Developing a hydrological monitoring and sub-seasonal to seasonal forecasting system for South and Southeast Asian river basins

Yifan Zhou[1], Benjamin F. Zaitchik[1], Sujay V. Kumar[2], Kristi R. Arsenault[2,3], Mir A. Matin[4], Faisal M. Qamer[4], Ryan A. Zamora[1], and Kiran Shakya[4]

[1]Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland, USA
[2]Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
[3]Science Applications International Corporation, Reston, Virginia, USA
[4]International Centre for Integrated Mountain Development, Kathmandu, Nepal

Correspondence: Benjamin F. Zaitchik (zaitchik@jhu.edu)

Abstract. South and Southeast Asia is subject to significant hydrometeorological extremes, including drought. Under rising temperatures, growing populations, and an apparent weakening of the South Asian monsoon in recent decades, concerns regarding drought and its potential impacts on water and food security are on the rise. Reliable sub-seasonal to seasonal (S2S) hydrological forecasts could, in principle, help governments and international organizations to better assess risk and act in the face of an oncoming drought. Here, we leverage recent improvements in S2S meteorological forecasts and the growing power of Earth observations to provide more accurate monitoring of hydrological states for forecast initialization. Information from both sources is merged in a South and Southeast Asia sub-seasonal to seasonal hydrological forecasting system (SAHFS-S2S), developed collaboratively with the NASA SERVIR program and end users across the region. This system applies the Noah-Multiparameterization (NoahMP) Land Surface Model (LSM) in the NASA Land Information System (LIS), driven by downscaled meteorological fields from the Global Data Assimilation System (GDAS) and Climate Hazards InfraRed Precipitation products (CHIRP and CHIRPS) to optimize initial conditions. The NASA Goddard Earth Observing System Model sub-seasonal to seasonal (GEOS-S2S) forecasts, downscaled using the National Center for Atmospheric Research (NCAR) General Analog Regression Downscaling (GARD) tool and quantile mapping, are then applied to drive 5 km resolution hydrological forecasts to a 9-month forecast time horizon.

Results show that the skillful predictions of root zone soil moisture can be made 1 to 2 months in advance for forecasts initialized in rainy seasons and up to 8 months when initialized in dry seasons. The memory of accurate initial conditions can positively contribute to forecast skills throughout the entire 9-month prediction period in areas with limited precipitation. This SAHFS-S2S has been operationalized at the International Centre for Integrated Mountain Development (ICIMOD) to support drought monitoring and warning needs in the region.

1 Introduction

South and Southeast Asia is one of the most populated areas in the world, and a significant portion of livelihoods depend directly or indirectly on smallholder agriculture. Agriculture is one of the most weather-dependent human activities (Hatfield et al., 2011), and smallholder systems are particularly vulnerable to weather variability, including extreme events such as drought. South and Southeast Asia has been experiencing anthropogenic warming since the 1950s (Sivakumar and Stefanski, 2010), and the warming is projected to continue in the near future (Barros and Field, 2014). The frequency of extreme weather events, including droughts, has also been increasing under this warming trend, with implications for food security and social stability in a conflict-prone region that already includes extensive marginal agriculture on semi-arid lands (Samaniego et al., 2018).

Sub-seasonal-to-seasonal hydrological forecast systems (S2S-HFS) have the potential to aid preparedness for these extreme events. Such systems have been implemented all over the world at the scale of large river basins (Yuan et al., 2016; Getirana et al., 2009), countries (Xia et al., 2012; Bell et al., 2017; Shah et al., 2017), continents (Wanders et al., 2019; Sheffield et al., 2014; Yuan et al., 2013) and the entire globe (Alfieri et al., 2013; Yuan et al., 2011; Wanders and Van Lanen, 2015).

An S2S-HFS generally includes three components – a meteorological forecast, a downscaling method, and a hydrological or land surface model (Yuan et al., 2015; Hao et al., 2017). The presence of a land surface model in this system means that hydrological forecasts can draw skill from both the quality of meteorological forecast and the accuracy of modeled initial hydrological states. The influence of initial hydrological states has a more substantial impact during the early prediction period, while meteorological forecasting dominates later months (Shukla and Lettenmaier, 2011). The initial condition, however, can still positively contribute to the forecast skill several months after the initialization (Samaniego et al., 2019). This contribution can come from the memory of deeper soil moisture, groundwater storage and cumulated snow pack in earlier seasons, which provides the potential for hydrological forecasts to have skills greater than meteorological S2S forecasts (Koster et al., 2010).

South and Southeast Asia presents both a challenge and an opportunity in this regard. The challenge is that S2S meteorological forecasts can be quite difficult in some areas. For example, de Andrade et al. (2019) study the skill of meteorological S2S forecasts and find that for multiple initialization dates, the forecasts lose meaningful skill within 4 weeks of initialization. Similarly, Jie et al. (2017) conclude that the operational climate models from the WCRP Seasonal to Subseasonal Prediction Project lack the skill for prediction of the Asian summer monsoon, a major precipitation source of most areas in South Asia, beyond 1 month in advance. The opportunity is that many rivers in major basins of South and Southeast Asia are sourced from mountainous regions with significant snowpack and seasonally frozen soil water, such that a properly initialized hydrological forecast has the potential to add considerable predictive power beyond the meteorological forecast.

In this study, we aim to establish an operational fine-resolution, continental hydrological monitoring and forecasting system for South and Southeast Asia: the South and Southeast Asia Hydrological Forecast System – S2S (SAHFS-S2S). This system includes downscaled meteorological fields from a dynamically based S2S meteorological forecast system, an advanced land surface model, and post-processing tools to derive drought indicators. The system aims to provide end-users in South and Southeast Asia with water resources information to help manage local drought risks and strengthen food security.

We first test the forecasting system in hindcast experiments against a retrospective run of the monitoring system, which is our best estimate of hydrological states, and available satellite-derived estimates of soil moisture, the critical variable for our drought monitoring goals. An additional control set of hindcast simulations are designed to evaluate the importance of initial conditions to the hydrological forecast skill. All settings in this set of control hindcast simulations are the same as the forecast system except that the initial conditions are set as the same climatological hydrologic states averaged from the entire hindcast period.

Section 2.1 describes the precipitation pattern in research areas and major river basins. Section 2.2 describes the dataset used in this study. Details of the monitoring and forecasting system are described in Sect. 2.3 and 2.4, respectively. The post-processing of both systems is described in Sect. 2.5. The comparisons among different simulations, the importance of the initial conditions, and validation of both systems and applications of the systems are discussed in Sect. 3. Finally, the implications, limitations, and possible future work are discussed in Sect. 4.

2 Methods

2.1 Research area and precipitation patterns

In this study, all simulations are performed on a domain within the South and Southeast Asia region, ranging from 8 to 45° N and 58 to 123° E. The analysis, however, focuses on a study domain, mainly including five major river basins and surrounding areas (Fig. 1a). The five major river basins in this South and Southeast Asia study domain are the Helmand, Indus, Ganges, Brahmaputra, and Mekong basins.

Among these five river basins, the Mekong, Brahmaputra, and Ganges have the highest average precipitation, annually: 1677, 1227, and 1108 mm, respectively. These three basins are strongly influenced by the South Asian monsoon pattern, and most precipitation falls in the summertime monsoon season (80 % in the Mekong basin, 70 % in the Brahmaputra basin, and 85 % in the Ganges basin). The monsoon seasons generally start in May, June, and July in the Mekong basin, the Brahmaputra basin, and the Ganges basin, respectively, and end in late September or early October. The Mekong and Brahmaputra basins have considerable precipitation variations between the upper and lower basins. Moisture that enters the basins from the south is naturally blocked by the Himalayan ranges and cannot reach the upper basins of these river systems. The upper Brahmaputra basin suffers from a direct rain shadow effect of the Himalayas and has precipitation amount less than one-fourth of the precipitation in the lower Brahmaputra basin (Immerzeel, 2008). Although the monsoon precipitation falls mostly in the lower Mekong basin, snow accumulated in the upper Mekong basin provides
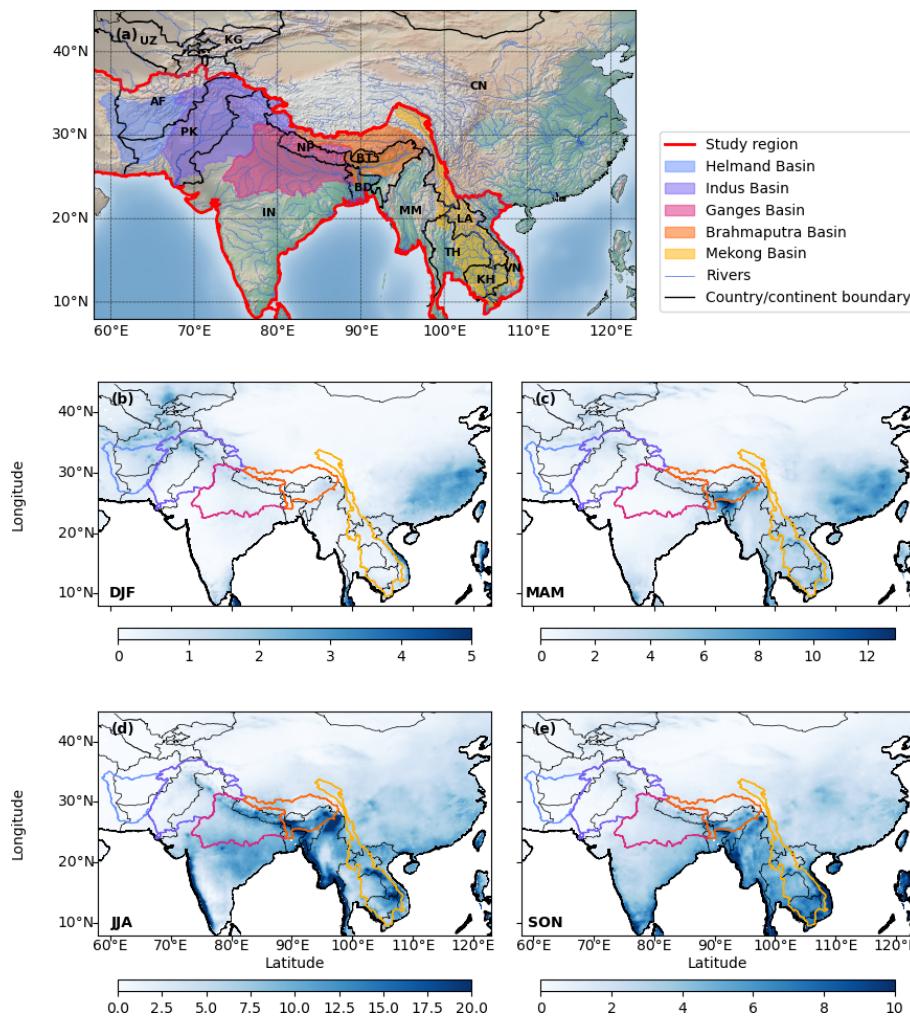
Figure 1. (a) The extent of simulation area, research domain and five major river basins in South and Southeast Asia; Seasonal precipitation climatology for 2000–2018 **(b)** December, January, and February; **(c)** March, April, and May; **(d)** June, July and August; **(e)** September, October and November. The seasonal precipitation climatology is estimated from Climate Hazards Center InfraRed Precipitation with Station (CHIRPS) data.

vital water resources for the entire basin during pre-monsoon and dry seasons (Syaukat, 2012).

In the Indus basin, the seasonal cycle is also influenced by the monsoon climate (Chen et al., 2016). Most precipitation falls during the monsoon season from July to September (199 mm), accounting for 51 % of the annual average precipitation (389 mm). The pre-monsoon and winter seasons, however, also contribute an essential part of the precipitation, while the post-monsoon period is the driest season. Despite providing less total precipitation than the summer monsoon, winter precipitation is vital for wheat and barley crops. Indus also has significant spatial precipitation variation. Most precipitation falls in the mountainous regions in the upper basin, while the lower basin lies in one of the driest areas in South Asia. Melting water from snow and glaciers in these mountain ranges provides a considerable amount of fresh water to the Indus river (Fowler and Archer, 2006).

The Helmand basin has a hyperarid to arid climate (Whitney, 2006), with an annual average precipitation of only 144 mm. The seasonal cycle of precipitation is different from the other four basins, as the maximum precipitation occurs during the winter season, which accounts for 76 % of the total precipitation. The Helmand basin also suffers from high temperatures and extreme wind which intensify the aridity in the basin.

