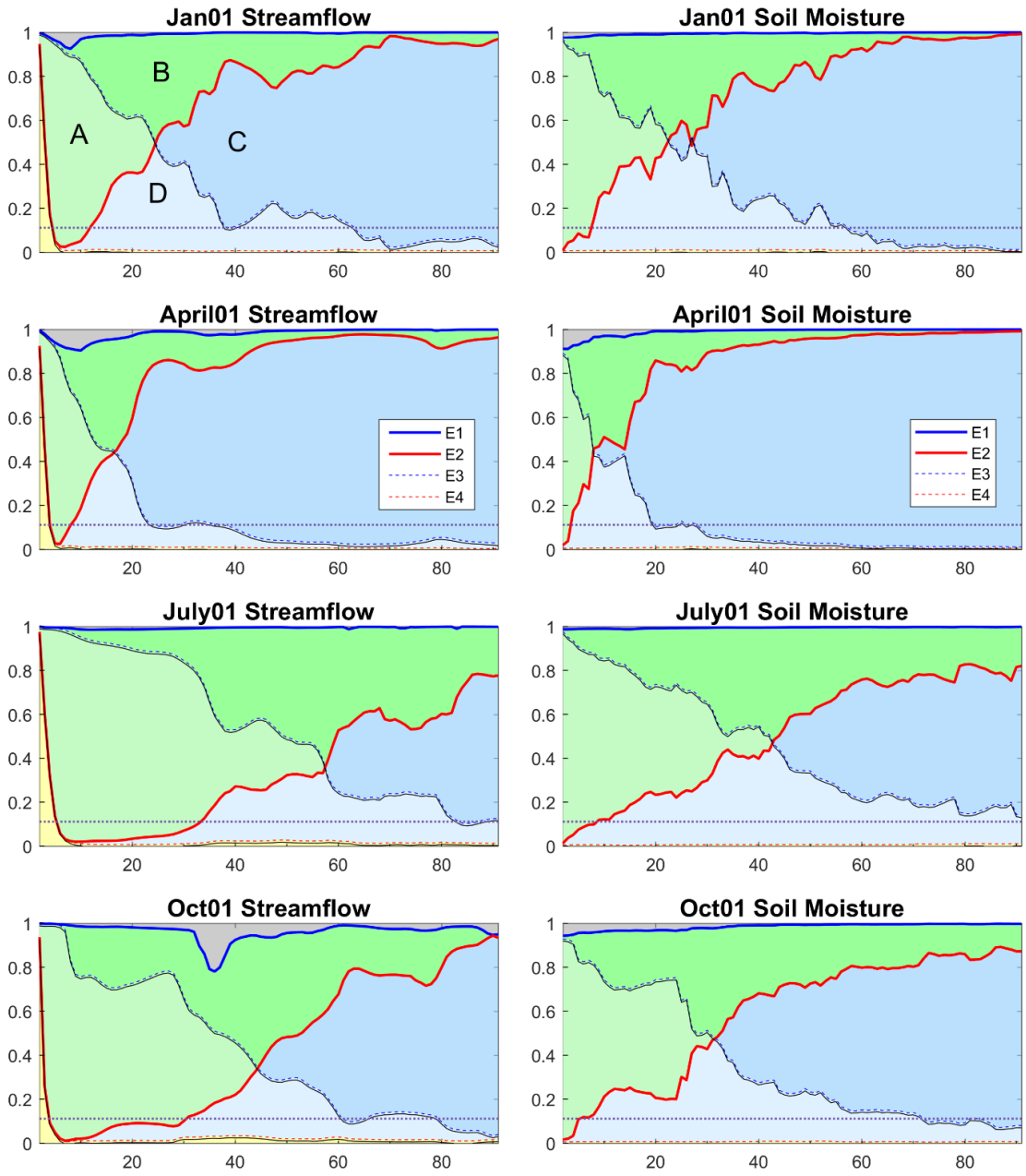


Figure 29: Signal to total ratio of streamflow (first column), 0-2 m soil moisture (second column) forecasts in the NWM seasonal forecast experiments. The solid and dotted blue lines represent realistic ISMS with the observed and random climate forcing, respectively (E1 and E3, Table 2). The solid and dotted red lines represent random ISMS with the observed and random climate forcing, respectively (E2, and E4). Green regions: A represent effect of ISMS only (with random climate forcing), and B is the effect ISMS under observed climate forcing. Blue regions: C represent effect of climate forcing only (with random ISMS), and D effect of climate forcing with realistic ISMS. See text for details. The dash purple line is threshold with 0.05 significant level. See text for details.

Figure 30 is for groundwater signal to total ratio. Initial soil moisture perturbations only contain soil moisture under 2-meters, so the biases in initial states is smaller than Figure 30. The yellow areas is predictability supported by other sources like 0-2 soil moisture. The ground water perturbations play a smaller role in the National Water Model as seen in a smaller green areas compared with blue (climate) and yellow areas (other sources). Predictability depends on seasonal climate condition. OND first 30 days are drier than JAS, deeper acquirer supports streamflow. It has longer memory, the effect of random groundwater remains longer time. As a result, groundwater conditions has a higher streamflow predictability in the OND than other seasons.

Comparing predictability from other sources (yellow region) in four seasons, when it transfers dry period to wet period, e.g. OND to DJF the climate forcing plays a larger role than ground water initial conditions and other sources. On the contrary, when wet period transition to dry period, e.g. AMJ to JAS then contribution of groundwater and other sources increase. Then convergence of dash red line and dash blue line prove that initial effect of groundwater perturbations is diluted with time increase.

