

# **A Probabilistic Drought Forecasting Framework: A Combined Dynamical and Statistical Approach**

Hongxiang Yan, Hamid Moradkhani, and Mahkameh Zarekarizi

Remote Sensing and Water Resources Lab,

Department of Civil and Environmental Engineering, Portland State University

## **Abstract**

In order to improve drought forecasting skill, this study develops a probabilistic drought forecasting framework comprised of dynamical and statistical modeling components. The novelty of this study is to seek the use of data assimilation to quantify initial condition uncertainty with the Monte Carlo ensemble members, rather than relying entirely on the hydrologic model or land surface model to generate a single deterministic initial condition, as currently implemented in the operational drought forecasting systems. Next, the initial condition uncertainty is quantified through data assimilation and coupled with a newly developed probabilistic drought forecasting model using a copula function. The initial condition at each forecast start date are sampled from the data assimilation ensembles for forecast initialization. Finally, seasonal drought forecasting products are generated with the updated initial conditions. This study introduces the theory behind the proposed drought forecasting system, with an application in Columbia River Basin, Pacific Northwest, United States. Results from both synthetic and real case studies suggest that the proposed drought forecasting system significantly improves the seasonal drought forecasting skills and can facilitate the state drought preparation and declaration, at least three months before the official state drought declaration.

**Keywords:** Probabilistic drought forecasting, Data assimilation, Initial condition uncertainty, Particle Markov chain Monte Carlo

## **1 Introduction**

Drought is a naturally occurring climate phenomenon driven by extreme macroclimatic variability originating from atmospheric interactions and feedback between the atmosphere, the oceans, and the land surface (Ahmadalipour et al., 2017; Keyantash and Dracup, 2002; Mishra and Singh, 2010; Yan, 2016). Contrary to other natural disasters such as floods, hurricanes, and tornados, droughts develop slowly over large spatial extents and last for multiple years. It is usually difficult to identify droughts until severe damage has already occurred (Luo and Wood, 2007; Mishra and Singh, 2011). Moreover, drought is the most costly extreme event among all natural disasters (NCDC, 2012; Ross and Lott, 2003; Sheffield et al., 2014). For instance, Howitt et al. (2014) estimated that the recent California (CA) extreme drought resulted in an economic loss of \$2.7 billion in 2014; Washington Department of Agriculture (2015) calculated that the economic loss of the 2015 state drought was more than \$335 million.

In order to mitigate the extensive negative socio-economic impacts, comprehensive preparation and effective response to drought is necessary. A drought forecasting system which provides forecasts of potential drought conditions in a timely manner is critical to stakeholders and decision-makers. In the U.S., there are few federal drought forecasting systems, including the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center's (CPC's) Seasonal Drought Outlook (SDO) (Steinemann, 2006) and the National Integrated Drought Information System's (NIDIS's) Drought Early Warning Systems (DEWS). Despite these efforts, the World Climate Research Program still recognizes that seasonal drought forecasting is one of the major research gaps in hydroclimatology (WCRP, 2010).

Drought forecasting generally relies on drought indicators computed by either dynamical model simulations or statistical methods (DeChant and Moradkhani, 2014; Kumar et al., 2014b;

Mao et al., 2015; Wood and Lettenmaier, 2008). One option to generate hydrologic prediction is the Ensemble Streamflow Prediction (ESP) framework (Day, 1985). The ESP is generated by deriving the hydrologic modeling using resampled historical meteorological variables and future drought condition is estimated accordingly. However, the sampled forcing data is not necessarily representative of the non-stationary climate (Milly et al., 2008; Yan and Edwards, 2013). Another option is to use dynamical forecasting products, such as the Climate Forecast System Version 2 (CFSv2) products (Yuan et al., 2011) and the North American Multimodel Ensemble (NMME) climate forecasts (Yuan, 2016). While dynamical forecast provides future climate information, the precipitation forecasts are subject to high uncertainty and exhibit low skill beyond one month lead time (Hayes et al., 2005; Lavers et al., 2009; Yuan et al., 2013). As an alternative, a recently developed probabilistic drought forecasting approach built on copula functions (Madadgar and Moradkhani, 2013) could be used. Recent studies have shown that such a statistical approach could lead to good seasonal drought forecasting skill (Chen et al., 2015; Madadgar and Moradkhani, 2016, 2014; Mishra and Desai, 2005).