2.2 Data

Datasets used in this research include the daily Climate Hazards Center InfraRed Precipitation data (CHIRP) and Climate Hazards Center InfraRed Precipitation with Station data (CHIRPS), National Oceanic and Atmospheric Administration (NOAA)'s Global Data Assimilation System (GDAS), NASA's Goddard Earth Observing System Model

sub-seasonal to seasonal forecast version 1 and version 2 (GEOS-S2S-V1 and V2), and the European Space Agency's Climate Change Initiative for Soil Moisture (ESA-CCI SM) (Table 1).

CHIRP is a precipitation dataset developed by the Climate Hazards Center at University of California, Santa Barbara. This dataset has a quasi-global (50° S–50° N and all longitudes) spatial coverage and is derived from satellite data (Funk et al., 2015). The data are available from 1981 to near real-time present with a 2–3 d latency. These data have 0.05° spatial resolution, and the daily precipitation product is used in this research. CHIRPS is a precipitation dataset derived by combining the CHIRP dataset with in situ station data (Funk et al., 2015). The specifications of the dataset are similar to CHIRP, except that this dataset has about 3-week latency. In this study, the CHIRP–CHIRPS dataset provides the precipitation field in the monitoring system (see Sect. 2.3 for details), and CHIRPS is used as the reference for downscaling precipitation forecasts in the forecasting system (see Sect. 2.4 for details).

GDAS is an atmospheric analysis system developed at the National Center for Environmental Prediction (NCEP) at NOAA (National Climatic Data Center, 2020). It is produced by assimilating surface observations into a Global Forecast System (GFS). The assimilated surface observation includes balloon data, wind profiler data, aircraft reports, buoy observations, radar observations, and satellite observations (https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-data-assimilation-system-gdas, last access: 17 December 2020). The GDAS dataset has a 6-hourly temporal resolution and produces short-term meteorological forecasts at 0, 3, 6 and 9 h lead time. GDAS data are available from 2000 onward and have real-time updates. The spatial resolution starts at about $1° \times 1°$ in the year 2000 and has gradually improved to an equivalent grid of $0.125° \times 0.125°$, since early 2015. In this study, seven meteorological variables from GDAS are used as forcings in the monitoring system (see Sect. 2.3 for details) and baselines for meteorological downscaling in the forecasting system (see Sect. 2.4 for details). These variables include downward long-wave radiation, downward shortwave radiation, air temperature, specific humidity, air pressure, and zonal and meridional wind speed fields.

GEOS-S2S-V1 is a meteorological forecast dataset produced by the GEOS atmospheric model at the Global Modeling and Assimilation Office (GMAO) at NASA Goddard Space Flight Center. GEOS-S2S-V1 forecasts were initialized about every 5 d in hindcast experiments and forecasts from 2000 to 2017 and produce daily forecasts for 9 months. In each month, one forecast initialization date contains 10 ensemble members while others only have one. GEOS-S2S-V1's spatial resolution is $1° \times 1.25°$.

Eight variables from daily GEOS-S2S-V1, including the same seven variables as in GDAS and precipitation rate variables (the eighth variable), are used in this research.

ESA-CCI SM is a global gridded surface soil moisture dataset derived from remote sensing observation (Gruber et al., 2019). The ESA-CCI SM contains three sub-datasets: ACTIVE dataset, PASSIVE dataset, and COMBINED dataset. The ACTIVE dataset is derived by merging satellite datasets measured by active scatterometer instruments. The ACTIVE dataset spans from August 1991 to December 2019. The PASSIVE dataset is derived by merging satellite datasets measured by passive radiometer instruments. The PASSIVE dataset spans from November 1978 to December 2019. The COMBINED dataset is merged and rescaled from the ACTIVE and the PASSIVE dataset. All three sub-datasets have a daily temporal resolution and 0.25° spatial resolution. The COMBINED dataset is used in this study to evaluate the hydrological monitoring and forecasting system.

2.3 Monitoring system

The monitoring system is an instance of a land data assimilation system (LDAS), which is a technique that merges observations with physically based models to produce optimal estimates of terrestrial hydrological states and fluxes (Rodell et al., 2004; Mitchell et al., 2004). For the purposes of monitoring hydrological extremes, our process consists of three steps (Fig. 2): (1) meteorological data processing, (2) land surface model simulations, and (3) post-processing to obtain relevant metrics. The first two steps are completed within NASA's Land Information System Framework (LIS; Kumar et al., 2006).

Our meteorological estimates consist of the seven GDAS variables listed above, plus precipitation. Total precipitation is extracted from CHIRPS at a daily scale for retrospective analysis. Daily precipitation is disaggregated to 6-hourly estimates based on the diurnal cycle of MERRA-2 precipitation, using LIS's Land surface Data Toolkit (LDT; Arsenault et al., 2018). The remaining meteorological forcing variables are extracted from GDAS and downscaled to the spatial scale of the monitoring system (5 km), using bilinear interpolation with lapse-rate and aspect-slope correction within LIS.

The land surface model uses information from meteorological forcing variables to estimate hydrologic variables through a physically based representation of hydrologic processes and land surface energy balance. For this study, we present a monitoring system that uses the Noah-Multiparameterization (Noah-MP) land surface model. Noah-MP is an augmentation of the Noah LSM, which was first implemented in the NCEP Eta Data Assimilation System (EDAS) mesoscale forecast suite and NCEP Global Forecast System (GFS) to provide land surface feedback to climate models (Ek et al., 2003). The Noah LSM provides feedback by simulating the water fluxes and energy balance among canopy, vegetation, soil, streamflow, and snow (Livneh et al., 2010; Ek et al., 2003). Noah-MP augments the model representation of physical processes in the surface energy balance,

Table 1. Summary of datasets used in this study. The meteorological forcing fields include precipitation (precip), downward long-wave radiation (LW), downward shortwave radiation (SW), air temperature (T_a), specific humidity (Q), surface air pressure (P), zonal (U), and meridional (V) wind speed.

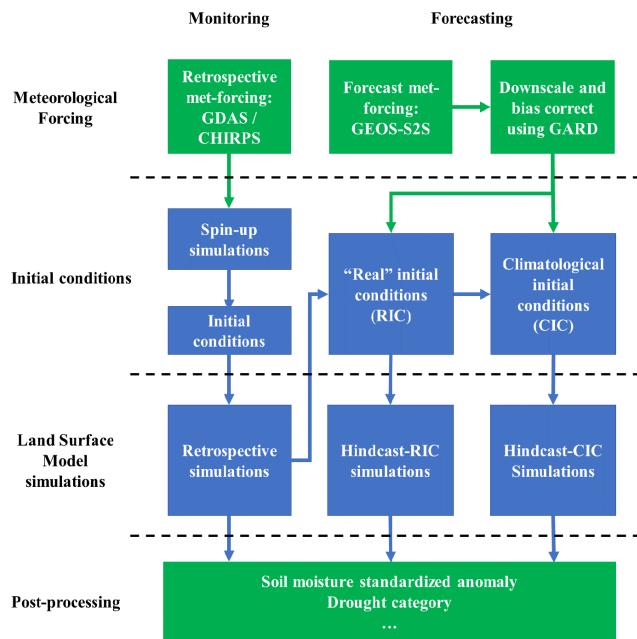| Data	Period of Record	Spatial resolution	Temporal resolution	Variable used	Reference
CHIRPS	1981–present	0.05°	1 d	Precip	Funk et al. (2015)
GDAS	2000–present	1° × 1° gradually improved to 0.125° × 0.125°	6 h	LW, SW, T_a, Q, P, U, V	National Climatic Data Center (2020)
GEOS-S2S-V1	January 1981–2018	1° × 1.25°	1 d	Precip, LW, SW, T_a, Q, P, U, V	Borovikov et al. (2019)
GEOS-S2S-V2	1981–present	0.5° × 0.5°	1 d	Precip, LW, SW, T_a, Q, P, U, V	Molod et al. (2020)
ESA-CCI SM	November 1978–December 2019	0.25° × 0.25°	1 d	Surface soil moisture	Gruber et al. (2019)

Figure 2. A schematic workflow of the South and Southeast Asia monitoring and forecasting system. Retrospective simulations (open-loop), hindcasts with "Real" initial condition (hindcast-RIC) simulations, and hindcasts with climatological initial condition (hindcast-CIC) simulations are designed to evaluate the monitoring and forecasting system.

snow and frozen soil, groundwater, runoff, and leaf dynamics. In addition, Noah-MP enables multiple options for a variety of physical processes within the model (Niu et al., 2011; Yang et al., 2011).

A variety of spin-up techniques has been used to acquire initial conditions for hydrologic models (Seck et al., 2015) and land surface models (Cai et al., 2014; Nie et al., 2018). Here, a 57-year offline spin-up is performed by simulating the period 2000–2018, three times, to obtain equilibrium hydrological states (e.g., groundwater storage) under prevailing climate patterns for our meteorological data sources. These equilibrium states, in principle, help monitoring simulations reach the best estimation of hydrological outputs once the monitoring period begins. The monitoring system is then initialized on 1 January 2000, owing to the availability of the GDAS data product and is run up to near real time. The system operates at a 15 min time step and generates water and energy fluxes and states at a 5 km spatial resolution that we save at daily temporal resolution, including soil moisture, evapotranspiration, terrestrial water storage, and snow water equivalent, among others. The system has satellite data assimilation capabilities (Getirana et al., 2020c; Kumar et al., 2020; Xue et al., 2019; Kumar et al., 2019), but the simulations presented here are open-loop simulations in which satellite data are integrated via parameter fields and meteorological forcings rather than through active land data assimilation.

Satellite-informed input parameters include the 1 km² resolution Moderate Resolution Imaging Spectroradiometer-International Geosphere Biosphere Program (MODIS-IGBP) land cover dataset (Friedl et al., 2010), 5 min FAO soil texture dataset (http://iridl.ldeo.columbia.edu/SOURCES/.NASA/.ISLSCP/.GDSLAM/.Hydrology-Soils/.soils/.dataset_documentation.html, last access: 17 December 2020), 30 m Shuttle Radar Topography Mission (SRTM) elevation dataset (Farr et al., 2007), 0.144° global albedo

maps (Csiszar and Gutman, 1999), and green vegetation fraction maps (Gutman and Ignatov, 1998) derived from measurements made by the advanced very high resolution radiometer (AVHRR) onboard NOAA's polar orbiting satellites. These parameters are represented as climatologies in our simulations.

We ran our monitoring system in retrospective mode from 2000 to 2017, which is initialized with a restart file from the spin-up simulation. This retrospective simulation serves as a baseline when evaluating the forecast system. From that time onward, the monitoring system has run in near-real-time mode. On account of the 3-week latency of the CHIRPS product, we use CHIRP precipitation to extend the retrospective simulation to real time. The simulations during this period are re-run with CHIRPS once CHIRPS becomes available.

2.4 Forecasting system

The forecasting system applies downscaled ensemble GEOS-S2S forecasts to drive Noah-MP simulations out to a 9-month forecast lead. The workflow and model specifications of the forecasting system are similar to the monitoring system except for meteorological forcing data processing (Fig. 2).

In the forecasting system, the same meteorological forcing variables as in the monitoring system are extracted from GEOS-S2S data products (see details of GEOS-S2S in Sect. 2.2). Due to the coarse resolution and inevitable bias of global climate model outputs, these forcing variables from GEOS-S2S-V1 are downscaled and bias-corrected to corresponding monitoring system forcing variables (i.e., precipitation from CHIRPS and other variables from GDAS) using the Generalized Analog and Regression Downscaling (GARD) algorithm (https://github.com/NCAR/GARD, last access: 3 April 2020; Gutmann et al., 2020). This downscaling algorithm takes a training dataset, prediction dataset, and observation dataset as inputs. The observation dataset contains records of variables (dependent variables) with targeted fine spatial resolution, for example, the precipitation from the CHIRPS dataset. The training dataset includes records of coarse-spatial-resolution variables (independent variables), which have the same time resolution as the dependent variables, for example, hindcasted GEOS-S2S-V1 precipitation. The prediction dataset contains the records of the same coarse-resolution variables as the training dataset but acquired in the forecast period – for example, new GEOS-S2S-V1 forecast precipitation. The prediction dataset contains the records of the same coarse-resolution variables as the training dataset but acquired in the forecast period – for example, new GEOS-S2S-V1 forecast precipitation. The GARD algorithm downscales the prediction dataset to the resolution of the observation dataset with an analog–regression approach.