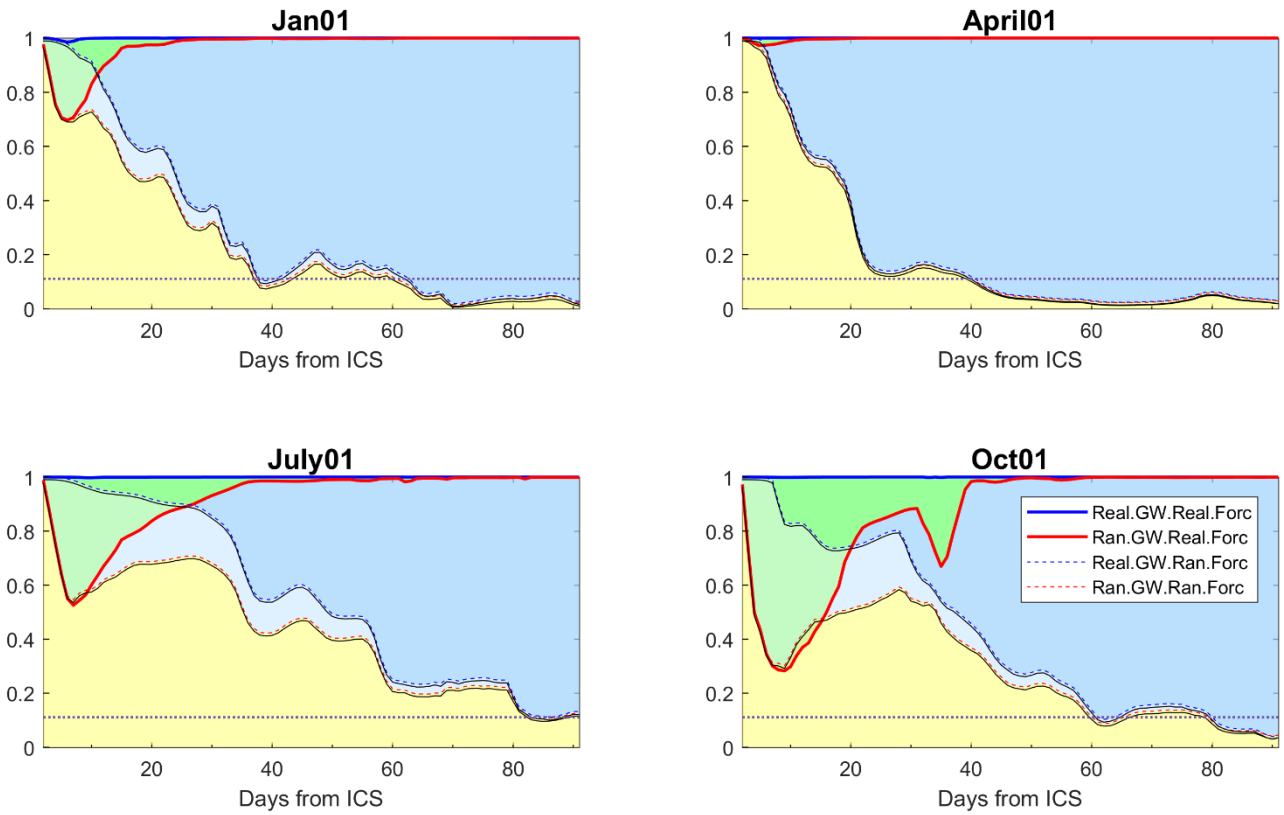


Figure 30: Signal to total ratio of streamflow forecasts in the NWM seasonal forecast experiments. It is similar with Figure 28, but for ground water initial states. The solid and dotted blue lines represent realistic groundwater initial condition with the observed and random climate forcing, respectively. The solid and dotted red lines represent random groundwater initial condition with the observed and random climate forcing, respectively. Green regions: A represent effect of groundwater initial condition only (with random climate forcing), and B is the effect groundwater initial condition under observed climate forcing. Blue regions: C represent effect of climate forcing only (with random initial condition), and D effect of climate forcing with realistic groundwater initial condition. See text for details. The dash purple line is threshold with 0.05 significant level. See text for details.

Chapter 5 SUMMARY AND CONCLUSIONS

This study finds that ISMS has a greater impact on streamflow predictability than soil moisture predictability at seasonal time scale. Effects of ISMS is comparable or even greater than that of CF for seasonal streamflow forecast in dry seasons. This is a new result compared with a previous study where authors conducted sub-seasonal streamflow and soil moisture forecast experiments using a simple water balance model and for 22 catchments in Switzerland (R Orth & S I Seneviratne, 2013). They found that ISMS has a greater impact on soil moisture predictability than on streamflow predictability. Our study used coupled hydrological model that included two way coupling between a physically based water and energy balance model (Noah-MP) and hydrology model for overland and sub-surface flow, and an additional stream/lake model for flow routing in this streams. Hence, we attributed a greater improvement in the streamflow forecast skill to the complex model structure that allows realistic streamflow simulations.

This study also provides a physically plausible explanation of a higher improvement in the streamflow forecast than soil moisture forecast. The explanation is based on lag effects between soil moisture and streamflow anomalies (Figure 23). A longer lag-time to reach peak correlation between soil moisture anomalies and streamflow anomalies and a slower decay of the cross-correlation lead to a higher impact on streamflow predictability, e.g. winter versus summer in Figure 23. Hence, we hypothesized that streamflow forecast skill can be significantly improved beyond the skill in the seasonal climate forecast from two additional sources: (1) soil moisture memory effects (Sanjiv Kumar et al., 2019), and (2) lag-effect between soil moisture and streamflow anomalies.

The seasonal cycle of soil moisture variability is muted in the NWM model compared with the SMAP data. This result is consistent with a previous study where authors have compared soil

moisture simulated by the Noah land surface model with the SCAN sites in Alabama [see Fig. 7 in (Xia, Ek, Wu, Ford, & Quiring, 2015)]. Dirmeyer et al. (2016) also found that Noah land surface model generally underestimate the soil moisture variability. In Alabama, the problem of underestimated soil moisture variability can be larger due to presence of Karst geologic feature that allows rapid drainage of soil water from the bottom of the soil column and is not parametrized in Noah land surface model (Kuniansky, Weary, & Kaufmann, 2016; Norton, 2018). The soil infiltration is not parameterized well in the model, it leads to baseflow simulation biases. The underestimated soil moisture variability can lengthen the soil moisture memory time scale, and thereby can affect streamflow predictability (Paul A Dirmeyer et al., 2016; R Orth & S I Seneviratne, 2013). Therefore, it is likely that our predictability estimate is somewhat positively biased. Model parametrization improvements can bring in further refinements in our predictability estimates.

The forecast error (RMSD) increases with increasing lead-time in the fully coupled climate model seasonal forecast experiment (P. A. Dirmeyer et al., 2013; Kumar, Dirmeyer, Lawrence, et al., 2014). In this study, we used the observed climate forcing to drive land surface and hydrology model, and hence the forecast error decreased with the increasing lead-time particularly in the winter and spring when climate forcing has a greater effect (Figures 4.2.1, and Table 6). The Land-atmosphere coupling can further improve the soil moisture forecast skill especially in the mid-latitude regions (S. Kumar et al., 2019). Furthermore, there is merits in developing fully coupled streamflow forecasting system as originally envisaged in the WRF-Hydro modeling framework (Gochis et al., 2018).

LIST OF REFERENCES

Amenu, G. G., Kumar, P., & Liang, X. Z. (2005). Interannual variability of deep-layer hydrologic memory and mechanisms of its influence on surface energy fluxes. *Journal of Climate, 18*(23), 5024-5045. Retrieved from <Go to ISI>://WOS:000234302400007.

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, R. D., . . . Wilson, T. (2013). US Climate Reference Network soil moisture and temperature observations. *Journal of Hydrometeorology, 14*(3), 977-988.

Berghuijs, W., Woods, R., & Hrachowitz, M. (2014). A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nature Climate Change, 4*(7), 583.

Cai, X. T., Yang, Z. L., David, C. H., Niu, G. Y., & Rodell, M. (2014). Hydrological evaluation of the Noah-MP land surface model for the Mississippi River Basin. *Journal of Geophysical Research-Atmospheres, 119*(1), 23-38. Retrieved from <Go to ISI>://WOS:000330930600003. doi:10.1002/2013jd020792

Campbell, G. S. (1974). A simple method for determining unsaturated conductivity from moisture retention data. *Soil science, 117*(6), 311-314.

Dirmeyer, P. A., Kumar, S., Fennessy, M. J., Altshuler, E. L., DelSole, T., Guo, Z. C., . . . Straus, D. (2013). Model Estimates of Land-Driven Predictability in a Changing Climate from CCSM4. *Journal of Climate, 26*(21), 8495-8512. Retrieved from <Go to ISI>://WOS:000330193800015.

Dirmeyer, P. A., Wu, J., Norton, H. E., Dorigo, W. A., Quiring, S. M., Ford, T. W., . . . Lawrence, D. M. (2016). Confronting Weather and Climate Models with Observational Data from Soil Moisture Networks over the United States. *Journal of Hydrometeorology, 17*(4), 1049-1067. Retrieved from http://journals.ametsoc.org/doi/10.1175/JHM-D-15-0196.1. doi:10.1175/JHM-D-15-0196.1

Entekhabi, D., Yueh, S., O'Neill, P. E., Kellogg, K. H., Allen, A., Bindlish, R., . . . Crow, W. T. (2014). SMAP handbook–soil moisture active passive: Mapping soil moisture and freeze/thaw from space.

Entin, J. K., Robock, A., Vinnikov, K. Y., Hollinger, S. E., Liu, S., & Namkhai, A. (2000). Temporal and spatial scales of observed soil moisture variations in the extratropics. *Journal of Geophysical Research: Atmospheres (1984–2012), 105*(D9), 11865-11877. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2000JD900051. doi:10.1029/2000JD900051

Felfelani, F., Pokhrel, Y., Guan, K., & Lawrence, D. M. (2018). Utilizing SMAP Soil Moisture Data to Constrain Irrigation in the Community Land Model. *Geophysical Research Letters, 45*(23), 12,892-812,902.