Recent works show that seasonal hydrologic forecast skill is mostly controlled by hydrologic initial condition (DeChant and Moradkhani, 2015, 2011; Koster et al., 2010; Li et al., 2009; Shukla and Lettenmaier, 2011; Wood et al., 2016; Wood and Lettenmaier, 2008; Yossef et al., 2013; Yuan et al., 2016). For instance, the initial condition refers to the root-zone soil moisture for agricultural drought forecasting. Initial condition uncertainty arises from the chaotic properties of the Earth system and is unavoidable due to inability to accurately observe land surface states (Stainforth et al., 2005). In the western U.S., the accurate estimation of initial condition uncertainty is particularly important due to ongoing coverage of drought in California and Pacific Northwest. However, the initial condition uncertainty is not considered in the

aforementioned drought forecasting systems, where a single deterministic hydrologic initial condition is utilized. Therefore, in order to improve drought forecasting skill, this study seeks the use of ensemble data assimilation (DA) to improve land initialization by quantifying the initial condition uncertainty (DeChant and Moradkhani, 2011; Yan et al., 2015; Yan and Moradkhani, 2016).

In summary, the scope of this study is to develop a probabilistic drought forecasting system, with the use of DA and a copula-based probabilistic drought forecasting method. It is hypothesized that an accurate quantification of uncertainty in hydrologic initial condition would lead to a better drought forecasting skill. The paper is organized as follows. Section 2 describes the framework of the proposed probabilistic drought forecasting system, including the dynamical hydrologic modeling, the DA implementation, and the statistical modeling using copula functions to generate the probabilistic forecasts. Section 3 applies the proposed forecasting system with a case study in the Columbia River Basin, Pacific Northwest. Finally, section 4 provides the conclusion.

## **2 Probabilistic Drought Forecasting System Framework**

The framework of the proposed probabilistic drought forecasting system is described here. The system is composed of three main components: (1) hydrologic modeling, (2) data assimilation, and (3) a copula-based probabilistic drought forecasting model. The hydrologic model is first calibrated for the study region. Then the copula-based probabilistic drought forecasting model is developed based on the long-term open-loop (OL) simulations that is simulation with no assimilation, of hydrologic observations. Next, the DA system is used to assimilate the remotely sensed and in-situ observations to improve hydrologic simulation and characterize the initial

condition uncertainty by estimating the probability density function (PDF) of the initial conditions. Last, the initial conditions sampled from the PDF are used in the copula-based probabilistic drought forecasting model to generate seasonal drought forecasts. Figure 1 displays the proposed drought forecasting framework.

-----

Please place Figure 1 here

-----

## **2.1 Hydrologic Modeling**

The Precipitation-Runoff Modeling System (PRMS) (Leavesley et al., 1983) is used in this study to model the land surface states. The PRMS is a modular deterministic, distributed-parameter, and physical-process watershed model (Markstrom et al., 2008). The land surface hydrologic process simulated by PRMS includes the evapotranspiration, runoff, infiltration, interflow, snowpack, and soil moisture. Instead of delineating a watershed into uniform grid cells, the PRMS partitions a watershed into hydrologic response units (HRUs) that are based on the physical attributes of the watershed such as land-surface elevation, slope and aspect, vegetation type, soil type, and spatiotemporal climate patterns (Markstrom et al., 2015). The physical attributes and hydrologic response of each HRU are assumed to be homogeneous. The meteorological forcing data for PRMS are precipitation, minimum temperature and maximum temperature. Excess runoff is routed to the outlet through the cascade and Muskingum routing methods.

The PRMS version 4.0.1 (PRMS-IV) released on March 15, 2015 was used in this study. The PRMS-IV takes soil moisture into account in three reservoir systems: the preferential-flow reservoir, the capillary reservoir, and the gravity reservoir. The preferential-flow reservoir represents the water content between preferential-flow threshold and total soil saturation; it is available for fast interflow and Dunnian surface runoff. The capillary reservoir represents the moisture content between wilting point and field capacity; it is only available for evapotranspiration and not for drainage. The gravity reservoir is limited to the water content between field capacity and preferential-flow threshold. The water content in this reservoir is available for slow interflow, groundwater recharge, and Dunnian surface runoff (Markstrom et al., 2015).

## **2.2 Data Assimilation Algorithm**

The ensemble DA system is used to quantify the initial condition uncertainty. The recently developed particle filter Markov chain Monte Carlo (PMCMC) (Moradkhani et al., 2012) is used in this study. The PMCMC is an extension of the particle filter with sampling importance resampling (PF-SIR) (Moradkhani et al., 2005) and uses the PF-SIR to design efficient high-dimensional proposal distributions for MCMC algorithm. Despite the reasonable performance of the EnKF, the PF was found to be more robust since it can relax the Gaussian error assumption, preserve water balance, and provide a more complete representation of state posterior distribution for a nonlinear non-Gaussian hydrologic system (DeChant and Moradkhani, 2012; Dong et al., 2015; Plaza et al., 2012; Yan and Moradkhani, 2016).