This downscaling algorithm takes a training dataset, prediction dataset, and observation dataset as inputs. The observation dataset contains records of variables (dependent vari-

ables) with targeted fine spatial resolution, for example, the precipitation from the CHIRPS dataset. The training dataset includes records of coarse-spatial-resolution variables (independent variables), which have the same time resolution as the dependent variables, for example, hindcasted GEOS-S2S-V1 precipitation. The prediction dataset contains the records of the same coarse-resolution variables as the training dataset but acquired in the forecast period – for example, new GEOS-S2S-V1 forecast precipitation. The GARD algorithm downscales the prediction dataset to the resolution of the observation dataset with an analog–regression approach.

In this application, we apply GARD to downscale each forecast variable as a function of the same variable in the training datasets (e.g., precipitation to predict precipitation). GARD has the capability to use multiple predictor variables to improve downscaling accuracy, but the influence of the different combinations of independent variables in GARD is beyond the scope of this paper.

GEOS-S2S-V1 products consist of a series of meteorological forecasts initialized about every 5 d. To construct proper training datasets, we only use the first several days (approximately 5 d) of each forecast simulation before the next forecast is initialized to create one record of independent variables per day from the year 2000 to 2017. This approach to creating training datasets aims to take advantage of frequent (approximately every 5 d) GEOS-S2S-V1 hindcasts and to avoid outliers created by mismatched GEOS-S2S-V1 variables with corresponding GDAS/CHIRPS variables due to the sharp forecast skill decline of climate models in later forecast periods (Jie et al., 2017). CHIRPS and GDAS are used to construct the observation dataset. Due to the changing resolution of GDAS data products, coarser-resolution data in the early years are downscaled to $0.125° \times 0.125°$ using bilinear interpolation with lapse-rate and aspect-slope correction to unify the spatial resolution. This $0.125° \times 0.125°$ resolution GDAS product is then aggregated into daily data to unite time intervals with the training dataset.

In order to correct the evolving bias between longer lead-time GEOS-S2S-V1 precipitation hindcasts and CHIRPS, the total precipitation variable is further post-processed using a cumulative distribution function (CDF) matching method (Yuan et al., 2014). In each month within the 9-month forecast period, daily GARD-downscaled GEOS-S2S-V1 precipitation and CHIRPS precipitation from 2000 to 2017 are used to construct the precipitation forecast and observation CDF function separately. We acknowledge that it is possible to achieve bias correction on a finer timescale by constructing a GARD training dataset and CDF functions at sub-monthly scale (e.g., weekly scale), but shortening the timescale will also reduce the size of the training data.

The downscaled meteorological variables are further disaggregated from the daily timescale to the 6-hourly time series to capture sub-daily time variation. Total precipitation and solar radiation are disaggregated using the ratio between

6-hourly and daily GDAS climatology data, while the other six variables are disaggregated using the difference between 6-hourly and daily climatology data.

The performance of the downscaling method is evaluated at the monthly timescale in the five major river basins in South Asia using root mean squared error (RMSE) (Figs. S1–S5). The RMSE is first calculated for GEOS before and after downscaling against the retrospective forcing (i.e., the combination of GDAS and CHIRPS) and then normalized by the range of the retrospective forcing.

Overall, the RMSE of air temperature, surface pressure, and relative humidity is greatly reduced after downscaling. In addition to the GARD algorithm, the applied CDF matching has further reduced the RMSE for precipitation. For other fields, the impacts of downscaling differ across basins. For instance, downscaling leads to reduced RMSE of wind speed for the Ganges basin while its impact on shortwave and longwave radiation is marginal.

A set of hindcast simulations has been initialized on 1 May from the year 2000 to 2017. The meteorological forcing of these simulations is GEOS-S2S-V1 initialized on 1 May, downscaled following the forecasting system workflow, and the initial conditions of these simulations are obtained from retrospective simulations (i.e., "real" initial condition; Fig. 2). Each initialized hindcast simulation has 10 ensemble members and lasts for 9 months (i.e., May to January of next year). 1 May was chosen as the initialization date of this experiment to capture the monsoon seasons in the monsoon regions and dry season in the western part of the domain (see Sect. 2.1 for details). This set of hindcast simulations with "real" initial conditions (hereafter referred to as hindcast-RIC simulations) is designed to evaluate the performance of the forecast workflow.

An additional set of control hindcast simulations with climatological initial conditions (hereafter referred to as hindcast-CIC simulations; Fig. 2) is designed to study the impact of initial conditions on the forecast skill. In this set of hindcast-CIC simulations, the workflow and all settings are the same as hindcast simulations except for initial conditions. All 18 hindcast-CIC simulations use the climatological initial conditions (CIC) calculated by averaging 1 May hydrological states from 2000 to 2017 obtained from retrospective simulations.

In additional to the hindcast simulations, the forecast system is also running operationally, monthly, with the same forecast workflow but driven by downscaled GEOS-S2S-V2 product instead of downscaled GEOS-S2S-V1 product. The GEOS-S2S-V2 dataset also provides 9-month daily meteorological forecasts about every 5 d, but with finer spatial resolution of $0.5° \times 0.5°$. Evaluation of the operational forecast system is planned for the future, however, because GEOS-S2S-V2 offers only a small ensemble prior to 2017, making evaluation less reliable at this time. This paper focuses on the evaluation of the hindcast simulations driven by downscaled GEOS-S2S-V1 from 2000 to 2017.

In the hydrological forecasting system, meteorological forecast forcing data are sourced from all ensemble members launched on a common GEOS-S2S forecast initialization date, i.e., the one date per month that offers the 10-member ensemble. It is possible to include other GEOS-S2S meteorological forecasts in the hydrological forecasting system, but proper bias-correction methods are needed to correct inconsistent forecast skill or bias among GEOS-S2S meteorological forecasts with different initialization dates. For example, in the hindcast-RIC hydrological simulations initialized at 1 May, a combination of 26 April and 1 May GEOS-S2S-V1 meteorological forecasts can be used to enlarge the size of the ensemble members. However, sub-monthly to monthly GARD and CDF post-processing methods would need to be applied to 26 April and 1 May forecasts separately.

2.5 Post-processing

In this study, the evaluation of the monitoring and forecast systems is performed on monthly scales. All simulations are first averaged from daily to a monthly timescale. The ensemble mean of forecast simulations generally shows much less variance than retrospective simulations due to averaging effects (Koster et al., 2004). To compare simulations from both systems and other datasets, all simulation results and data products are standardized before comparison.

Climatological monthly mean and standard deviations for retrospective simulations are calculated for each month separately (January to December) with simulation results from 2000–2017. The retrospective simulation results are then standardized in the form of the standardized anomaly with the following equation,

$$\text{stdamly}_{\text{yyyymm}} = \frac{\text{val}_{\text{yyyymm}} - \overline{\text{val}_{\text{mm}}}}{\text{SD}_{\text{mm}}},$$

where $\text{stdamly}_{\text{yyyymm}}$ is the standardized anomaly for year "yyyy" and month "mm", $\text{val}_{\text{yyyymm}}$ denotes the original values from simulations and analysis datasets for year "yyyy" and month "mm", $\overline{\text{val}_{\text{mm}}}$ is the climatology mean for month "mm", and SD_{mm} is the climatological standard deviation for month "mm". We note that GDAS is an operational analysis system chosen for its low latency. It is not a time-consistent reanalysis product. This means that over the study period GDAS underwent several significant changes in input and data structure, including changes in spatial resolution. All GDAS data were regridded to a common resolution, and lapse rate and slope-aspect corrections were applied to downscale to topography. Nevertheless, the use of this operational product does mean that anomalies calculated against the long-term mean can contain some statistical artifacts. As the primary purpose of these anomalies is application to drought monitoring, we accept this limitation with the understanding that the system can be used to capture significant drought events but that it is not optimized for trend detection or for precise ranking of event intensities over time.

The forecast simulation results are standardized similarly. For simplicity, the results presented in the following sections are only based on the standardized ensemble mean for forecast simulation related analyses, for which the standardization is applied after the calculation of ensemble mean.

3 Results and discussion

3.1 Retrospective simulations vs. hindcast-RIC simulations

For purposes of agricultural drought prediction, we are most concerned with our ability to predict soil moisture anomalies. As large-scale networks of root zone soil moisture (RZSM) observations are rare in South and Southeast Asia, we evaluate the prediction skill of the forecast system of SAHFS-S2S by comparing RZSM estimates in hindcast-RIC simulations to RZSM in the retrospective simulations. The depth of root zone soil moisture is defined as 1 m in this study. Since the retrospective run and the hindcast-RIC simulations use the same land surface model, these comparisons aim to evaluate the impact of meteorological forcing on the prediction skill of RZSM. The comparison is performed by calculating the inter-annual correlation coefficients (R) of monthly RZSM at different lead times, from May (1-month lead time) to January (9-month lead time). Figure 3 shows maps of these correlation coefficients across forecast lead times. Higher correlations indicate more skillful RZSM hindcasts. In May (1-month lead time), the correlation between hindcast-RIC and the retrospective simulations is positive in most of the region, and the correlation is significant at 0.95 significant level except for northeast India. As expected, correlations drop in strength and significance as lead time increases. In June (2-month lead time), the correlation remains positive and significant in the Indochinese Peninsula and the west of the research domain. The prediction skill of the RZSM drops quickly and becomes insignificant in most of the regions except for the northwestern regions of the domain 2 or 3 months after the forecast initialization date. The high forecast skill in the northwestern regions is drawn from the "real" initial conditions (i.e., long hydrological memory) due to low precipitation. This phenomenon is further explained in detail at the basin scale below and in Sect. 3.2.

We repeat this analysis at the basin scale for the five major river basins of the South and Southeast Asia area (highlighted in Fig. 1). Figure 4 shows RZSM and precipitation comparison in the Ganges basin, which receives considerable precipitation during the summer monsoon. Please note that the time axis in Fig. 4 is rearranged, so that data of a specific lead month are grouped. Figure 4a shows the monthly precipitation time series averaged for the Ganges basin. In each month, the magnitudes of precipitation from retrospective (i.e., CHIRPS) and hindcast simulations (i.e., downscaled GEOS5-S2S-V1) are similar to each other due to the ap-

plication of CDF matching (see Sect. 2.4). The inter-annual variation of the hindcast precipitation and retrospective precipitation both are higher in wet seasons than in dry seasons. The inter-annual variation of hindcast precipitation, however, is smaller than that of the retrospective precipitation, especially in months with large lead time. This difference in inter-annual variability is the result of averaging a across large forecast ensemble spread in later months.

Figure 4b shows the climatology RZSM (i.e., average RZSM from the year 2000 to the year 2017) for May to January from retrospective and hindcast-RIC simulations. The climatologies of RZSM from retrospective and hindcast-RIC simulations have similar magnitude and seasonality. The magnitudes of inter-annual variability of each month, which are represented by inter-annual standard deviation and are shown as error bars in 4b, are smaller in hindcast-RIC simulations. The reason for this magnitude difference is, again, the effect of averaging the large ensemble spread of hindcast-RIC simulations. Climatologically, as the monsoon season picks up in July, RZSM increases dramatically due to intense precipitation. The RZSM climatology peaks in August, indicating some delay between the precipitation peak and the annual maximum in soil water storage.

Figure 4c shows the RZSM standardized anomaly and rainfall inter-annual correlations between retrospective and hindcast-RIC simulations for each month in the Ganges basin. Precipitation has positive correlations during the rainy seasons, but this prediction skill is never statistically significant. At the end of the monsoon (October), precipitation skill drops dramatically to near zero. In contrast, the correlation for RZSM (yellow line in Fig. 4b) fluctuates around the significant line (dashed line in Fig. 4b, $R = 0.484$) as lead time increases. This fluctuation of RZSM skill within the monsoon season is closely driven by the performance of the meteorological forecast of precipitation (green line in Fig. 4b). Although the skill in precipitation drops rapidly after the monsoon season, the RZSM forecast skill maintains around the significant line with a mild declining trend. This difference indicates that the low precipitation amount after monsoon seasons has little influence on the inter-annual variability of the RZSM. The high basin-scale RZSM correlation during July to December (yellow line in Fig. 4b) is a contrast to the relatively low pixel-scale RZSM in the Ganges basin (Fig. 3c–h). This difference reflects the fact that small spatial mismatch between retrospective and hindcast-RIC simulations, which is likely due to the same mismatch of meteorological forcings, especially precipitation (see blue line in Figs. 4b and S6), can lead to disagreement at pixel scale that averages out at basin scale.