Givati, A., Gochis, D., Rummler, T., & Kunstmann, H. (2016). Comparing one-way and two-way coupled hydrometeorological forecasting systems for flood forecasting in the Mediterranean region. *Hydrology, 3*(2), 19.

Gochis, D. J., M. Barlage, A. Dugger, K. FitzGerald, L. Karsten, M. McAllister, . . . Yu, W. (2018). The WRF-Hydro modeling system technical description, (Version 5.0). *NCAR Technical Note*, 107. Retrieved from https://ral.ucar.edu/sites/default/files/public/WRFHydroV5TechnicalDescription.pdf.

Greuell, W., Franssen, W. H., & Hutjes, R. W. (2019). Seasonal streamflow forecasts for Europe-Part 2: Sources of skill. *Hydrology and Earth System Sciences, 23*(1), 371-391.

Guo, Z. C., Dirmeyer, P. A., & DelSole, T. (2011). Land surface impacts on subseasonal and seasonal predictability. *Geophysical Research Letters, 38*. Retrieved from <Go to ISI>://WOS:000298759700004. doi:Artn L2481210.1029/2011gl049945

Han, E., Crow, W. T., Holmes, T., & Bolten, J. (2014). Benchmarking a soil moisture data assimilation system for agricultural drought monitoring. *Journal of Hydrometeorology, 15*(3), 1117-1134.

Hooper, R. P., Nearing, G., & Condon, L. S. (2017). *Using the National Water Model as a Hypothesis-Testing Tool.* Paper presented at the CUAHSI Hydroinformatics Conference (2017: Tuscaloosa, Alabama).

Kam, J., & Sheffield, J. (2016). Changes in the low flow regime over the eastern United States (1962–2011): variability, trends, and attributions. *Climatic change, 135*(3-4), 639-653.

Kerandi, N., Arnault, J., Laux, P., Wagner, S., Kitheka, J., & Kunstmann, H. (2018). Joint atmospheric-terrestrial water balances for East Africa: a WRF-Hydro case study for the upper Tana River basin. *Theoretical and applied climatology, 131*(3-4), 1337-1355.

Koster, R. D., Mahanama, S. P., Livneh, B., Lettenmaier, D. P., & Reichle, R. H. (2010). Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nature Geoscience, 3*(9), 613.

Koster, R. D., & Suarez, M. J. (2001). Soil moisture memory in climate models. *Journal of Hydrometeorology, 2*(6), 558-570.

Kumar, S., Dirmeyer, P. A., & Kinter, J. L. (2014). Usefulness of ensemble forecasts from NCEP Climate Forecast System in sub-seasonal to intra-annual forecasting. *Geophysical Research Letters, 41*(10), 3586-3593. Retrieved from <Go to ISI>://WOS:000337610200036. doi:10.1002/2014gl059586

Kumar, S., Dirmeyer, P. A., Lawrence, D. M., DelSole, T., Altshuler, E. L., Cash, B. A., . . . Straus, D. M. (2014). Effects of realistic land surface initializations on subseasonal to seasonal soil moisture and temperature predictability in North America and in changing climate simulated by CCSM4. *Journal of Geophysical Research-Atmospheres, 119*(23), 13250-13270. Retrieved from <Go to ISI>://WOS:000346907100012.

Kumar, S., & Merwade, V. (2009). Impact of Watershed Subdivision and Soil Data Resolution on SWAT Model Calibration and Parameter Uncertainty 1. *JAWRA Journal of the American Water Resources Association, 45*(5), 1179-1196.

Kumar, S., & Merwade, V. (2011). Evaluation of NARR and CLM3.5 outputs for surface water and energy budgets in the Mississippi River Basin. *Journal of Geophysical Research-Atmospheres, 116*. Retrieved from <Go to ISI>://WOS:000289853100002.

Kumar, S., Merwade, V., Kam, J., & Thurner, K. (2009). Streamflow trends in Indiana: Effects of long term persistence, precipitation and subsurface drains. *Journal of Hydrology, 374*(1-2), 171-183. Retrieved from <Go to ISI>://WOS:000269071000014. doi:10.1016/j.jhydrol.2009.06.012

Kumar, S., Nemwman, M., Lawrence, D. M., Lo, M.-H., Akula, S., C-W., L., . . . Lombardozzi, D. (2019). The GLACE-Hydrology Experiment: Effects of Land-Atmosphere Coupling on Soil Moisture Variability and Drought Predictability. *Jounal of Climate, Submitted*.

Kumar, S., Newman, M., Wang, Y., & Livneh, B. (2019). Potential reemergence of seasonal soil moisture anomalies in North America. *Journal of Climate*(2019).

Kumar, S., Zwiers, F., Dirmeyer, P. A., Lawrence, D. M., Shrestha, R., & Werner, A. T. (2016). Terrestrial contribution to the heterogeneity in hydrological changes under global warming. *Water Resources Research, 52*(4), 3127-3142. Retrieved from <Go to ISI>://WOS:000377432800044. doi:10.1002/2016wr018607

Kuniansky, E. L., Weary, D. J., & Kaufmann, J. E. (2016). The current status of mapping karst areas and availability of public sinkhole-risk resources in karst terrains of the United States. *Hydrogeology Journal, 24*(3), 613-624.

Larocque, M., Fortin, V., Pharand, M., & Rivard, C. (2010). Groundwater contribution to river flows–using hydrograph separation, hydrological and hydrogeological models in a southern Quebec aquifer. *Hydrology and Earth System Sciences Discussions, 7*(5), 7809-7838.

Li, D., Lettenmaier, D. P., Margulis, S. A., & Andreadis, K. (2019). The value of accurate high-resolution and spatially continuous snow information to streamflow forecasts. *Journal of Hydrometeorology*(2019).

Li, H., Luo, L., Wood, E. F., & Schaake, J. (2009). The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *Journal of Geophysical Research: Atmospheres, 114*(D4).

Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D., & Reichle, R. (2012). Soil moisture, snow, and seasonal streamflow forecasts in the United States. *Journal of Hydrometeorology, 13*(1), 189-203.

Mahanama, S. P., Koster, R. D., Reichle, R. H., & Zubair, L. (2008). The role of soil moisture initialization in subseasonal and seasonal streamflow prediction–A case study in Sri Lanka. *Advances in Water Resources, 31*(10), 1333-1343.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernandez-Prieto, D., . . . Verhoest, N. E. C. (2017). GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development, 10*(5), 1903-1925. Retrieved from <Go to ISI>://WOS:000401485700001. doi:10.5194/gmd-10-1903-2017

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

Milly, P. C., Kam, J., & Dunne, K. A. (2018). On the sensitivity of annual streamflow to air temperature. *Water Resources Research, 54*(4), 2624-2641.

Mitra, S., Srivastava, P., & Singh, S. (2016). Effect of irrigation pumpage during drought on karst aquifer systems in highly agricultural watersheds: example of the Apalachicola-Chattahoochee-Flint river basin, southeastern USA. *Hydrogeology Journal, 24*(6), 1565-1582. doi:10.1007/s10040-016-1414-y

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE, 50*(3), 885-900.