### 2.2.1 PF-SIR Algorithm

Following Moradkhani (2008), the state-space models that describe the generic non-linear earth system are as follows:

$$y_t = h(x_t) + v_t \quad (1)$$

$$x_t = f(x_{t-1}, u_t, \theta) + w_t \quad (2)$$

where  $x_t \in \mathbb{R}^n$  is a vector of the uncertain state variables at current time step,  $y_t \in \mathbb{R}^m$  is a vector of observation data,  $u_t$  is the uncertain forcing data,  $\theta \in \mathbb{R}^d$  is a vector of model parameters,  $h(\cdot)$  is a non-linear function that relates the states  $x_t$  to the observations  $y_t$ ,  $w_t$  represents the model error, and  $v_t$  indicates the observation error. The errors  $w_t$  and  $v_t$  are assumed to be white noise with mean zero and covariance  $Q_t$  and  $R_t$ , respectively.

According to Moradkhani et al. (2005), the posterior distribution of the state variables  $x_t$  given a realization of the observations  $y_{1:t}$  is written as follows:

$$p(x_t|y_{1:t}) = p(x_t|y_{1:t-1}, y_t) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} \quad (3)$$

$$= \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{\int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t}$$

$$p(x_t|y_{1:t-1}) = \int p(x_t, x_{t-1}|y_{1:t-1})dx_{t-1} = \quad (4)$$

$$\int p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1}$$

where  $p(y_t|x_t)$  is the likelihood,  $p(x_t|y_{1:t-1})$  is the prior distribution, and  $p(y_t|y_{1:t-1})$  is the normalization factor. The marginal likelihood  $p(y_{1:t})$  can be computed as:

$$p(y_{1:t}) = p(y_1) \prod p(y_t|y_{1:t-1}) \quad (5)$$

where the normalization factor  $p(y_t|y_{1:t-1})$  is:

$$p(y_t|y_{1:t-1}) = \int p(y_t, x_t|y_{1:t-1})dx_t = \int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t \quad (6)$$

In practice, Equation (3) does not have an analytical solution except for few special cases (e.g., the linear system with Gaussian assumption). Instead, the posterior distribution  $p(x_t|y_{1:t})$  is approximated using a set of Monte Carlo (MC) random samples as:

$$p(x_t|y_{1:t}) \approx \sum_{i=1}^N w_t^{i+} \delta(x_t - x_t^i) \quad (7)$$

where  $w_t^{i+}$  is the posterior weight of the  $i$ th particle,  $\delta$  is the Dirac delta function, and  $N$  is the ensemble size. The normalized weights are calculated as follows:

$$w_t^{i+} = w_t^{i-} \frac{p(y_t|x_t^i)}{\sum_{i=1}^N w_t^{i-} p(y_t|x_t^i)} \quad (8)$$

where  $w_t^{i-}$  is the prior particle weights, and  $p(y_t|x_t^i)$  can be computed from the likelihood  $L(y_t|x_t^i)$ . Generally, a Gaussian distribution is used to estimate  $L(y_t|x_t^i)$ :

$$L(y_t|x_t^i) = \frac{1}{\sqrt{(2\pi)^m |R_t|}} \exp \left[ -\frac{1}{2} (v_t^i)^T R_t^{-1} (v_t^i) \right] \quad (9)$$

where  $v_t^i = y_t - h(x_t^i)$  is the residual. The estimate of the normalization factor  $p(y_t|y_{1:t-1})$  in Equation (6) is given by:

$$p(y_t|y_{1:t-1}) \approx \sum_{i=1}^N w_t^{i-} p(y_t|x_t^i) = \sum_{i=1}^N w_t^{i-} L(y_t|x_t^i) \quad (10)$$

To obtain approximate samples from  $p(x_t|y_{1:t})$ , a resampling operation is necessary to reduce the weight degeneration problem where all but one of the importance weights are close to zero. The SIR algorithm suggests resampling the particles with a probability greater than the uniform probability. After resampling, all the particle weights are set equal to  $1/N$ . Based on the above descriptions, the PF-SIR algorithm can thus be summarized as follows:

At time  $t = 1$ ,



- Initialize the system by sampling the model states  $x_t^i, i = 1, \dots, N$  from a uniform distribution
- Assign the particle weights uniformly  $w_t^{i+} = w_{t+1}^{i-} = 1/N, i = 1, \dots, N$