Figure 4d shows the monthly RZSM standardized anomaly (see section 2e) from retrospective and hindcast-RIC simulations. The ensemble means of RZSM from hindcast-RIC simulations are shown in red dots, and the error bars represent the standard deviation of the ensemble. The spread of the ensemble members is relatively small in
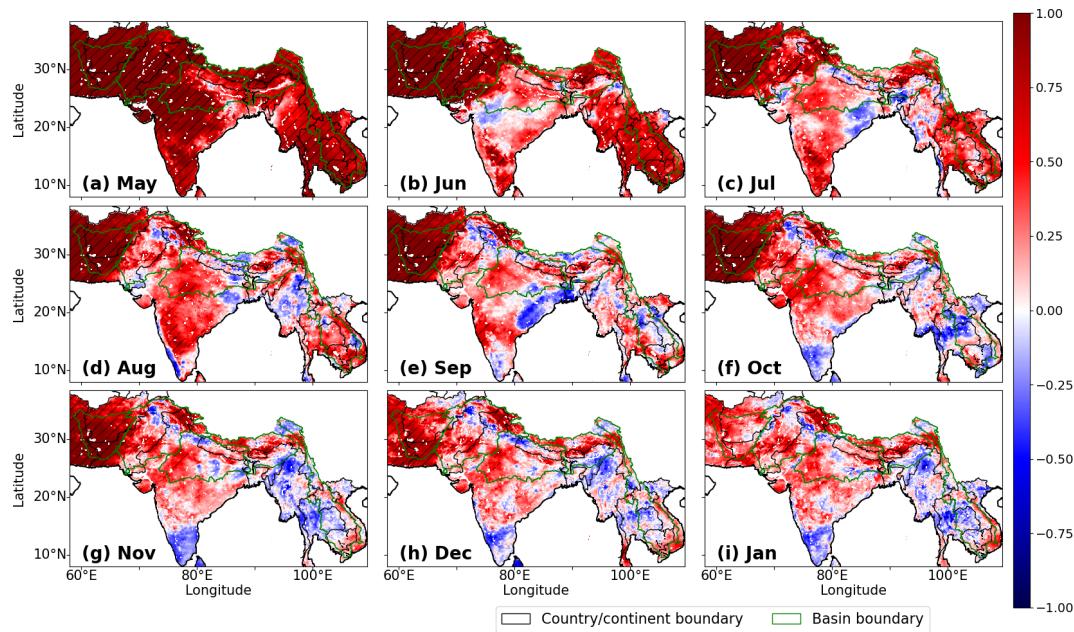
Figure 3. Inter-annual correlation of RZSM between retrospective simulations and hindcasting simulations which is computed using RZSM data from year 2000 to 2017 in months May (1-month lead time) to January (9-month lead time) **(a–i)**. The hatches denote the areas with statistically significant correlation at 0.95 confidence level.

May, which represents high confidence in the prediction from the forecast system. This confidence originates from the relatively low uncertainty of meteorological forcing variables in May. The spread increases quickly in June (i.e., 2-month lead time) and later lead months.

The timing of the Indus basin rainy season is similar to the Ganges, but the magnitude of precipitation is far smaller (Fig. 5). As a result, the average RZSM from all years is much lower in the Indus (Figs. 4b and 5b). Precipitation amount increases from May to June, but high evapotranspiration (ET) in June causes a deficit in the water budget that causes a decrease in RZSM in June (Fig. S7). Though climatology of precipitation is low from November to January, monthly RZSM climatology still increases during this period due to low ET. Hindcast-RIC simulations underestimated climatological RZSM in every month, and the underestimation grows in later dry months. This underestimation is, in part, due to the high bias of ET in the hindcast-RIC simulations, which is an issue of ongoing study. The precipitation correlations between hindcast-RIC and retrospective simulations are low and statistically insignificant for all months, as the maximum correlation is 0.22 in May (red line in Fig. 5c). Though hindcast-RIC simulations show lower prediction skill in precipitation in the Indus than in the Ganges basin, the RZSM correlations in the Indus are still significant in the first 3 months (orange line in Fig. 5c). This relatively strong RZSM prediction skill is again a result of a drier climate. The relatively low amount of precipitation in

the Indus basin allows the skill derived from initial conditions to persist longer.

In the Helmand basin, the magnitude of precipitation is smaller than in either the Ganges or Indus basin. More importantly, the seasonal precipitation cycle is different as well (Fig. 6a). The Ganges and Indus basins experience peak precipitation during the summertime South Asian monsoon, such that our May forecast initialization captures the onset of the rainy season. The Helmand basin receives precipitation in winter–spring, with the dry season setting in May, just as our forecasts initialize (Fig. 6a). This allows us to see the behavior of a forecast that is nearly pure persistence: there is some precipitation in May, for which hindcasts show skill, but for the following 5–6 forecast months, there is effectively no precipitation, and soil moisture shows a steady decline (Fig. 6b). This allows skill present in the initial conditions to persist, as seen in the high RZSM correlation between hindcast-RIC and retrospective simulations through November (Fig. 6c). This high correlation starts to drop from November but remains significant until December, even though precipitation has low prediction skills in November and December. While this result might not be indicative of seasonal forecast skill in the Helmand basin in general, since forecasts initialized for the rainy season might be of broader relevance, the result does provide a useful example of how our forecast system behaves in basins initialized during the dry season.
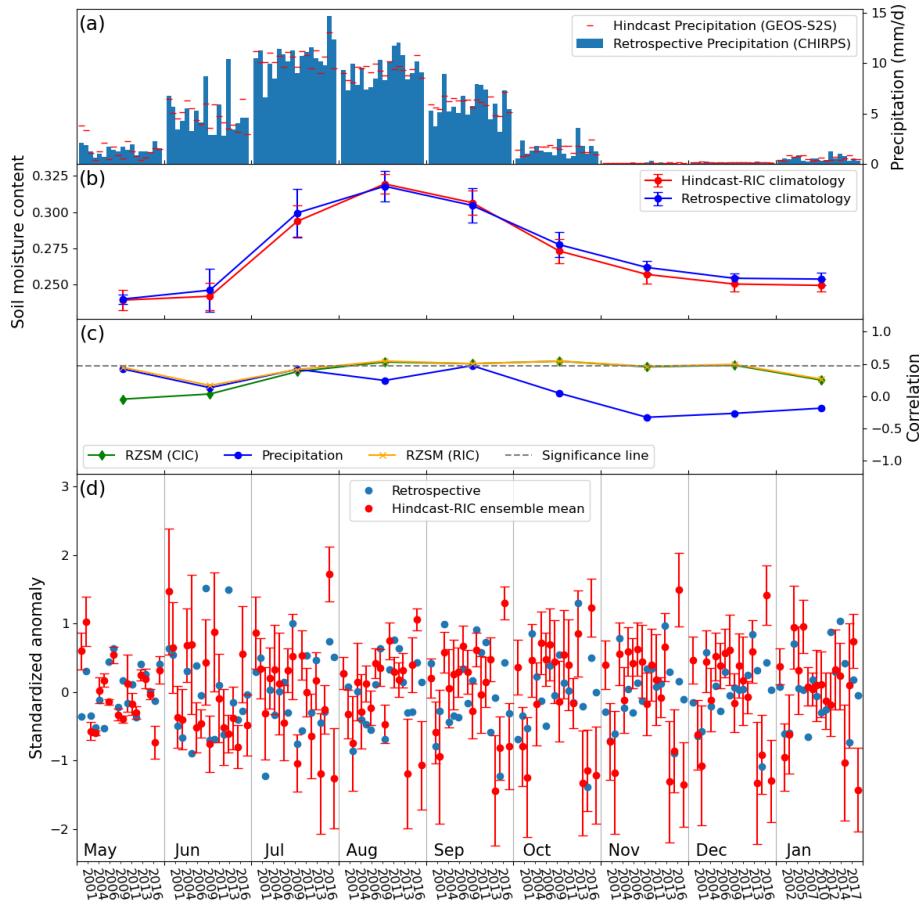
Figure 4. Comparison between retrospective simulations and hindcast simulations in the Ganges basin of **(a)** monthly time series of precipitation, **(b)** monthly climatology of root zone soil moisture (RZSM), **(c)** inter-annual correlations of basin-averaged RZSM standardized anomaly and precipitation, and **(d)** monthly standardized anomaly of RZSM (the red error bars represent for the standard deviation of separate hindcast ensemble members). Please note that the time axis for monthly time series are rearranged so that the data for the same month are grouped.

3.2   Impact of initial conditions

The RZSM forecast skill is a product of the accuracy of both the forecast meteorological forcing variables and initial hydrological states. In this section, we examine the impact of initial conditions on prediction skills by comparing hindcast-RIC simulations to our hindcast-CIC simulations (see section 2d for details). We focus on RZSM correlations with the retrospective simulation as our performance metric, since we are most concerned with the ability of quality forecast initialization to improve the simulation of inter-annual RZSM variability. Figure 7 shows the contribution of initial conditions as the difference in correlation with retrospective RZSM between hindcast-RIC and hindcast-CIC simulations. In these maps, shades of red indicate areas and times where hindcast-RIC have improved skills relative to hindcast-CIC simulations. For lead-one forecasts (May), we see a positive contribution of initial conditions to RZSM forecast skills primarily in India. For lead-two forecasts (June), the area with the

largest correlation difference moves to the west of the domain, and the area with a significant correlation difference in southern India also shrinks. In later months, from July to December, the areas with significant differences of correlations continue to shrink because precipitation, along with other meteorological forcing variables, gradually reduces or even eliminates the impact of initial conditions. In portions of the west of the domain (Helmand basin), where the precipitation is low from May to November, the difference in correlations remains significant until December.

Figure 8 shows the number of lead months before the difference in RZSM forecast skill – measured as correlation with the retrospective simulation – from the two sets of hindcast simulations (hindcast-RIC and hindcast-CIC) drops to a negligible value (in this case, 0.01). Given that the only difference in the two sets of hindcast simulations is the initial conditions, Fig. 8 indicates the duration of the impact of the initial conditions on the performance of hindcast simulations. A blueish color in this figure indicates areas with short
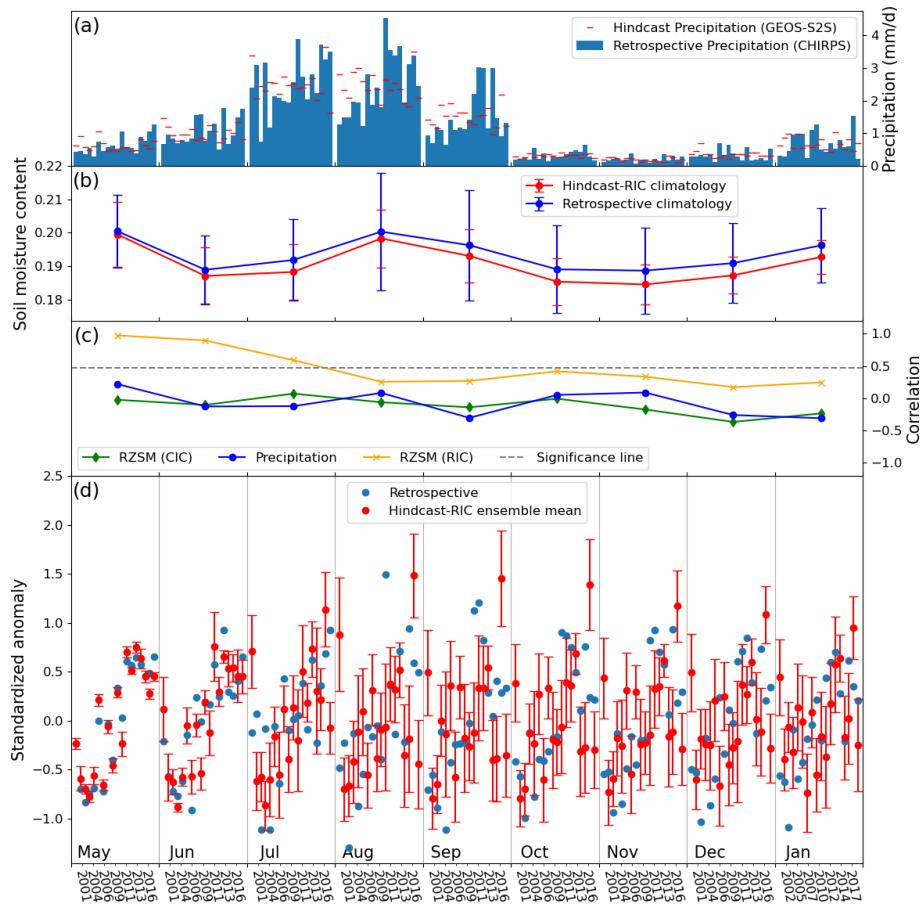
Figure 5. Same as Fig. 4, but for the Indus basin.

memories of the initial condition. We note that our hydrological hindcasts are based on "offline" simulations in which surface conditions like soil moisture do not affect meteorological forcings. The influence of the initial conditions, then, is limited to direct impacts of water storage on the water balance and does not include any potential land–atmosphere feedbacks. Short memories generally present in areas receiving considerable summer monsoon precipitation, such as the west coast and northeast parts of India, the southern slopes of the Himalayas, the west coast, and the southeast of the Indochinese peninsula (Figs. 8, S8). The soil in these wet areas becomes saturated during the monsoon seasons due to intense precipitation. The memories of the initial conditions are thus weakened or eliminated. In contrast, having less precipitation, the yellow areas have longer memories of the initial conditions. This result is specific to our initialization month and will differ with respect to the timing of the rainy season.