Naabil, E., Lamptey, B., Arnault, J., Olufayo, A., & Kunstmann, H. (2017). Water resources management using the WRF-Hydro modelling system: Case-study of the Tono dam in West Africa. *Journal of Hydrology: Regional Studies, 12*, 196-209.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology, 10*(3), 282-290.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., & Williams, J. R. (2011). *Soil and water assessment tool theoretical documentation version 2009*. Retrieved from

Nicolai-Shaw, N., Gudmundsson, L., Hirschi, M., & Seneviratne, S. I. (2016). Long-term predictability of soil moisture dynamics at the global scale: Persistence versus large-scale drivers. *Geophysical Research Letters, 43*(16), 8554-8562. doi:10.1002/2016gl069847

Niu, G. Y., Yang, Z. L., Dickinson, R. E., & Gulden, L. E. (2005). A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models. *Journal of Geophysical Research: Atmospheres, 110*(D21).

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., . . . Xia, Y. L. (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, 116*. Retrieved from <Go to ISI>://WOS:000292111500002. doi:Artn D1210910.1029/2010jd015139

Norton, H. (2018). *Soil Moisture Memory (Smm) of Karst and Non-karst Soils.* George Mason University,

Orth, R., & Seneviratne, S. I. (2012). Analysis of soil moisture memory from observations in Europe. *Journal of Geophysical Research-Atmospheres, 117*. Retrieved from <Go to ISI>://WOS:000307753000002. doi:Artn D1511510.1029/2011jd017366

Orth, R., & Seneviratne, S. I. (2013). Predictability of soil moisture and streamflow on subseasonal timescales: A case study. *J. Geophys. Res., 118*(19), 10,963-910,979. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/jgrd.50846/full. doi:10.1002/jgrd.50846

Orth, R., & Seneviratne, S. I. (2013). Propagation of soil moisture memory to streamflow and evapotranspiration in Europe. *Hydrology and Earth System Sciences, 17*(10), 3895-3911. Retrieved from http://www.hydrol-earth-syst-sci.net/17/3895/2013/. doi:10.5194/hess-17-3895-2013

Priest, S. (2004). *Evaluation of ground-water contribution to streamflow in coastal Georgia and adjacent parts of Florida and South Carolina*: US Department of the Interior, US Geological Survey.

Rodell, M., Houser, P., Berg, A., & Famiglietti, J. (2005). Evaluation of 10 methods for initializing a land surface model. *Journal of Hydrometeorology, 6*(2), 146-155.

Schaefer, G. L., Cosh, M. H., & Jackson, T. J. (2007). The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *Journal of Atmospheric and Oceanic Technology, 24*(12), 2073-2077. Retrieved from <Go to ISI>://WOS:000252001300007. doi:10.1175/2007jtecha930.1

Schaefer, G. L., & Paetzold, R. F. (2001). SNOTEL (SNOwpack TELemetry) and SCAN (Soil Climate Analysis Network). Automated Weather Stations for Applications in Agriculture and Water Resources Management: Current Use and Future Perspectives. *K. G. Hubbard and M. V. K. Sivakumar, Eds., AGM-3, WMO/TD 1074, WMO, 187–194*.

Schilling, K. E. (2016). Comment on "Climate and agricultural land use change impacts on streamflow in the upper midwestern United States" by Satish C. Gupta et al. *Water Resources Research, 52*(7), 5694-5696.

Senatore, A., Mendicino, G., Gochis, D. J., Yu, W., Yates, D. N., & Kunstmann, H. (2015). Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales. *Journal of Advances in Modeling Earth Systems, 7*(4), 1693-1715.

Shukla, S., Sheffield, J., Wood, E. F., & Lettenmaier, D. P. (2013). On the sources of global land surface hydrologic predictability. *Hydrology & Earth System Sciences, 17*(7).

Singh, S., Mitra, S., Srivastava, P., Abebe, A., & Torak, L. (2017). Evaluation of water-use policies for baseflow recovery during droughts in an agricultural intensive karst watershed: Case study of the lower Apalachicola–Chattahoochee–Flint River Basin, southeastern United States. *Hydrological Processes, 31*(21), 3628-3644.

US Army Corps of Engineers, M. D. (2014). Environmental Impact Assessment, Update of the Water Control Manuals for the Alabama-Coosa-Tallapoosa River Basin in Georgia and Alabama.

Vinnikov, K. Y., Robock, A., Speranskaya, N. A., & Schlosser, A. (1996). Scales of temporal and spatial variability of midlatitude soil moisture. *Journal of Geophysical Research-Atmospheres, 101*(D3), 7163-7174. Retrieved from <Go to ISI>://WOS:A1996UB98600040.

Wanders, N., Thober, S., Kumar, R., Pan, M., Sheffield, J., Samaniego, L., & Wood, E. F. (2019). Development and Evaluation of a Pan-European Multimodel Seasonal Hydrological Forecasting System. *Journal of Hydrometeorology, 20*(1), 99-115.

Wood, A. W., Hopson, T., Newman, A., Brekke, L., Arnold, J., Clark, M., . . . Clark, M. (2016). Quantifying Streamflow Forecast Skill Elasticity to Initial Condition and Climate Prediction Skill. *http://dx.doi.org/10.1175/JHM-D-14-0213.1, 17*(2), 651-668. Retrieved from http://journals.ametsoc.org/doi/10.1175/JHM-D-14-0213.1. doi:10.1175/JHM-D-14-0213.1

Wu, W., Geller, M. A., & Dickinson, R. E. (2002). The Response of Soil Moisture to Long-Term Variability of Precipitation. *Journal of Hydrometeorology, 3*(5), 604-613. Retrieved from http://journals.ametsoc.org/doi/abs/10.1175/15257541%282002%29003%3C0604%3ATROSMT%3E2.0.CO%3B2. doi:10.1175/1525-7541(2002)003<0604:TROSMT>2.0.CO;2

Xia, Y., Ek, M. B., Wu, Y., Ford, T., & Quiring, S. M. (2015). Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part I: Comparison and analysis. *Journal of Hydrometeorology, 16*(5), 1962-1980.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., . . . Mocko, D. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres, 117*(D3), n/a-n/a. doi:10.1029/2011jd016048

Xiao, M., Udall, B., & Lettenmaier, D. P. (2018). On the causes of declining Colorado River streamflows. *Water Resources Research, 54*(9), 6739-6756.

Yossef, N. C., Winsemius, H., Weerts, A., Beek, R., & Bierkens, M. F. (2013). Skill of a global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resources Research, 49*(8), 4687-4699.

Yucel, I., Onen, A., Yilmaz, K., & Gochis, D. (2015). Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall. *Journal of Hydrology, 523*, 49-66.

Zhang, Y.-K., & Schilling, K. (2006). Increasing streamflow and baseflow in Mississippi River since the 1940 s: Effect of land use change. *Journal of Hydrology, 324*(1-4), 412-422.

Zheng, D., van der Velde, R., Su, Z., Wang, X., Wen, J., Booij, M. J., . . . Chen, Y. (2015). Augmentations to the Noah model physics for application to the Yellow River source area. Part I: Soil water flow. *Journal of Hydrometeorology, 16*(6), 2659-2676.

Appendix 1
Model set up and calibration

In general, WPS delineates research domain. Then run GIS-PT to prepare original set up files. Next step, its drainage density will be used in SWAT. SWAT River network, USGS gage, subwatershed, flow accumulation, and flow direction generated in will replace files generated in GIS-PT. This set up method is suitable for this project. Following is the reason and outline.

Research domain and mesh are defined by WPS in geo_em.d0X.nc file. X represents domain number. It includes mesh, geographic and projection information, and land use categories. They will be used to define coarse & fine meshes, and land use raster in WRF-Hydro preprocessing tool. Usually, the next step is running WRF-Hydro GIS preprocessing tool (GIS-PT) to get all set up files prepared. Beside it, some aspects are also set up in SWAT model. Model river network, drainage density, drainage area, and lake grids are important for model accuracy. In GIS-PT, the number of routing grid cells to define stream is 200 (default value).