At time  $t \geq 2$ ,

- Propagate the  $N$  model states forward in time  $x_t^i$  through model operator
- Update the particle weights  $w_t^{i+} = w_t^{i-} \frac{p(y_t|x_t^i)}{\sum_{i=1}^N w_t^{i-} p(y_t|x_t^i)}$
- Resample to obtain  $N$  new equally-weighted particles  $x_t^i$  with  $w_t^{i+} = 1/N, i = 1, \dots, N$

### 2.2.2 PMCMC Algorithm

The PMCMC is an extension of the particle filter with sampling importance resampling (PF-SIR) and uses the PF-SIR to design efficient high-dimensional proposal distributions  $p(x_t|y_{1:t})$  for MCMC algorithm. The PMCMC has the following three steps:

(1) Initialization ( $j = 0$ ): run PF-SIR targeting  $p(x_t|y_{1:t})$ , sample  $X_t(0) \sim p(x_t|y_{1:t})$  and let  $p(y_{1:t})(0)$  denote the corresponding marginal likelihood estimate.

(2) Iteration ( $j \geq 1$ ): sample  $X_t^* \sim p(x_t|y_{1:t})$  again and let  $p(y_{1:t})^*$  denote the corresponding marginal likelihood estimate. (3) Calculate the acceptance ratio as:

$$\min \left\{ 1, \frac{p(y_{1:t})^*}{p(y_{1:t})(j-1)} \right\} \quad (11)$$

and set  $X_t(j) = X_t^*$  and  $p(y_{1:t})(j) = p(y_{1:t})^*$ ; otherwise set  $X_t(j) = X_t(j-1)$  and  $p(y_{1:t})(j) = p(y_{1:t})(j-1)$ . The marginal likelihood is estimated based on the Equations (5) and (10).

### 2.3 Copula-based Probabilistic Drought Forecasting Model

A newly developed probabilistic drought forecasting method based on copula functions (Madadgar and Moradkhani, 2014, 2013) is used here which is coupled with the initial condition uncertainty through DA. Since drought indicators are usually statistically dependent on their past status, they can be expressed within the Bayesian networks, which are capable of describing the conditional dependencies of two random variables. The core of this forecasting model is to apply a multivariate distribution function to forecast future drought conditions given the current or past drought status.

Copulas are multivariate distribution functions on the  $n$ -dimensional unit cube. The variables of copula functions are uniformly distributed with  $[0, 1]$ . Following Sklar's theorem (Sklar, 1959), a multivariate distribution  $P(x_1, \dots, x_n)$  can be expressed by copula functions as follows:

$$P(x_1, \dots, x_n) = C[P(x_1), \dots, P(x_n)] = C(u_1, \dots, u_n) \quad (12)$$

where  $C$  is the cumulative distribution function (CDF) of copula, and  $P(x_i)$  is the marginal distribution of  $x_i$  being uniform with  $[0, 1]$ , which is denoted by  $u_i$ . If the  $C[P(x_1), \dots, P(x_n)]$  is absolutely continuous, its PDF can be written as:

$$c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n} \quad (13)$$

Using the PDF of the copula, the joint PDF of  $p(x_1, \dots, x_n)$  can be expressed as:

$$p(x_1, \dots, x_n) = c(u_1, \dots, u_n) \prod_{i=1}^n p(x_i) \quad (14)$$

For the drought forecasting application, only two random variables are modeling ( $n = 2$ ). As a result, the conditional probability distribution of  $x_{t+1}$  given  $x_t$  is written as:

$$p(x_{t+1}|x_t) = \frac{p(x_{t+1}, x_t)}{p(x_t)} \quad (15)$$

Replacing the joint PDF  $p(x_{t+1}, x_t)$  in Equation (15) with Equation (14), the conditional probability of Equation (15) can be revised as:

$$p(x_{t+1}|x_t) = \frac{p(x_{t+1}, x_t)}{p(x_t)} = \frac{c(u_{t+1}, u_t)p(x_{t+1})p(x_t)}{p(x_t)} = c(u_{t+1}, u_t)p(x_{t+1}) \quad (16)$$