At basin scale, we see significant variability in the role that initial conditions play in RZSM forecasts. For the Ganges basin, the correlation of RZSM between the retrospective simulations and the hindcast-CIC simulation is -0.04 in May (Fig. 4c green line) due to low precipitation correlation and inconsistent initial condition between the retrospective and hindcast-CIC simulations. A significant difference between hindcast-RIC RZSM (Fig. 4c orange line) and hindcast-CIC RZSM (Fig. 4c green line) in terms of correlation with retrospective RZSM in May suggests that an accurate initial condition is crucial to have meaningful RZSM prediction. This difference, however, decreases in June and July. The RZSM predictions from hindcast-RIC and hindcast-CIC simulations become almost identical after July as the memory of the initial condition is eliminated by the accumulating effects of meteorological forcing variables. The elimination of the initial condition is also observed rather quickly in June in the Mekong basin (Fig. S9)

In the drier Indus basin, the difference in RZSM prediction skills between the hindcast-RIC and hindcast-CIC (Fig. 5c yellow and green lines) reduces over the first 3 months. This difference, however, is never eliminated. The difference between these two correlations becomes roughly constant from August to January. In January, 9 months after initialization, the RZSM from hindcast-RIC simulations still shows skillful prediction while the RZSM correlation between the hindcast-CIC and retrospective simulations becomes negative. This result indicates that in a drier basin, the impact of the initial conditions is reduced but can still positively contribute to the
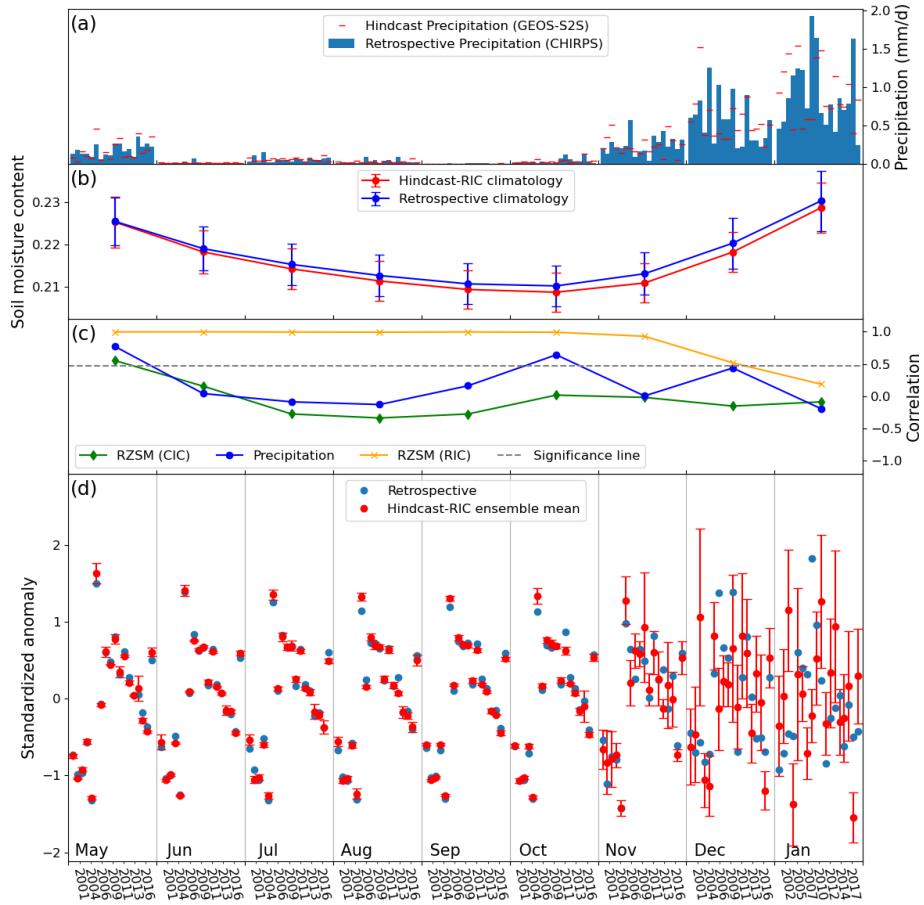
Figure 6. Same as Fig. 4, but for the Helmand basin.

prediction skills of RZSM even 9 months after the initialization date. In the Brahmaputra basin, the impact of the initial condition also lasts for the entire 9-month prediction period (Fig. S10).

In the Helmand basin (see the location of Helmand basin in Fig. 1), both hindcast-RIC and hindcast-CIC simulations have significant correlations with RZSM in retrospective simulations at the lead-one month (May) due to high agreement between hindcast and retrospective precipitation products, but the hindcast-RIC simulation is still significantly better than hindcast-CIC in this month. The predictability of RZSM in hindcast-CIC simulations brought by skillful precipitation prediction in May is quickly lost in June and drops to negative values from July to January, while the hindcast-RIC simulation has skillful RZSM predictions through December on account of the long memory of initial conditions through the protracted Helmand basin dry season.

The impact of initial conditions has also been studied using ensemble streamflow prediction (ESP) and reverse-ESP methods (Yuan et al., 2016; Shukla and Lettenmaier, 2011; Luo et al., 2008; Ma et al., 2018). These studies have yielded similar conclusions regarding the fact that the initial condition has a longer impact when the forecast is initialized in a

dry season and a shorter impact when the forecast is initialized in a wet season.

3.3 Comparison with satellite estimates

The top-10 cm surface soil moisture (SSM) from retrospective and hindcast-RIC simulations is also evaluated against the ESA-CCI SM products (see section 2b for details). SSM values extracted from the ESA-CCI SM product, retrospective simulations, and hindcast-RIC simulations are pre-processed to a monthly timescale before comparison. The monthly ESA-CCI SM data are calculated by averaging all available daily data for that month. To make a fair comparison, we calculate retrospective and hindcast-RIC monthly SSM data by first upscaling daily data to the same resolution as ESA-CCI SM ($0.25° \times 0.25°$) and then averaging to monthly timescale using only data from days when ESA-CCI SM daily data are also available.

Figure 9 shows the inter-annual correlation map between retrospective monthly SSM and ESA-CCI SM (left) and the inter-annual correlations between hindcast-RIC monthly SSM and ESA-CCI SM (right) using data in May, July, and September between 2000–2017. The retrospective SSM
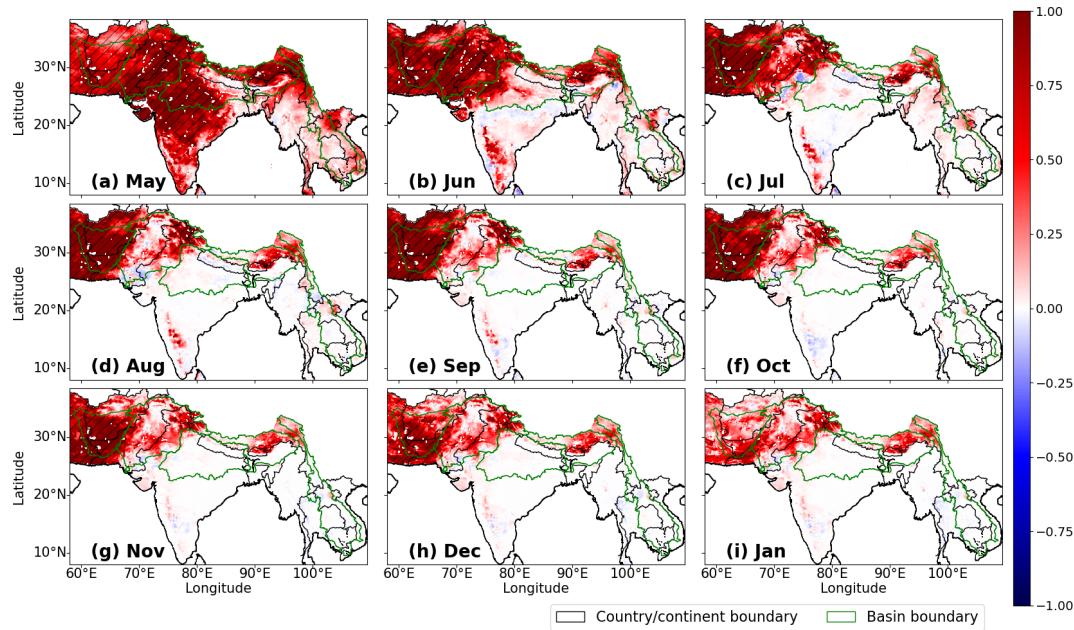
Figure 7. Inter-annual correlation difference of root zone soil moisture (1 m) between hindcast simulations using real initial condition (RIC) and climatological initial condition (CIC) against the retrospective simulations. The difference is computed as the correlation of RZSM between hindcast simulations and retrospective simulations minus the correlation between hindcast simulations and retrospective simulations. The hatched areas denote the statistically significant correlation difference at 0.95 confidence level.

Figure 8. The smallest lead month when the forecast skills of root zone soil moisture (1 m) between hindcast simulations using real initial condition and climatological initial condition (CIC) against retrospective simulations is less than 0.01.

in May has a high agreement with ESA-CCI SM in areas in south India, the Indochinese Peninsula, and northern Afghanistan. In July and September, the retrospective simulation captures more spatial patterns of ESA-CCI SM than in May in Pakistan and India, but less in Afghanistan and the Indochinese Peninsula.

Hindcast-RIC SSM generally has a lower correlation with ESA-CCI SM than retrospective SSM. In May, hindcast-RIC SSM has a significant positive correlation with ESA-CCI SM in the Indochinese Peninsula and northern Afghanistan. In July and September, the forecast skill of SSM dramatically
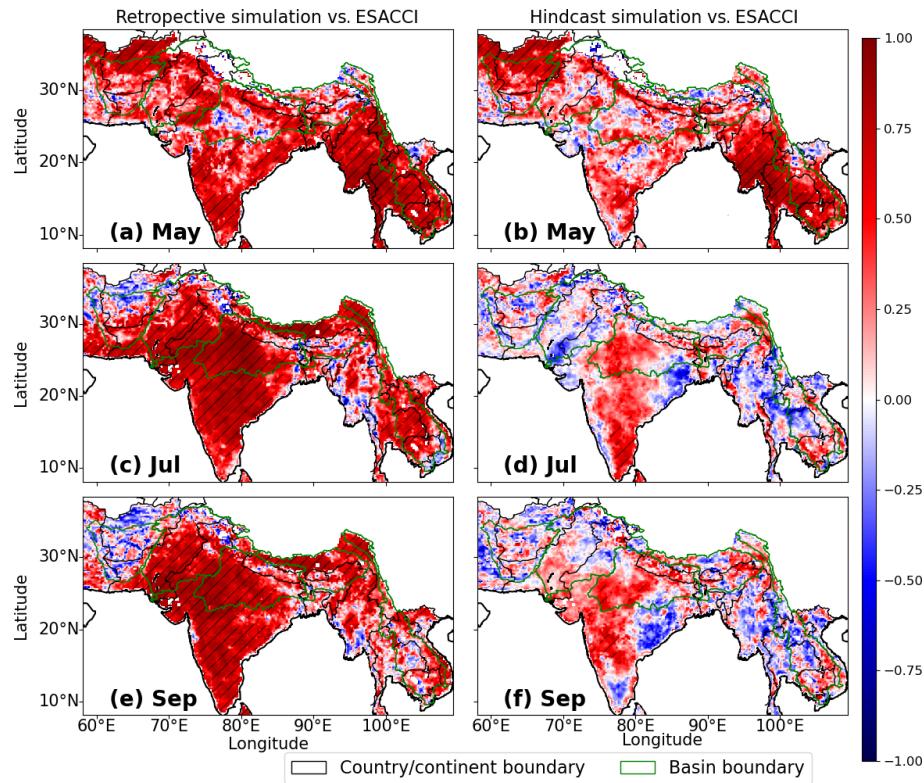
Figure 9. Surface soil moisture (top 10 cm) correlation between European Space Agency Climate Change Initiative (ESA CCI) soil moisture product and **(a, c, e)** retrospective simulations, and **(b, d, f)** forecasting simulations in May, July, and September. The hatched areas denote the statistically significant correlation at 0.95 confidence level.

drops. Areas with significant SSM only appear sporadically in India, Nepal and the Indochinese Peninsula.