For gages No.1, No.2, No.5, and No.9 (see Figure10 for subwatershed sequence number), compared with observation, simulated streamflow contains large biases (figure not shown). For No.14 subwatershed and watershed outlet, the results look good. In SWAT model, true river network could be burned in. Then, generates DEM-based flow direction and accumulation. The common threshold area to define river is watershed area/100, but it lose too many river branches. If use watershed area*3/1000, the simulation results are good enough except No.14 subwatershed and watershed outlet. Summarize the above phenomenon; the GIS-PT drainage density is good to generate accurate outlet streamflow. River network generated in SWAT is accurate for subwatershed with small drainage area. At the same time, the biases mentioned above from GIS-PT and SWAT cannot be diminished by any way of calibration. So combining the two method together is one possible way. Specifically, we will use GIS-PT drainage density, to finish all other

set up process in SWAT. That is to say adjusting threshold area to define stream network in SWAT to match the drainage density got in GIS-PT. Even the river network in SWAT now does not match GIS-PT network absolutely, the model can generate reasonable and acceptable streamflow from upstream to downstream.

When adjusting parameters one by one, evaluate model performance in each of the 15 gages, watershed outlet streamflow is decisive. Based on the river network, USGS gage should be nudged to the nearest reach to define subwatershed. Overall, river network, flow accumulation, flow direction, USGS gage are generated in SWAT. Terrain routing mesh and other parameters are generated in GIS-PT. Stream order is generated by ArcGIS Hydrology toolbox stream order tool. Model will run with the mask of subwatershed delineated by 15 USGS streamflow gages (see Fig 10 for mask delineated by subwatershed). Finally, define lake area threshold before input lake shapefile into GIS-PT. Because too many lakes decrease watershed outlet streamflow, especially when large area lakes surround a streamflow gage, streamflow will approach zero. The outline of model set up is described above. Here is specific step by step set up method but not in extreme detail. Please refer to tools (WPS, GIS-PT, GIS, SWAT) manual.

Set up in WPS system

WPS will delineate research domain and the course mesh (LSM mesh). Detailed information is in WRF ARW User's Guide and WRF ARW Online Tutorial. There are many information in the output file. Only mesh, geographic projection information, and land use categories will be used in the following workflow. Users should pay special attention to land use categories, USGS and MODIS-20 are supported. Because WRF-Hydro parameter files MPTABLE.TBL does not contain MODIS-21 parameters. In the WPS preparing process, if

MODIS-21 is used, model maybe works well sometimes, but it uses MODIS-20 parameter actually.

Set up in GIS-PT

1. The first step is running WRF-Hydro GIS preprocessing tool (GIS-PT) to get all set up files prepared. USGS gage XY Geographic Coordinate Systems is NAD27. Z Coordinate Systems is national Geodetic Vertical Datum 1929 (NGVD29). Refer to material named Downloading USGS Streamflow Data from Internet into ArcGIS. It is necessary to make sure the USGS gages are projected into the right position under WPS coordinate system. Since GIS-PT tool will match output file with projected coordinate system in WPS. After this process, gage longitude and latitude could be used in GIS-PT.
2. Select “mask CHANNELGRID variable to forecast basins” and “create lake parameter (LAKEPARM) file”. Input reservoirs shapefile. The rule of thumb of lake area threshold is 0.75~1.0 km. Too many lakes will decrease streamflow apparently. Define regridding factor and adjusting threshold grid cells for stream definition to acquire an acceptable watershed outlet streamflow. This drainage density will be used in SWAT. So keep the river network shapefile.

Set up in SWAT

1. SWAT should be run twice. The first time is setting river network drainage density a little bit greater than the density generated in GIS-PT. Make ensure all SWAT input files projection systems are the same, including DEM setup file and mask. DEM resolution is 30-meters, so it should be clipped into research domain (mask). Then input GIS-PT river network shapefile into SWAT.

2. Besides the usual way to set up watershed in SWAT, the NHD plus true river network could be burn in SWAT. Select DEM-based stream definition. Adjust threshold area to define stream, and make sure the drainage density greater than GIS-PT drainage density. Then generate flow direction and accumulation. Stop here, save river network, it will be used later.
3. Repeat step1 &2 processes but use DEM generated in GIS-PT. GIS-PT DEM resolution is terrain routing grid resolution (360*360 for this project). The burn in river network generated in former step (step 2). Adjust threshold area to define stream, and make sure the generated drainage density correspondent with GIS-PT drainage density. Then click flow direction and accumulation button, and create streams and outlets. Delete outlet generated in this step, add subwatershed outlet manually. Specifically, input GIS-PT USGS gage shapefile, add the gage point to river network generated in this step. The gage point should be nudged into the nearest location of gage point in GIS-PT.
4. Delineate subbasin watershed and whole watershed outlets. Keep all output files. Pay attention that output files projection systems are the same as GIS-PT/ WPS.

Set up in GIS-PT

1. Run GIS-PT again from the very beginning, create mask but do not create lake parameter this time.
2. Open river network and flow direction generated in SWAT to calculate new stream order. The ArcGIS hydrology stream order tool will do this.
3. Write flow direction, flow accumulation, stream network, gage points, subwatershed area in SWAT and stream order generated in former step (step 2) to replace the variables from

step 1. The subbasin series number should correspond with outlet sequence, gage location should be within subwatershed.

4. Use GIS-PT utilities " Build Groundwater Inputs" to generate subbasin parameter file since Fulldom_hires.nc basin mask and gage variables are replaced in former step (step 3).
5. Zip all files, input it into GIS-PT utilities "Add Lake Parameters", to form lake parameters and mask the river network raster file when lake overlaps channel. That is the reason why we should set up lake very carefully.

Several trouble shootings

1. After several times running, there are fragment files in the GIS-PT directory. They will lead to unexpected failure. So zipped out original GIS-PT package and replace the one currently used from time to time.
2. GIS-PT contains Python projection function, if the ArcGIS installation folder is not enough (less than 100GB), the tool will show 999999 error without any other hint. Pay attention that GIS-PT tool corresponding ArcGIS version and ArcSWAT corresponding ArcGIS version. These two should be the same.
3. 'Fulldom_hires.nc' file latitude and longitude maybe empty or filled with missing value sometimes. This error can not be found in "Examine Outputs of GIS Preprocessor" tool. So check them in the end manually.
4. 'GWBASINS.nc data' may not be correct sometimes. So check it, and rerun "Build Groundwater Inputs" to fix the problem.

5. Do not use reprojected DEM to run GIS-PT. Sometimes, reproject will lead to unexpected error. It means do not reproject DEM to match coordinate system defined in WPS, since GIS-PT will project data when it is necessary.
6. Land cover classification is USGS 24 type land cover, since WRF-Hydro parameter file doesn't contain MODIS-21 land cover parameters. If the later one is used, unless any of the missing 3 types is accessed by WRF-Hydro, it will show error information, otherwise the default is hard to find.

Model calibration

In general, we use measured parameter firstly, then manually adjust several sensitive parameters by repeatedly trial and error. First, adjust parameters in CHANPARM.TBL since measured values are available. Based on it, Manning's Roughness coefficients for different types of channel are available in Open-Channel Hydraulics (Chow 1988). The stream order 1 to 4 channels in this study are main channels, its roughness coefficient is around 0.03. Since simulation is warm start, default initial depth of water in the channe l is enough. Channel bottom width and channel side slope could be found from DEM (30-m resolution). Figure 1 is the DEM cross section used to find bottom width for each stream order. Combine them and channel surface width, we get channel slope parameters. Stream order 1 cross section figure is similar with stream order 2 figures. In ArcGIS, selected several representative points for each stream order, get cross section profile graph, then calculate side slope and bottom width directly. Use each stream order average value, the value varies for different cross section. This way is referred to the method in Merwade and Coonrod (2008). Calibrating reservoir parameters in LAKEPARM should refer to US Army Corps of Engineers water control manuals (like http://water.sam.usace.army.mil/rfhenry.pdf). They show

reservoir basic parameters and figures, In the manual, FLOOD RISK MANAGEMENT SLUICE parameters are used to estimate orifice cross section and elevation. Combine with dam photo, if manual doesn't contain sluice parameters, default value will be used. Weir coefficient is from Broad-Crested Weir Coefficients table (Brater et al. 1976). So weir coefficient, length, elevation, and orifice coefficient, area, elevation are available. WRF-Hydro model uses one orifice one broad-crest weir model to represent reservoir, and simulation does not contain manual operation. Approximate calculation is necessary to adjust real world reservoir parameters to fit into physical model (Broad-Crested Weirs with one orifice). Then maximum lake elevation should not be greater too much than weir elevation. Now, the corrected channel and lake parameters lead to better control of hydrograph shape and peak flow.