Using Equation (16), the probabilistic forecast of drought states in time  $t + 1$  given the drought condition at time  $t$  can be examined. The mode of the  $p(x_{t+1}|x_t)$  is correspondent to the maximum likelihood estimation (MLE) of  $x_{t+1}$ . For seasonal drought forecasting, the  $x_{t+1}$  represents the drought condition in the coming season, and  $x_t$  indicates the initial drought condition in the current season. Following Madadgar and Moradkhani (2013), to build the conditional PDF of Equation (16), the Monte Carlo (MC) approach is used to sample from the copula density function  $c(u_{t+1}, u_t)$ . Different copula functions can be used to join the marginal distributions of correlated and dependent variables. Among various copula families, the Archimedean and elliptical families are usually used in hydrological applications (Madadgar and Moradkhani, 2016). The Gaussian and  $t$  copulas from the elliptical family, and the Gumbel and Clayton from the Archimedean family, can be tested with parameters estimated by the inference function for margins (IFM) method (Joe, 1997). To select the appropriate copula function among the various copula functions, a goodness-of-fit (GOF) test can be used, such as the parametric bootstrapping GOF test (Genest and Rémillard, 2008).

Based on the descriptions in Sections 2.1-2.3, the proposed dynamical-statistical drought forecasting system can thus be summarized as follows:

At each drought forecasting initialization date  $t$ ,

- Download the forcing data until the initialization date

- Run the dynamical hydrologic model to estimate long-term retrospective root-zone soil moisture until the initialization date (modeling period: >30 years)
- Develop the copula-based statistical forecasting model based on the long-term retrospective dataset
- Download the remotely sensed surface soil moisture until the initialization date
- Run the DA system to assimilate remotely sensed surface soil moisture until the initialization date (assimilation period: several months)
- Quantify the initial condition uncertainty in terms of a PDF through the ensemble members from DA
- Sample initial condition from the PDF and input into the copula-based forecasting model
- Generate probabilistic drought forecasts

## **2.4 Seasonal to Inter-annual Drought Forecasting Framework**

Based on Figure 2, the proposed probabilistic drought forecasting system has three components: (1) *Seasonal drought forecasting*. This is the official seasonal drought forecasting product for decision-makers and water resources managers. The system generates forecasts for the current season based on last season and is performed on 1-January (winter forecasting), 1-April (spring forecasting), 1-July (summer forecasting), and 1-October (fall forecasting). (2) *Seasonal drought outlook*. The seasonal outlook products offer the seasonal drought information in an earlier time manner (up to six-month preparation time). There are three seasonal outlook products in the proposed system: the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> seasonal outlook. The 1<sup>st</sup> seasonal outlook products provide drought information for the next season given the drought status in last season. When more

information becomes available, the system updates and generates the 2<sup>nd</sup> and 3<sup>rd</sup> seasonal drought outlook products with four and five months of information. Due to the persistence of soil moisture memory, the seasonal drought forecasting is expected to provide better results than seasonal drought outlook. (3) *Inter-annual drought forecasting*. The inter-annual drought forecasting products estimate the drought information for the next half a year. It is performed on 1-January, 1-April, 1-July, and 1-October. The inter-annual drought forecasting can provide long-term drought information for decision-makers.

-----  
Please place Figure 2 here  
-----

### **3 Case Study**

In general, droughts can be classified as meteorological drought (precipitation deficit), agricultural drought (soil moisture deficit), and hydrological drought (streamflow/groundwater deficit) (Van Loon, 2015). In this study, we examine the impacts of quantifying the root-zone soil moisture initial condition uncertainty with the assimilation of satellite surface soil moisture for improving agricultural drought forecasting. Similar to what has been explained here, other remotely sensed or in-situ data, e.g. total water storage and precipitation, can also be assimilated in the same way to improve the forecasting of hydrological and meteorological droughts.

### 3.1 Study Area

The Columbia River Basin (CRB), located in Pacific Northwest (PNW), covers about 674,500 km<sup>2</sup> of U.S. (~85%) and Canada (~15%) (Beechie and Imaki, 2014). In the U.S., the CRB spans seven states, including Washington, Oregon, Idaho, Montana, Utah, Wyoming, and Nevada (Figure 3). The CRB encompasses a wide range of physiographic provinces and ecoregions ranging from semiarid in central plateaus to wet forests in the Cascade Mountains (Omernik and Bailey, 1997). The mean annual precipitation ranges from about 200 mm in central plateaus to about 3,550 mm in Cascade Mountains (Daly et al., 2002). The Columbia River is the 4<sup>th</sup> largest river in North America, originating in the Rocky Mountain and flowing to the Pacific Ocean, with a mean annual discharge of 247 million cubic meters (Cosens and Williams, 2012). More than 450 dams have been built on this river system to provide hydroelectricity, flood control, irrigation, and stream regulation. The Columbia River is dominated by snow hydrology: snow accumulating in winter and melting in spring. It generally shows a characteristic of low flow in winter and peak flow in spring (Hamlet and Lettenmaier, 1999). Recently, the CRB droughts have received increasing attention due to the low snowpack and rising temperature. For instance, the Washington Department of Agriculture (2015) calculated that the economic loss of the 2015 state drought was more than \$335 million.