Figure 10 shows basin-scale comparisons with ESA-CCI SM, with precipitation seasonality included for context (Fig. 10a). At this scale we see high correlations between the retrospective simulation and ESA-CCI SM in most basins in most months: correlations are significant in the Indus basin in all months; in the Helmand basin in all months but September; in the Brahmaputra basin in all months except for November; in the Ganges basins in all months except for November and December; and in the Mekong basin in all months except for August and September.

As expected, correlations are lower for the hindcast-RIC simulations. In May, there are significant skills in all basins but the Ganges basin, where the hindcast-RIC had also shown low skill relative to retrospective simulations (see Fig. 4). In June, skill drops off surprisingly quickly in the Indus and Helmand basins, both of which showed extended RZSM forecast skill when evaluated against the retrospective simulation (see Figs. 5 and 6). This occurs despite the fact that ESA-CCI SM and the retrospective run have a relatively high correlation in these months. One reason that skill relative to ESA-CCI SM might drop off so quickly in these drier basins is that the memory of SSM, as opposed to RZSM, is short, especially in dry areas, and SSM can be highly sensitive to modest rainfall. This could make the ESA-CCI SM evaluation more sensitive to errors in forecast precipitation (see precipitation correlation in Figs. 5c, 6c) than the RZSM evaluation was, leading to a rapid loss of skill due to imperfect precipitation forecasts in the first few months of the simulations. From July onward the basin-scale hindcast-RIC evaluations relative to ESA-CCI SM are noisy; all basins drop to non-significant correlations in July and August, but there is a rebound to skillful prediction in the Mekong and Helmand in September due to the rebound of the precipitation forecast skill (see precipitation correlation in Figs. 6c and S9c). We also note that in the Helmand basin, the SSM correlation in October shows the opposite tendency from the precipitation correlation (Fig. 6c). The difference could be due to errors in ESA-CCI SM or CHIRPS, the influence of missing data in ESA-CCI SM, or noise in the forecasts. By November, the correlation between ESA-CCI and hindcast-RIC SSM drops to insignificant levels for all basins. It is worth noting that although on the basin scale we see significant SSM correlation in the Brahmaputra basin in September, spatial SSM correlations suggest lower forecast skills (Fig. 9). This difference in correlation suggests a spatial mismatch in the SSM predictions. A similar difference in correlation of spatial precipitation and basin-scale precipitation is also found in the Brahmaputra basin (Figs. 5c and S6).
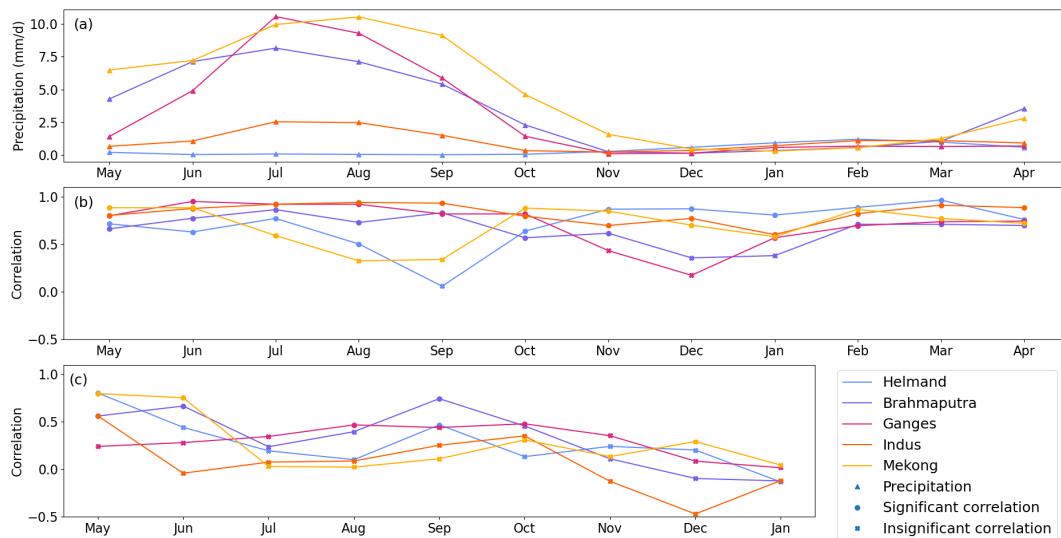
Figure 10. Comparison of **(a)** monthly retrospective precipitation climatology, **(b)** inter-annual correlation between monthly ESA-CCI SM product and monthly retrospective surface soil moisture (top 10 cm) and **(c)** inter-annual correlation between monthly ESA-CCI SM product and monthly forecasting SSM in five major river basins in South and Southeast Asia. Please note that the SSM data for a basin in a year are flagged as missing data if more than 50 % of data points in a basin are missing in ESA-CCI SM monthly data. The criterion of significance of correlation are then different due to different sample sizes.

3.4 Case study of the 2015 South and Southeast Asia drought

The 2015 El Niño event caused widespread drought in South to Southeast Asia. This drought had significant impacts on health, food security, and fire risks in more than nine countries in south–southeast Asia (Van Der Schrier et al., 2016; Qian et al., 2019). Figure 11 shows the performance of our system in monitoring and forecasting the development of the 2015–2016 drought in the Ganges and Mekong basins. This hindcast is initialized on 1 May 2015.

In the Ganges basin, the precipitation (Fig. 11a) and RZSM (Fig. 11e) values from the hindcast generally match the magnitudes of the retrospective simulations. There is some month-to-month discrepancy in precipitation hindcast relative to the CHIRPS record used in the retrospective simulation, but over the course of the monsoon season, these differences nearly average out – the hindcast only slightly overestimates precipitation (and thus underestimates the severity of the drought). This results in an RZSM forecast that is generally consistent with the observations, albeit somewhat less dry overall and noisier month-to-month, up until the final month of the forecast.

In the Mekong basin, in contrast, the forecast underestimates monsoon season precipitation (overestimates drought severity) (Fig. 11b, d). This results in a hindcast in which drought is both more severe and more persistent than observed (Fig. 11f, h). Thus, the direction of the drought is captured in this basin, but in application the hindcast might overestimate the predicted impacts of this 2015 drought.

The forecast of the RZSM standardized anomaly also captures the spatial pattern in the first 2 months in both the Ganges basin and the Mekong basin (Figs. S11, S12).

3.5 Application of drought indicator

Applying a hydrological monitoring and forecast system to drought applications requires that the ensemble simulation output be converted to meaningful and interpretable drought indicators. Here we use the Mekong and the Helmand basins as examples to illustrate one method for doing so. This process starts with the monthly RZSM values from both retrospective and hindcast-RIC simulations. In each month, the percentile of the monthly RZSM values is calculated on a gridded basis, based on a normal distribution fitted from the 18-year data record at each grid cell. The severity of drought at each grid cell is then categorized based on the same drought categories used in the United States Drought Monitor (Svoboda et al., 2002): the severity is classified as exceptional drought (D4), extreme drought (D3), severe drought (D2), moderate drought (D1), or abnormally dry (D0) using RZSM percentile thresholds of 2 %, 5 %, 10 %, 20 %, and 30 %, respectively.

Figures 12 and 13 show the fractional area of each drought category for the retrospective simulation and the hindcast-RIC simulations, for all lead times, in the Mekong and the Helmand basins, respectively. In the Mekong basin, the hindcast simulations capture the major drought events reasonably well in May (1-month lead time), though they overestimate the drought areas in 2005 and underestimate them in 2010 (Fig. 12a). From June to August (2-month lead time to 4-
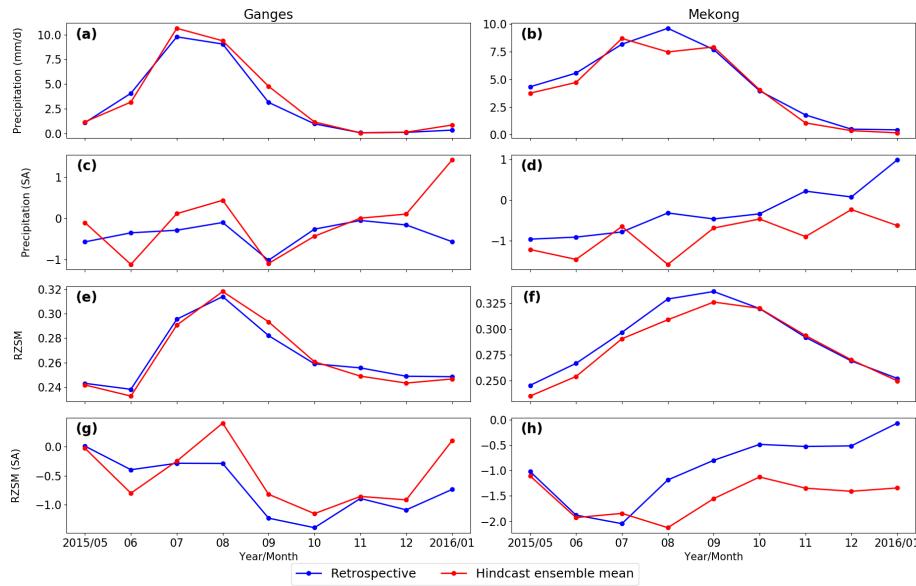
Figure 11. Comparison of monitoring and forecasting system on 2015–2016 drought event in Ganges basin (left panels) and Mekong basin (right panels) in terms of precipitation **(a, b)**, precipitation standardized anomaly **(c, d)**, root zone soil moisture (RZSM) **(e, f)**; and RZSM standardized anomaly **(g, h)**.

month lead time), the hindcast simulations predict significantly larger drought areas in 2005 and 2016 compared to retrospective simulations, while hindcast-RIC and retrospective simulations estimate relatively similar drought areas in other drought years (Fig. 12b–d). After August, the drought areas become poorly estimated in hindcast-RIC simulations compared to the retrospective simulations (Fig. 12e–i) due to unskillful hindcast-RIC RZSM estimates in the Mekong basin (Fig. S2c). If we focus on the 2015 drought event, the hindcast-RIC and retrospective simulations agree well in the first 3 months (Fig. 12a–c). Starting from August, the hindcast-RIC simulations estimate larger drought areas than the retrospective simulations, especially in the "exceptional drought" category (Fig. 12d–i). This overestimation of the drought area is consistent with the lower prediction of the RZSM standardized anomaly in Fig. 11, and this overestimation is mainly attributed to the lower precipitation forecast from the downscaled GEOS-S2S-1 (Fig. S9). The progress of a specific drought event can also be tracked by comparing drought categories across different months (see red rectangle in Figs. 12 and S5 for 2015 drought). The SAHFS-S2S shows similar skills forecasting drought categories in the Ganges (Fig. S13), Brahmaputra (Fig. S14), and the Indus (Fig. S15) basins.

Consistent with a high correlation of RZSM between retrospective and hindcast simulations (Fig. 6b), the fractional drought area in the Helmand basin agrees well for the first 7 months (Fig. 13a–g). From the year 2000 to 2004, the Helmand basin experienced severe drought conditions. In 2005, above-average rainfall ended this prolonged period of drought conditions. In 2016, the drought condition was rel-

atively stable from May to October and then became more severe in November and December according to retrospective and hindcast-RIC simulations. In January, however, retrospective simulations report a less severe drought than December, while hindcast-RIC simulations report a more severe drought. It is worth noting that an intensification of drought from one month to the next does not necessarily mean that the soil moisture is drier than the previous month. Each month's drought indices are calculated relative to the distribution of historical soil moisture conditions in that month.