Most of calibration references show similar sensitive parameters and calibration sequence. Baseflow and flow volume are sensitive to saturated soil conductivity (SATDK). To increase baseflow, SATDK is increased little by little. Then increase groundwater bucket model coefficient to increase baseflow. Other sensitive parameters include saturated hydraulic conductivity parameterization (REFDK), saturated hydraulic infiltration parameterization (REFKDT). Since the model does not compile with spatial distributed soil property, they remain default for this project. Some parameters do not make too much difference. They are surface retention depth (RETDEPRTFAC), saturated soil lateral conductivity (LKSATFAC), bucket model zmax, exponent, and zinit.

Reservoir model is Broad-Crested Weirs with one orifice under the weir.

Broad-Crested Weirs discharge equation is

$$Q = CLH^{1.5}$$

C = Broad-crested weir coefficient (values between 2.6 and 3.1 are normal)

L = Broad-crested weir length. (Unit: ft)

H = Head above weir crest. (Unit: ft)

Orifices discharge equation is

$$Q = CA(2gH)^{0.5}$$

C = Orifice coefficient (a value of 0.6 is usually appropriate)

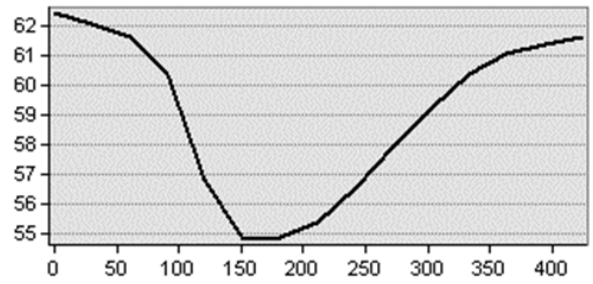
A = Area of orifice, square feet (ft^2)

g = Acceleration due to gravity, 32.2ft/sec^2

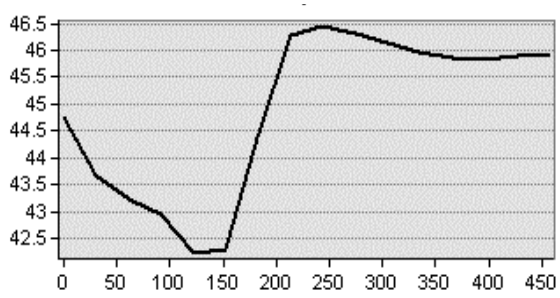
H = Head above orifice centroid, ft

Table 1. Calibrated lake parameters in NWM

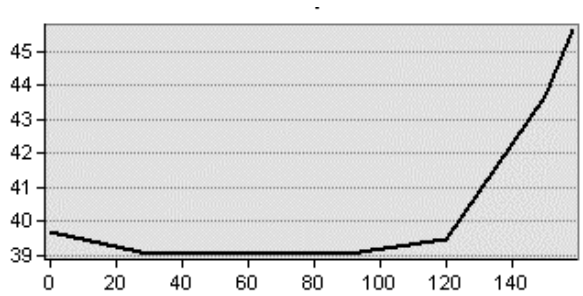
Parameter	Long name
LkMxE	Maximum lake elevation (m ASL) (above see level)
WeirC	Weir coefficient
WeirL	Weir length (m)
WeirE	Weir elevation (m ASL)
OrificeC	Orifice coefficient
OrificeA	Orifice cross-sectional area (sq. m)
OrificeE	Orifice elevation (m ASL)