-----

Please place Figure 3 here

-----

### 3.2 Data Sources

The meteorological forcing data, precipitation, maximum and minimum temperature (January 1, 1979 to December 31, 2015) were acquired from the Phase 2 of the North American Land Data Assimilation Systems (NLDAS-2) (Xia et al., 2014). The majority of NLDAS-2 atmospheric forcing data are derived from the North American Regional Reanalysis (NARR) which features a 32 km spatial resolution and a three-hour temporal resolution. The NLDAS software is used to interpolate the coarse resolution NARR data to the finer scale  $1/8^{\text{th}}$  degree NLDAS grid and to the one-hour NLDAS temporal resolution (Xia et al., 2012). In this study, the hourly NLDAS-2 primary forcing data were first aggregated into daily time step and then upscaled/downscaled for each HRU in CRB.

The blended microwave soil moisture climate change initiative (CCI) products v02.2 released on February 2016 (Liu et al., 2012) and the Advanced Microwave Scanning Radiometer2 (AMSR2) soil moisture products (Imaoka et al., 2010) were used in this study. The CCI soil moisture were merged from four passive and two active microwave products (Dorigo et al., 2015), including the Scanning Multichannel Microwave Radiometer (SMMR), Special Sensor Microwave Imager (SSM/I), Tropical Rainfall Measure Mission (TRMM) Microwave Imager (TMI), Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E), Advanced Microwave Instrument (AMI), and Advanced Scatterometer (ASCAT). The blended CCI data cover the period of 1978-2014. The AMSR2 is onboard the Global Change Observation Mission1-Water (GCOM-W1) satellite which was launched in May 2012 by the Japan Aerospace Exploration Agency (JAXA). The AMSR2 soil moisture products, generated using the Land Parameter Retrieval Model (LPRM) developed by Vrije Universiteit (VU) Amsterdam and NASA Goddard Space Flight Center (GSFC) (Owe et al., 2008), were

employed in this study. These two products are selected due to the data availability. The latest Soil Moisture Active Passive (SMAP) (Entekhabi et al., 2010) L-band soil moisture products issued from March 31, 2015 will be used for future operational drought forecasting. All the satellite soil moisture products were upscaled/downscaled for each HRU in CRB. The inverse distance weighting (IDW) spatial interpolation technique is used in this study to downscale/upscale the satellite soil moisture into HRU scale.

Calibration of the PRMS was performed on a daily timescale utilizing a combination of unregulated U.S. Geological Survey (USGS) streamflow data, as well as No Regulation No Irrigation (NRNI) streamflow data provided by Bonneville Power Administration (BPA) [<https://www.bpa.gov/power/streamflow/>]. The BPA-NRNI data cover daily streamflow from the period July 1, 1928 to September 30, 2008, and were generated in a joint effort between the U.S. Army Corps of Engineers (USACE) and the Bureau of Reclamation (BOR). The NRNI datasets emulate daily discharge gauges, typically at manmade control structures, where estimated and measured inputs and outputs can be summed to produce daily streamflow data.

### **3.3 Model Calibration**

Due to the CRB's significant extent into British Columbia, the HRU delineation was completed in two ways. The HRUs for the U.S. portion were provided by the PRMS Geospatial Fabric (Viger, 2014). For the area of the CRB inside British Columbia, the ESRI ArcMap 10.3.1 was used along with a digital elevation model to produce stream segment lines as well as watershed delineations. Due to a lack of data to calibrate the Canadian portion of the CRB, the HRUs are rather large, relative to the U.S. portion. The two spatial shapefiles (U.S. and Canadian) were then stitched together, ensuring no overlapping HRU areas and continuity between stream



segments between the U.S. and Canada border. As a result, a total of 7,739 HRUs and 4,019 stream segments were delineated for the CRB.