4 Conclusions

In this study, we present a high-resolution soil moisture monitoring system and sub-seasonal to seasonal forecasting system for a South and Southeast Asia region, SAHFS-S2S. SAHFS-S2S consists of a physically based land surface model, analysis and observation-based meteorological forcing datasets, and downscaled dynamically based meteorological forecasts. We compare 18 hindcast-RIC simulations, each of which is initialized on 1 May in a year from 2000 to 2017, with corresponding retrospective simulations. The comparisons show that the RZSM in hindcast-RIC simulations have considerable skill for the first 2 months, especially in the western part of the study domain, the Indochinese Peninsula, and southern India. The hindcast-RIC simulations continue to have high skills in the western part of the research domain – which are the driest areas in the domain – for another 5 months while showing generally low skill in other regions. Results presented here only capture system
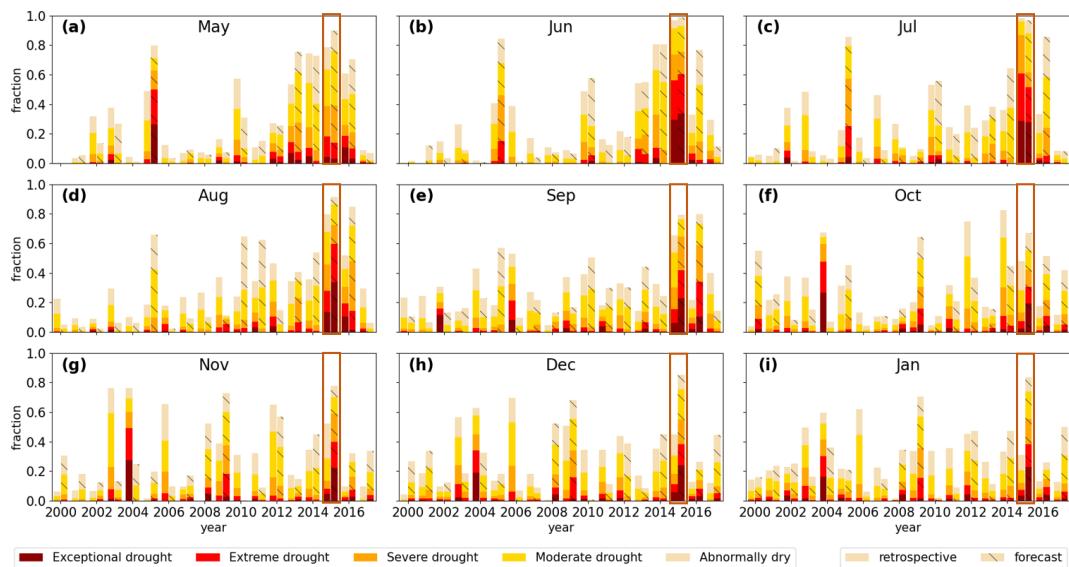
Figure 12. Fractions of areas of different drought categories in Mekong basin calculated using RZSM data from retrospective and hindcast-RIC simulations from **(a)** May to **(i)** January.
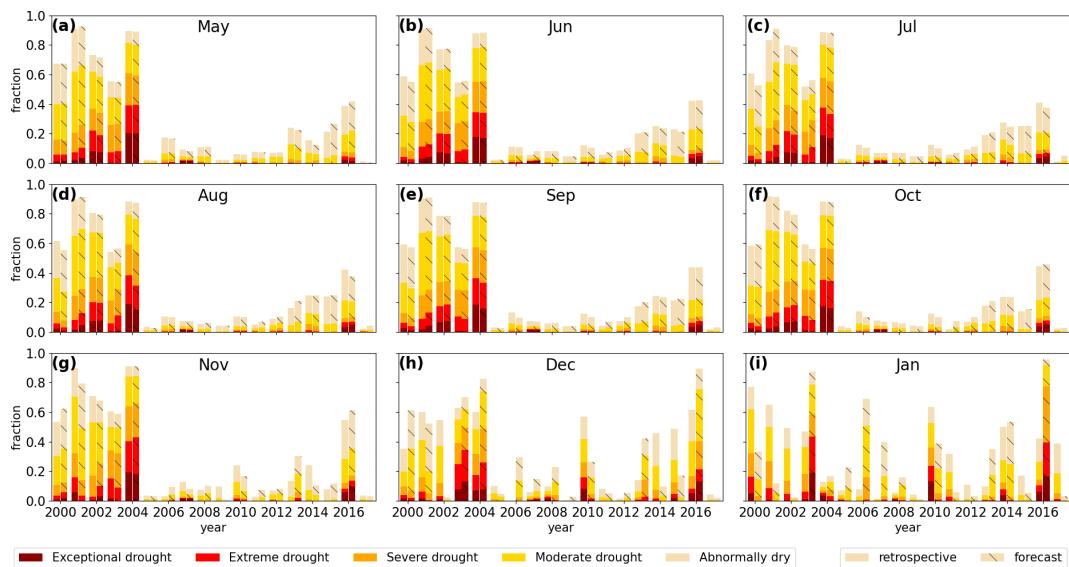

Figure 13. Same as Fig. 12 but for the Helmand basin.

skill for forecasts initialized in May, but diverse seasonality across the study domain allows us to examine forecast performance for both wet and dry season initialization dates.

To study the impact of the hydrological initial conditions on forecast skill, we designed a set of control hindcast simulations, which are initialized with climatological hydrological conditions (hindcast-CIC). In May (1-month lead time), the hindcast-RIC simulations outperform hindcast-CIC simulations in most parts of the domain except for the Indochinese Peninsula, where heavy precipitation quickly eliminates the memory of the initial conditions. The difference between the hindcast-RIC and hindcast-CIC simulations decreases as

lead month increases, and the accumulated influence of meteorological forcing gradually overwhelms the impact of the initial condition. In the Indochinese Peninsula, India, and surrounding areas, where precipitation is relatively intense in summer seasons, the influence of initial condition on the forecast skill is eliminated after 2 to 3 months. The correlation of precipitation dominates the prediction skills of RZSM when such a considerable precipitation influence emerges. When precipitation is low, however, the prediction skill of RZSM depends on the RZSM prediction in the previous month, with relatively less influence from other meteorological forcing variables. This pattern becomes particularly im-

portant in regions where the initialization dates (i.e., 1 May) of hindcast-RIC simulations are around the beginning of dry seasons. For example, in the Helmand basin, where the precipitation is mainly from mountain precipitation during the winter and early spring months (see Sect. 2.1), the accurate initial conditions help the prediction skills remain statistically significant for 8 months.

We also compared the surface soil moisture (SSM) from retrospective simulations and hindcast-RIC simulations with the ESA-CCI SM data product in the five major river basins in the study domain. The comparison shows that the retrospective simulations capture inter-annual variability in most of the months within the five basins. The SSM hindcasts in four out of five major basins (all but the Ganges) generally have a high correlation with ESA-CCI SM data in the first 1 or 2 months. The correlation then decreases for another 1 or 2 months, then increases again and reaches a second maximum after the peak monsoon rains (for which precipitation forecast skill is limited) before decreasing for the rest of the forecast period.

The prediction skill of the forecasting system of SAHFS-S2S depends on the land surface model, initial conditions, and meteorological forecasts. In this study, an accurate initial condition has been shown to have a positive contribution for prediction skills over much of the simulation domain, and particularly in dry areas and seasons. Future effort should be made to improve the accuracy of initial conditions estimated by the land surface model. This could include (1) assimilating ground or satellite-based observations into the land surface model (Getirana et al., 2020a, c; Wanders et al., 2014) and (2) better representation of anthropogenic influences, for example, irrigation (Nie et al., 2019) and reservoirs (Wanders and Wada, 2015; Getirana et al., 2020b), in the land surface model or hydrological model. Errors in the meteorological forecast are also a clear limitation on the forecast skill. Improved dynamically based S2S meteorological forecast systems (Pegion et al., 2019), advanced statistical-dynamical forecast methods (Madadgar et al., 2016; Shukla et al., 2014), and improved bias correction and downscaling methods (Rodrigues et al., 2018) are all areas of significant research effort. The results presented here show that current capabilities offer meaningful skill over shorter time horizons for much of the domain, and also that performance can be improved as each component of the forecast system improves.

Due to the difficulty of acquiring reliable, long-term streamflow observations, it is difficult to evaluate the streamflow monitoring and forecasting in this study domain. However, Yang et al. (2011) evaluated Noah-MP forced by the Global Land Data Assimilation System (GLDAS) in the Mekong and Ganges basins in this study domain and found that Noah-MP captured the seasonality of the streamflow in both basins and magnitude in the Mekong basin but underestimated the magnitude in the Ganges basin. Ghatak et al. (2018) found that the Noah LSM forced by GDAS and CHIRPS captured the timing of a 1-year flood event in the Indus basin but underestimated the magnitude and captured both seasonal cycles and magnitude of streamflow in the Kosi basin within the Ganges basin. The streamflow from the SAHFS-S2S monitoring system has been found to have similar results to those in Ghatak et al. (2018).

The SAHFS-S2S has been operationally implemented at the International Centre for Integrated Mountain Development (ICIMOD). The monitoring system updates about every 10 days and the forecast system launches a 9-month hydrological forecast at the beginning of each month. The outputs of SAHFS-S2S have been integrated to ICIMOD's Regional Drought Monitoring and Outlook System (RDMOS), which focuses on crop and drought conditions mainly within countries in the Hindu Kush and Himalaya regions. This integrated system is designed to support local decision-makers and government agencies concerned with food security and drought preparation.

Review statement. This paper was edited by Xing Yuan and reviewed by two anonymous referees.

References

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., and Pappenberger, F.: GloFAS – global ensemble streamflow forecasting and flood early warning, Hydrol. Earth Syst. Sci., 17, 1161–1175, https://doi.org/10.5194/hess-17-1161-2013, 2013.

Arsenault, K. R., Kumar, S. V., Geiger, J. V., Wang, S., Kemp, E., Mocko, D. M., Beaudoing, H. K., Getirana, A., Navari, M., Li, B., Jacob, J., Wegiel, J., and Peters-Lidard, C. D.: The Land surface Data Toolkit (LDT v7.2) – a data fusion environment for land data assimilation systems, Geosci. Model Dev., 11, 3605–3621, https://doi.org/10.5194/gmd-11-3605-2018, 2018.

Barros, V. R. and Field, C. B.: Climate change 2014: impacts, adaptation, and vulnerability. Part B: regional aspects, Cambridge University Press, Cambridge, UK, 2014.

Bell, V. A., Davies, H. N., Kay, A. L., Brookshaw, A., and Scaife, A. A.: A national-scale seasonal hydrological forecast system: development and evaluation over Britain, Hydrol. Earth Syst. Sci., 21, 4681–4691, https://doi.org/10.5194/hess-21-4681-2017, 2017.

Borovikov, A., Cullather, R., Kovach, R., Marshak, J., Vernieres, G., Vikhliaev, Y., Zhao, B., and Li, Z.: GEOS-5 seasonal forecast system, Clim. Dynam., 53, 7335–7361, 2019.

Cai, X., Yang, Z. L., David, C. H., Niu, G. Y., and Rodell, M.: Hydrological evaluation of the Noah-MP land surface model for the Mississippi River Basin, J. Geophys. Res.-Atmos., 119, 23–38, 2014.

Chen, G., Yang, Y., Yang, Z., Xie, J., Guo, J., Gao, R., Yin, Y., and Robinson, D.: Accelerated soil carbon turnover under tree plantations limits soil carbon storage, Sci. Rep., 6, 19693, https://doi.org/10.1038/srep19693, 2016.

Csiszar, I. and Gutman, G.: Mapping global land surface albedo from NOAA AVHRR, J. Geophys. Res.-Atmos., 104, 6215–6228, 1999.

de Andrade, F. M., Coelho, C. A., and Cavalcanti, I. F.: Global precipitation hindcast quality assessment of the Subseasonal to Seasonal (S2S) prediction project models, Clim. Dynam., 52, 5451–5475, 2019.

Ek, M., Mitchell, K., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res.-Atmos., 108, 8851, https://doi.org/10.1029/2002JD003296, 2003.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., and Roth, L.: The shuttle radar topography mission, Rev. Geophys., 45, RG2004 https://doi.org/10.1029/2005RG000183, 2007.

Fowler, H. and Archer, D.: Conflicting signals of climatic change in the Upper Indus Basin, J. Climate, 19, 4276–4293, 2006.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X.: MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, Remote Sens. Environ., 114, 168–182, 2010.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., and Hoell, A.: The climate hazards infrared precipitation with stations – a new environmental record for monitoring extremes, Scientific Data, 2, 1–21, 2015.

Getirana, A., Jung, H. C., Arsenault, K., Shukla, S., Kumar, S., Peters-Lidard, C., Maigari, I., and Mamane, B.: Satellite gravimetry improves seasonal streamflow forecast initialization in Africa, Water Resour. Res., 56, e2019WR026259, https://doi.org/10.1029/2019WR026259, 2020a.

Getirana, A., Jung, H. C., Van Den Hoek, J., and Ndehedehe, C. E.: Hydropower dam operation strongly controls Lake Victoria's freshwater storage variability, Sci. Total Environ., 726, 138343, https://doi.org/10.1016/j.scitotenv.2020.138343, 2020b.

Getirana, A., Rodell, M., Kumar, S., Beaudoing, H. K., Arsenault, K., Zaitchik, B., Save, H., and Bettadpur, S.: GRACE Improves Seasonal Groundwater Forecast Initialization over the United States, J. Hydrometeorol., 21, 59–71, 2020c.