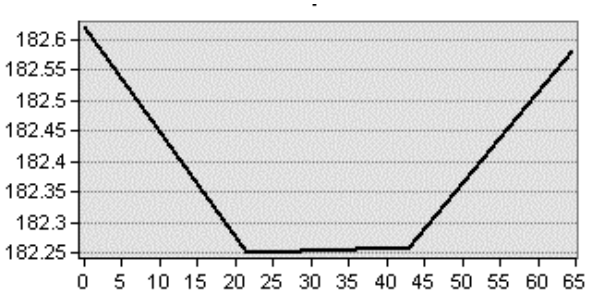


(a) stream order 2 profile graph

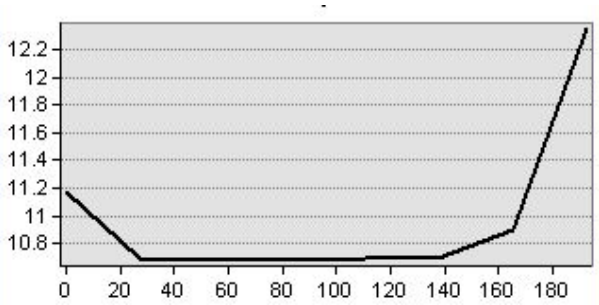


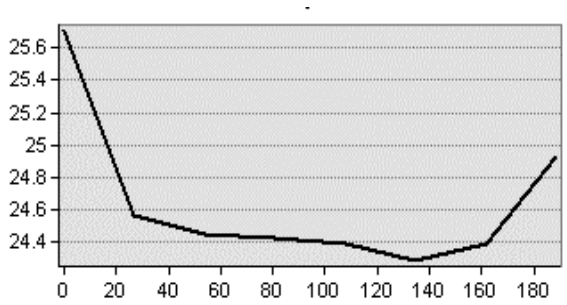
(b) stream order 3 profile graph



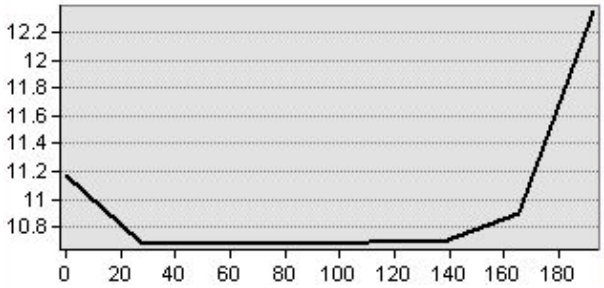


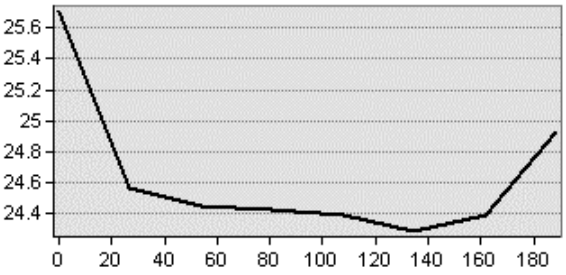
(c) stream order 3 profile graph





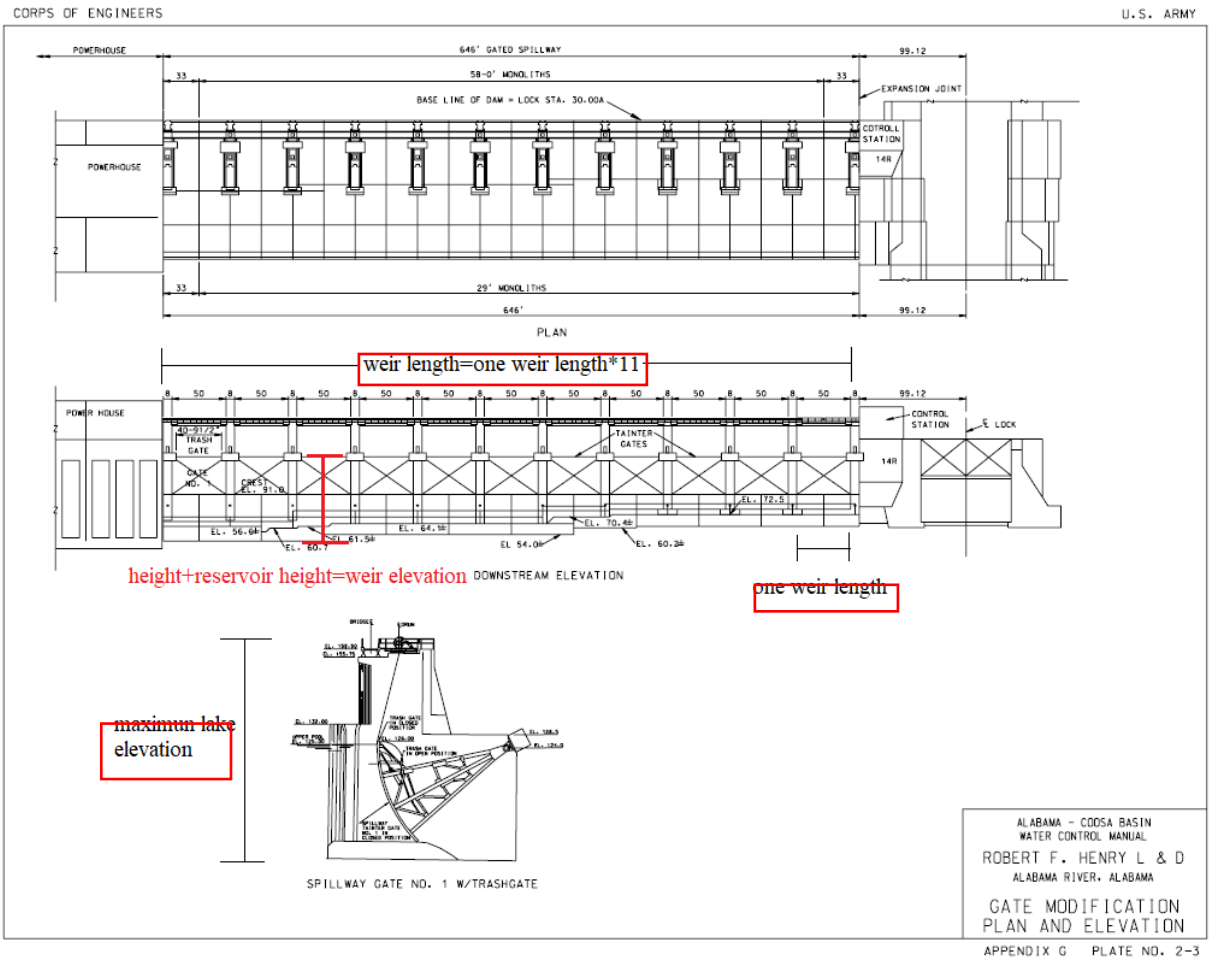
(d) stream order 4 profile graph





(e) stream order 5 profile graph

Figure 1 DEM cross section profile graph



Robert F. Henry L &D dam manual figure and parameters

Appendix 2

Noah-MP Soil Moisture and Runoff Parametrizations

The default total volumetric soil moisture is 60% of saturation. Then model will use simulated soil moisture instead of default value as model initial field for following calculation. Following is soil moisture calculation process and its role in runoff simulation. Run Noah-MP land surface model, the input variables are forcing and land states like vegetation, land use fraction, top soil texture categories, green vegetation fraction, and land surface albedo in each grid cell and so on.

Noah-MP soil moisture flux in soil layers

1.1 The soil water flux in the first layer

$$q = \lambda_1 \times z_1 + k_1 - I_{sfc} + ET_1 + E_{sfc} \text{ (top layer)}$$

$$q = \lambda_{i-1} \times z_{i-1} + k_{i-1} + ET_i + Q_{drain} \text{ (bottom layer)}$$

$$q = \lambda_i \times z_i + k_i - (\lambda_{i-1} \times z_{i-1} + k_{i-1}) + ET_i \text{ (layers in between)}$$

1. λ is soil water diffusivity, $\lambda = [\lambda_{sat} \times \left[\max\left(0.01, \frac{\theta^n}{\theta_e}\right)\right]^{B+2}] \times (1 - f_{cr})$. When consider soil moisture

$$\lambda = \left[\frac{1}{1 + (500 \times \theta_{soil-ice})^3}\right] \times \left[\lambda_{sat} \times \left[\max\left(0.01, \frac{\theta^n}{\theta_e}\right)\right]^{B+2}\right] + \left[1 - \frac{1}{1 + (500 \times \theta_{soil-ice})^3}\right] \times \lambda_{sat} \times \left(\frac{2}{\theta_c}\right)^{2+B}$$

θ_{soil} is volumetric soil water content. θ_e is soil effective porosity $\theta_e = \theta_{c-}\theta_{soil-ice}$. θ_c is the soil total porosity.

2. $z_i = 2 \times (\theta_i - \theta_{i+1})/[(Z_b)_{i+1} - (Z_b)_{i-1}]$, represents soil moisture content in unit soil depth. Z is the height above some datum.

3. K is hydrological conductivity $K = [K_{sat} \times \left[\max\left(0.01, \frac{\theta^n}{\theta_e}\right)\right]^{2\times B+3}] \times (1 - f_{cr})$

K_{sat} is saturated hydrological conductivity, $K_{sat} = \frac{J}{I}$ J is flux (unit: m/s), I is hydraulic gradient (unit: m^3/m^3). f_{cr} impermeable fraction due to frozen soil.

4. I_{sfc} is surface infiltration rate, ET is transpiration rate. Q_{drain} is bottom drainage. $E_{sfc}$is the soil surface evaporation rate.

1.2 Depth of water table

$$\Delta z = \frac{1}{\theta_e}\left[\frac{Q_{net}}{A} - R_{rc}\right] \times \Delta t$$

Δz is water table change, θ_e is effective soil porosity. Q_{net} is net lateral flow of saturated subsurface moisture. A is area, R_{rc} is soil column recharge rate from injection.

Noah-MP SIMGM water table equation is

$$z_\nabla = (-z_4 + 25) - \frac{W_a}{1000 \times 0.2}\,(z_\nabla < z_4)$$

$$z_\nabla = (-z_4) - \frac{W_a - 1000 \times 0.2 \times 25}{1000 \times \theta_{e4}}\,(z_\nabla\ within\ z_4)$$

1. Z is the height above some datum.
2. W_a is aquifer water storage. $\Delta W_a = Q - R_{sb}$ Q is aquifer recharge. R_{sb} is subsurface runoff rate.
3.

$$Q = -K_{bot}\frac{-z_\nabla - (f_{mic}\psi_{bot} - z_{bot})}{z_\nabla - z_{bot}}$$

Where bottom refers to the bottom soil layer adjacent to acquifer. ψ_{bot} is bottom layer matric potential. f_{mic} is the minimum fraction of micropore content in bottom layer.

4. θ_{e4} is the forth layer effective soil porosity.

1.3 Infiltration

$$I_{max} = \frac{\left[P_x \times \left[\frac{d_{tot} \times (1 - e^{(-K_{dt} \times \Delta t_1)})}{P_x + d_{tot} \times (1 - e^{(-K_{dt} \times \Delta t_1)})}\right]\right]}{\Delta t}$$

I_{max} is maximum infiltration rate, the boundary is I_{max}.