In this study, 146 NRNI data sets were used as the primary control points for calibration along with over 300 selected USGS streamflow gauges. The NRNI data from January 1, 1979 to December 31, 2000 were used for calibration and January 1, 2001 to September 30, 2008 for validation. Calibrations for USGS streamflow gauges varied in length due to lapsing streamflow gauge operation, however only data within the range of January 1, 1979 to December 31, 2010 were used for calibration/validation. Monthly averaged normal incident solar radiation atlas data [[http://www.nrel.gov/gis/data\\_solar.html](http://www.nrel.gov/gis/data_solar.html)] and monthly averaged evaporation atlas (Farnsworth et al., 1982) were used for calibration of solar radiation (SR) and potential evapotranspiration (PET) parameters in the U.S. portion of the model. The SR and PET parameters for the Canadian portion of the CRB were extrapolated from the U.S. portion.

A combination of two programs created by the USGS was used in calibration, LUCA and LUMEN (Hay and Umemoto, 2006). LUMEN assists in model structure during calibration so that the model can more easily be calibrated using a top-down approach. LUCA uses the Shuffled Complex Evolution (SCE) global search algorithm to calibrate the 31 model parameters (Duan et al., 1994). Figure 3 presents the Kling-Gupta efficiency (KGE) (Gupta et al., 2009) values for the 146 NRNI control points from January 1, 1984 to September 30, 2008. The majority of the gauges show the KGE values greater than 0.7, which indicates the good performance of the calibrated model.

### 3.4 Synthetic Study

To objectively assess the potential benefit of quantification of initial condition uncertainty in drought forecasting, a synthetic study is first conducted. Following Moradkhani (2008), the synthetic study includes the following steps: (1) a “truth” run of PRMS with the pre-calibrated model parameters; (2) synthetic satellite soil moisture observations, which are generated from the truth run by incorporating the observation errors; (3) ensemble OL run with perturbed forcing data without DA; (4) the DA integration that assimilates the synthetic satellite observations. The probabilistic initial conditions from OL and DA are then used in the copula model to generate drought forecasting products. Finally, the OL and DA forecasting results are compared against the true simulations to evaluate the impact of initial condition uncertainty through DA. The copula probabilistic drought forecasting model is developed based on the truth run of root-zone soil moisture. It is noted that in the following discussions, the “DA” experiment indicates the results with the assimilation of satellite surface soil moisture, while the “OL” indicates the results without assimilation of satellite observations, hence the uncertainty in initial condition is not accounted for. Both DA and OL results are achieved through hydrologic modeling and copula-based statistical modeling.

The DA was performed by assimilation of the synthetic satellite surface soil moisture for the period of October 1, 2014 to September 30, 2015. We focused on hindcasting the drought events in 2015, since the CRB received historically low snowpack in this year and drought emergencies had been declared in Oregon (OR) and Washington (WA) states in spring 2015. Considering the satellite data availability, we employed the CCI soil moisture products for DA in 2014 and AMSR2 retrievals for 2015. Contrary to the small ensemble size (12~20) used in the majority of previous satellite soil moisture DA studies (De Lannoy et al., 2012; Kumar et al.,

2014b, 2009; Pan and Wood, 2010; Reichle et al., 2010), a large ensemble size of 200 was used in this study to fully quantify the soil moisture posteriors. For the purpose of better visualization, all PRMS simulation results were downscaled/upscaled into NLDAS 1/8<sup>th</sup> degree grid cells. Results are presented only for the U.S. portion of CRB in this study.

In the DA implementation, the precipitation was perturbed with a lognormal distribution with a coefficient of variation of 0.25, and the minimum and maximum temperature were assumed to follow normal distributions with a coefficient of variation of 0.25. These values are suggested to account for errors in meteorological measurements due to spatial heterogeneity and sensor errors (DeChant and Moradkhani, 2015; Yan et al., 2015). The white noise (standard deviation) for the CCI and AMSR2 soil moisture were 0.04 and 0.08 m<sup>3</sup>/m<sup>3</sup>, respectively, according to Kumar et al. (2014b). Prior to DA, we scaled the CCI and AMSR2 standard deviations by the ratio of the soil moisture time series standard deviation of the PRMS model as suggested by Reichle et al. (2007) and Liu et al. (2011). After scaling, the synthetic satellite soil moisture observations were generated by perturbing the synthetic truths with a normal distribution with the scaled standard deviation.