Getirana, A. C., Bonnet, M.-P., Calmant, S., Roux, E., Rotunno Filho, O. C., and Mansur, W. J.: Hydrological monitoring of poorly gauged basins based on rainfall-runoff modeling and spatial altimetry, J. Hydrol., 379, 205–219, 2009.

Ghatak, D., Zaitchik, B., Kumar, S., Matin, M., Bajracharya, B., Hain, C., and Anderson, M.: Influence of Precipitation Forcing Uncertainty on Hydrological Simulations with the NASA South Asia Land Data Assimilation System, Hydrology, 5, 57, https://doi.org/10.3390/hydrology5040057, 2018.

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., and Dorigo, W.: Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology, Earth Syst. Sci. Data, 11, 717–739, https://doi.org/10.5194/essd-11-717-2019, 2019.

Gutman, G. and Ignatov, A.: The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models, Int. J. Remote Sens., 19, 1533–1543, 1998.

Gutmann, E. D., Hamman, J. J., Clark, M. P., Eidhammer, T., Wood, A. W., Arnold, J. R., and Nowak, K.: Evaluating the effect of regional climate inference methodologies in a common framework, in preparation, 2020.

Hao, Z., Yuan, X., Xia, Y., Hao, F., and Singh, V. P.: An overview of drought monitoring and prediction systems at regional and global scales, B. Am. Meteorol. Soc., 98, 1879–1896, 2017.

Hatfield, J. L., Boote, K. J., Kimball, B., Ziska, L., Izaurralde, R. C., Ort, D., Thomson, A. M., and Wolfe, D.: Climate impacts on agriculture: implications for crop production, Agron. J., 103, 351–370, 2011.

ICIMOD: Regional Drought Monitoring and Outlook System for South Asia, available at: http://tethys.icimod.org/apps/regionaldrought/current/, last access: 17 December 2020.

Immerzeel, W.: Historical trends and future predictions of climate variability in the Brahmaputra basin, International Journal of Climatology, Q. J. Roy. Meteorol. Soc., 28, 243–254, 2008.

Jie, W., Vitart, F., Wu, T., and Liu, X.: Simulations of Asian Summer Monsoon in the Sub-seasonal to Seasonal Prediction Project

(S2S) database, Q. J. Roy. Meteorol. Soc., 143, 2282–2295, https://doi.org/10.1002/qj.3085, 2017.

Koster, R. D., Suarez, M. J., Liu, P., Jambor, U., Berg, A., Kistler, M., Reichle, R., Rodell, M., and Famiglietti, J.: Realistic initialization of land surface states: Impacts on subseasonal forecast skill, J. Hydrometeorol., 5, 1049–1063, 2004.

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

Kumar, S., Mocko, D., Vuyovich, C., and Peters-Lidard, C.: Impact of Surface Albedo Assimilation on Snow Estimation, Remote Sens., 12, 645, https://doi.org/10.3390/rs12040645, 2020.

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., Houser, P. R., Geiger, J., Olden, S., Lighty, L., Eastman, J. L., Doty, B., and Dirmeyer, P.: Land information system: An interoperable framework for high resolution land surface modeling, Environ. Modell. Softw., 21, 1402–1415, 2006.

Kumar, S. V., Mocko, M. D., Wang, S., Peters-Lidard, C. D., and Borak, J.: Assimilation of Remotely Sensed Leaf Area Index into the Noah-MP Land Surface Model: Impacts on Water and Carbon Fluxes and States over the Continental United States, J. Hydrometeorol., 20, 1359–1377, 2019.

Livneh, B., Xia, Y., Mitchell, K. E., Ek, M. B., and Lettenmaier, D. P.: Noah LSM snow model diagnostics and enhancements, J. Hydrometeorol., 11, 721–738, 2010.

Luo, L., Sheffield, J., and Wood, E.: Towards a Global Drought Monitoring and Forecasting Capability, in: 33rd NOAA Annual Climate Diagnostics and Prediction Workshop, 20–24 October 2008, Lincoln, Niger, 2008.

Ma, F., Luo, L., Ye, A., and Duan, Q.: Seasonal drought predictability and forecast skill in the semi-arid endorheic Heihe River basin in northwestern China, Hydrol. Earth Syst. Sci., 22, 5697–5709, https://doi.org/10.5194/hess-22-5697-2018, 2018.

Madadgar, S., AghaKouchak, A., Shukla, S., Wood, A. W., Cheng, L., Hsu, K. L., and Svoboda, M.: A hybrid statistical-dynamical framework for meteorological drought prediction: Application to the southwestern United States, Water Resour. Res., 52, 5095–5110, 2016.

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Cosgrove, B. A., Sheffield, J., Duan, Q., and Luo, L.: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, J. Geophys. Res.-Atmos., 109, 1–32, https://doi.org/10.1029/2003JD003823, 2004.

Molod, A., Hackert, E., Vikhliaev, Y., Zhao, B., Barahona, D., Vernieres, G., Borovikov, A., Kovach, R. M., Marshak, J., and Schubert, S.: GEOS-S2S Version 2: The GMAO High-Resolution Coupled Model and Assimilation System for Seasonal Prediction, J. Geophys. Res.-Atmos., 125, e2019JD031767, https://doi.org/10.1029/2019JD031767, 2020.

National Climatic Data Center: NCEP EDAS and GDAS (FNL) Model Data (DSI-6141), NESDIS, NOAA, https://doi.org/10.5065/D65Q4T4Z, 2020.

Nie, W., Zaitchik, B. F., Rodell, M., Kumar, S. V., Anderson, M. C., and Hain, C.: Groundwater withdrawals under drought: Reconciling GRACE and land surface models in the United States High Plains Aquifer, Water Resour. Res., 54, 5282–5299, 2018.

Nie, W., Zaitchik, B. F., Rodell, M., Kumar, S. V., Arsenault, K. R., Li, B., and Getirana, A.: Assimilating GRACE into a Land Surface Model in the presence of an irrigation-induced groundwater trend, Water Resour. Res., 55, 11274–11294, https://doi.org/10.1029/2019WR025363, 2019.

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., and Rosero, E.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, J. Geophys. Res.-Atmos., 116, D12109, https://doi.org/10.1029/2010JD015139, 2011.

Pegion, K., Kirtman, B. P., Becker, E., Collins, D. C., LaJoie, E., Burgman, R., Bell, R., DelSole, T., Min, D., and Zhu, Y.: The Subseasonal Experiment (SubX): A Multimodel Subseasonal Prediction Experiment, B. Am. Meteorol. Soc., 100, 2043–2060, 2019.

Qian, X., Qiu, B., and Zhang, Y.: Widespread decline in vegetation photosynthesis in Southeast Asia due to the prolonged drought during the 2015/2016 El Niño, Remote Sens., 11, 910, https://doi.org/10.3390/rs11080910, 2019.

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., and Bosilovich, M.: The global land data assimilation system, B. Am. Meteorol. Soc., 85, 381–394, 2004.

Rodrigues, E. R., Oliveira, I., Cunha, R., and Netto, M.: DeepDownscale: a deep learning strategy for high-resolution weather forecast, Proceedings of the IEEE 14th International Conference on e-Science, 29 October–1 November, Amsterdam, Netherlands, 415–422, 2018.

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E. F., and Marx, A.: Anthropogenic warming exacerbates European soil moisture droughts, Nat. Clim. Change, 8, 421–426, 2018.

Samaniego, L., Thober, S., Wanders, N., Pan, M., Rakovec, O., Sheffield, J., Wood, E. F., Prudhomme, C., Rees, G., and Houghton-Carr, H.: Hydrological forecasts and projections for improved decision-making in the water sector in Europe, B. Am. Meteorol. Soc., 100, 2451–2472, 2019.

Seck, A., Welty, C., and Maxwell, R. M.: Spin-up behavior and effects of initial conditions for an integrated hydrologic model, Water Resour. Res., 51, 2188–2210, 2015.

Shah, R., Sahai, A. K., and Mishra, V.: Short to sub-seasonal hydrologic forecast to manage water and agricultural resources in India, Hydrol. Earth Syst. Sci., 21, 707–720, https://doi.org/10.5194/hess-21-707-2017, 2017.

Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., and Demuth, S.: A drought monitoring and forecasting system for sub-Sahara African water resources and food security, B. Am. Meteorol. Soc., 95, 861–882, 2014.

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

Shukla, S., Funk, C., and Hoell, A.: Using constructed analogs to improve the skill of National Multi-Model Ensemble March–April–May precipitation forecasts in equatorial East Africa,

Environ. Res. Lett., 9, 094009, https://doi.org/10.1088/1748-9326/9/9/094009, 2014.

Sivakumar, M. V. and Stefanski, R.: Climate change in South Asia, in: Climate change and food security in South Asia, edited by: Lal, R., Mannava, V. K., Sivakumar, S. M. A., Faiz, A. H. M, Rahman, M., and Islam, K. R., Springer, London, England, 13–30, 2010.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., and Stooksbury, D.: The drought monitor, B. Am. Meteorol. Soc., 83, 1181–1190, 2002.

Syaukat, Y.: Irrigation in Southern and Eastern Asia in figures AQUASTAT Survey-2011, FAO the United Nation, Rome, Italy, 2012.

Van Der Schrier, G., Klein Tank, A. M., Van Den Besselaar, E. J., and Swarinoto, Y.: Observed trends and variability in climate indices relevant for crop yields in Southeast Asia, J. Climate, 29, 2651–2669, 2016.

Wanders, N. and Van Lanen, H. A. J.: Future discharge drought across climate regions around the world modelled with a synthetic hydrological modelling approach forced by three general circulation models, Nat. Hazards Earth Syst. Sci., 15, 487–504, https://doi.org/10.5194/nhess-15-487-2015, 2015.

Wanders, N. and Wada, Y.: Human and climate impacts on the 21st century hydrological drought, J. Hydrol., 526, 208–220, 2015.

Wanders, N., Karssenberg, D., de Roo, A., de Jong, S. M., and Bierkens, M. F. P.: The suitability of remotely sensed soil moisture for improving operational flood forecasting, Hydrol. Earth Syst. Sci., 18, 2343–2357, https://doi.org/10.5194/hess-18-2343-2014, 2014.

Wanders, N., Thober, S., Kumar, R., Pan, M., Sheffield, J., Samaniego, L., and Wood, E. F.: Development and evaluation of a pan-European multimodel seasonal hydrological forecasting system, J. Hydrometeorol., 20, 99–115, 2019.

Whitney, J. W.: Geology, water, and wind in the lower Helmand Basin, Southern Afghanistan U.S. Geological Survey, Reston, Virginia, USA, 2006.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., and Meng, J.: 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, J. Geophys. Res.-Atmos., 117, D03109, https://doi.org/10.1029/2011JD016048, 2012.

Xue, Y., Houser, P. R., Maggioni, V., Mei, Y., Kumar, S. V., and Yoon, Y.: Assimilation of Satellite-Based Snow Cover and Freeze/Thaw Observations Over High Mountain Asia, Front. Earth Sci., 7, 115, https://doi.org/10.3389/feart.2019.00115, 2019.

Yang, Z. L., Niu, G. Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., and Tewari, M.: The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins, J. Geophys. Res.-Atmos., 116, D12110, https://doi.org/10.1029/2010JD015140, 2011.

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

Yuan, X., Wood, E. F., Roundy, J. K., and Pan, M.: CFSv2-based seasonal hydroclimatic forecasts over the conterminous United States, J. Climate, 26, 4828–4847, 2013.

Yuan, X., Wood, E. F., and Liang, M.: Integrating weather and climate prediction: Toward seamless hydrologic forecasting, Geophys. Res. Lett., 41, 5891–5896, 2014.

Yuan, X., Wood, E. F., and Ma, Z.: A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development, WiRes. Water, 2, 523–536, 2015.

Yuan, X., Ma, F., Wang, L., Zheng, Z., Ma, Z., Ye, A., and Peng, S.: An experimental seasonal hydrological forecasting system over the Yellow River basin – Part 1: Understanding the role of initial hydrological conditions, Hydrol. Earth Syst. Sci., 20, 2437–2451, https://doi.org/10.5194/hess-20-2437-2016, 2016.

Zhou, Y., Zaitchik, B. F.., Kumar, S. V., Arsenault, K. R., and Zamora, R. A.: "Data associated with publication: Developing a hydrological monitoring and sub-seasonal to seasonal forecasting system for South and Southeast Asian river basins", https://doi.org/10.7281/T1/JYAHTN, Johns Hopkins University Data Archive, V1, 2020.