1. $P_x = \max(0,\ Q_{wat} \times \Delta t)$ Q_{wat} is surface input water.
2. d_{tot}=$\sum_{i=1}^{nsoil}(d_{max})_i$

$(d_{max})_i$=$(((Z_b)_{i-1} - (Z_b)_i) \times \theta_{c-wp}) \times (1 - ((\theta_{soil})_i + (\theta_{soil-ice})_i - \theta_{wp})/\theta_{c-wp})$ Z_b is the depth of soil layer bottom; $\theta_{c-wp} = \theta_c - \theta_{wp}$, θ_{wp} is permanent wilting point porosity. $(\theta_{soil})_i + (\theta_{soil-ice})_i - \theta_{wp}$ represents the available soil water ratio (m^3/m^3) in this layer. Similarly, d_{ice}=$\sum_{i=1}^{nsoil}[d_{ice} - ((Z_b)_{i-1} - (Z_b)_i) \times (\theta_{soil-ice})_i]$

3. Δt_1 is the time interval to calculate soil infiltration. K_{dt} is hydraulic conductivity.

Infiltration boundary condition is

$$I_{max} = I_{max} \times f_{cr},\ I_{max} = \max(I_{max}, K),\ I_{min} = \min(I_{max} \times P_x)$$

The surface runoff and infiltration relation is $R_s = \max(0, Q_{wat} - I_{max})$

1.4 Soil layers accumulated saturation

The soil water infiltration subroutine calculates surface infiltration excess based on top soil layer water and precipitation. Ponded water evaporation subroutine generates ponding water evaporation. Infiltration excess and ponded water evaporation results are delivered to channel and reservoir routing section. In this research area (Southeastern US.) snow melt is negligible. So the soil layers accumulated saturation excess is

$$\theta_{total} = \int_{i=1}^{nsoil} (\theta_{soil} - \theta_e)\, dD_{soil}$$

θ_{soil} is volumetric soil water content. θ_e is soil effective porosity $\theta_e = \theta_{c-}\theta_{soil-ice}$. θ_c is the soil total porosity. The boundary condition is $\theta_{soil} = min[\theta_e, \theta_{soil}]$, $\theta_{soil} = max[\theta_{soil}, \theta_{soil-min}]$. The parameters used in Noah-MP is soil categories properties involve soil thermal diffusivity, maximum volumetric soil moisture, saturated soil hydraulic conductivity and so on.

Interaction with Runoff

In higher infiltration rate zone, Hortonian overland flow rarely happens except heavy rainfall, subsurface flow is the primary mechanism in rainfall runoff process. The total subsurface flow comprises two parts. One is groundwater flow in saturated zone, another is unsaturated subsurface flow seeping from the hillside above the water table. Both of them depend on soil moisture saturation condition. In routing processes i.e., overland flow and subsurface flow routing, model outputs are subsurface flow, overland flow, depth water table, baseflow and so on. In subsurface flow routing, soil moisture will be updated.

1.1 Subsurface runoff rate

$$R_{sb} = (1 - f_{cr-max}) \times f_{base} \times e^{-\wedge - K_{runoff} \times (z\nabla - z_{bot})}$$

1. f_{cr-max} is the maximum impermeable fraction due to the frozen soil. Impermeable fraction due to frozen soil is

$$f_{cr} = \frac{\max(0.0, e^{(-A \times (1 - f_{ice}))} - e^{(-A)})}{(1 - e^{(-A)})}$$

A=4, f_{ice} is the ice fraction in soil $f_{ice} = min\left[1.0, \frac{\theta_{soil-ice}}{\theta_c}\right]$, $\theta_{soil-ice}$ is the ratio between ice volume and total soil volume. θ_c is the soil porosity. The boundary soil ice and liquid are defined as following

$$\theta_{ice-max} = max[\theta_{soil-max}, \theta_{soil-ice}],\ f_{cr} = min[f_{cr}, f_{cr-max}],$$

2. f_{base} is base flow coefficient, $f_{base} = 4.\ (1 - f_{cr-max}) \times f_{base}$ is the maximum subsurface runoff rate, global reference value is 5.0×10^{-4} mm/s.

3. $\wedge$ is grid cell mean wetness index. $Wetness\ Index = \ln(\alpha / tan\beta)$, α is grid catchment area, β is local

slope.

4. K_{runoff} is runoff decay factor, $K_{runoff} = 2.0$. The global value is 6.

5. $z\nabla - z_{bot}$ is water table depth above bottom.

1.2 Overland flow (surface runoff)

$$\mathrm{R_s} = Q_{wat} \times [(1 - f_{cr}) \times f_{sat} + f_{cr}]$$

1. Q_{wat} is surface input water.

2. f_{sat} is fractional saturated area.

$$f_{sat} = (1 - f_{cr}) \times f_{max} \times e^{\left(-0.5 \times K_{runoff} \times (z\nabla - z_{bot})\right)} + f_{cr}$$

f_{max} is the maximum saturated area, it means the sum of areas where water table depth is zero. Other parameters are the same as former section. Another f_{sat} definition is

$$f_{sat} = max\left[0.01, \frac{\theta_{2m-ave}}{\theta_c}\right]^4$$

θ_{2m-ave} is 2-m averaged soil moisture, $\theta_{2m-ave} = \frac{\int_{i=1}^{nsoil}(\theta^N)dD_{soil}}{D_{soil}}$

If water content is less than theoretical infiltrated water, the saturation excess should be accumulated in each iteration, surface runoff equals to saturation excess in each step.

Soil moisture updated by subsurface flow routing

When runoff removes soil moisture R_{rm}

$$\mathrm{R}_{rm} = R_s \times \Delta t \times \frac{K_i \times D_{soil-i}}{\int_{i=1}^{nsoil} K_i \times d(D_{soil-i})}$$

R_s is overland flow. K is hydrological conductivity. D is soil depth. So soil moisture variability depends on removed soil moisture:

$$\Delta\theta_{soil} = \frac{R_{rm}}{D_{soil}}$$

The total soil moisture mass is

$$m_{soil} = \theta_{soil} \times D_{soil}$$

Subsurface runoff rate equals with soil moisture variability rate

$$\Delta R_{sb} = \frac{\Delta m_{soil}}{D_{soil} \times \Delta t}$$

Soil moisture content variablis equals with subsurface runoff

$$\Delta\theta_{soil} = \frac{\Delta m_{soil}}{D_{soil}}$$

For each layer, saturation soil water content change is

$$\Delta(\theta_{\mathrm{sat}})_{\mathrm{i}} = \frac{-\lambda_i \times 2}{(-Z_b)_{i+1} \times (-Z_b)_i{}^2} \times \Delta t \quad (1 < \mathrm{i} < \mathrm{nsoil})$$

$$\Delta(\theta_{\mathrm{sat}})_1 = \frac{\lambda_1 \times 2}{(-Z_b)_{i+1} \times (-Z_b)_i} \times \Delta t \quad (\mathrm{i} = 1)$$

$$\Delta(\theta_{\mathrm{sat}})_{\mathrm{n}} = 0 \quad (\mathrm{i} = \mathrm{n})$$

References

Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E. and Tewari, M., 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, 116(D12).

Yang, Z.L., Cai, X., Zhang, G., Tavakoly A. A., Jin Q., Meyer H. L., Guan X., 2011. The community Noah land surface model with multi parameterization options (Noah-MP): Technical Description.

Niu, G.Y., Yang, Z.L., 2011. The community Noah land-surface model (LSM) with multi-physics options: User's Guide.

Gochis, D., J., Chen, F., Hydrological Enhancements to the community Noah land surface model: NCAR Scientific Technical Report.