*The effects of DA on soil moisture estimation.* The improvement or degradation of DA on soil moisture predictions are assessed using the normalized information contribution (NIC) (Kumar et al., 2014b). The NIC for root-mean-square-error (RMSE) is defined as follows:

$$NIC = \frac{RMSE_{OL} - RMSE_{DA}}{RMSE_{OL}} \quad (17)$$

where  $RMSE_{OL}$  indicates the RMSE values between OL and synthetic truth,  $RMSE_{DA}$  indicates the RMSE values between DA and synthetic truth. If  $NIC > 0$ , the DA improves the OL skill; if  $NIC = 0$ , the DA does not add any skill; if  $NIC < 0$ , the DA degrades the OL skill; and if  $NIC = 1$ , the DA achieves the maximum skill. Figure 4 presents the NIC values in the surface and root-

zone soil moisture and their spatial distributions across the CRB. It is noted that the NIC values for OL and DA are generated using the ensemble mean estimates. From Figure 4, the majority of the grid cells show the positive NIC values, which indicates the added-value of the DA. Generally, the improvements in the surface soil moisture field are consistent with the improvements in the root-zone soil moisture field. For surface soil moisture, the daily basin-averaged RMSE ( $\text{m}^3/\text{m}^3$ ) for the OL was 0.0213, and it decreased to 0.0114 with DA. Similarly, the daily basin-averaged root-zone soil moisture RMSE value decreased from 0.0194 in the OL to 0.0116 in the DA. The improvements in surface field are higher than root-zone field, which is consistent with the previous soil moisture synthetic studies (Kumar et al., 2014a, 2012, 2009).

-----  
Please place Figure 4 here  
-----

*The effects of DA on drought forecasting.* Prior to investigating the drought forecasting results, it is necessary to compare the initial conditions generated by the OL and DA. This comparison can be seen in Figure 5. In this figure, the initial conditions for seasonal drought forecasting beginning on January 1 (winter forecasting), April 1 (spring forecasting), July 1 (summer forecasting), and October 1, 2015 (fall forecasting) are presented. Each sub-plot contains the basin-averaged daily root-zone soil moisture for the synthetic truth, shown as a single value; the OL and the DA, shown as a distribution of values, which represent the probability distribution of initial root-zone soil moisture values. From Figure 5, it is observed that the OL and DA display very different behavior for each season. For all four seasons, the DA

root-zone soil moisture reduce the uncertainty of OL estimations and the mean root-zone soil moisture is closer to the synthetic truth for the DA.

-----  
Please place Figure 5 here  
-----

In this study, drought is characterized with the root-zone soil moisture percentile (Mao et al., 2015; Shukla et al., 2011; Wang et al., 2011) and drought intensity is classified based on the National Drought Mitigation Center (NDMC) United States Drought Monitoring (USDM) (Svoboda et al., 2002). Five categories are defined: D0 (abnormally dry, percentile  $\leq 30\%$ ), D1 (moderate drought, percentile  $\leq 20\%$ ), D2 (severe drought, percentile  $\leq 10\%$ ), D3 (extreme drought, percentile  $\leq 5\%$ ), and D4 (exceptional drought, percentile  $\leq 2\%$ ). Leaving out the first five years as model spin-up period, the copula drought forecasting model was developed based on the true simulations from January 1, 1984 to December 31, 2014. The probabilistic forecast of drought status in the following season given the drought condition in the current season was examined using root-zone soil moisture for each grid cell. The dependencies between aggregated root-zone soil moisture were modeled by a Gaussian copula while their marginal distributions are modeled with lognormal distributions as suggested by Madadgar and Moradkhani (2014). The copula forecasted drought conditions using the OL and DA initial conditions were compared against the corresponding synthetic truth. It is noted that the OL and DA initial conditions were sampled as the posterior mean values.

Figure 6 presents the spatial distributions of seasonal drought forecasting probabilities of OL and DA for winter/spring/summer/fall 2015. The probabilistic drought conditions in each season were forecasted using the root-zone soil moisture in the previous season. The absolute drought extent bias (%) between the synthetic truth and MLE forecasted droughts are shown in the top panel of Figure 7. Generally, compared with the synthetic truth, both OL and DA seasonal forecasting products showed high probabilities (>30%) for the major drought locations in the following seasons. These results indicate the efficiency of the copula model in seasonal drought forecasting. Based on the drought extent bias, the DA estimates showed systematic improvements over the OL estimates in the four seasons. For instance, the drought extent bias between the MLE-OL and synthetic truth was 29.87% for winter 2015 forecasting, and it decreased to 20.71% with MLE-DA. Similarly, drought extent bias decreased from 4.73% with MLE-OL to 0.95% with MLE-DA for 2015 summer.

-----  
Please place Figures 6-7 here  
-----

Figure 8 shows the spatial distributions of three seasonal drought outlook and inter-annual drought forecasting probabilities of OL and DA for summer 2015 and April-September (spring and summer) 2015, respectively. The absolute drought extent bias (%) between the synthetic truth and MLE forecasted droughts are shown in the bottom panel of Figure 7. For seasonal outlook, the drought extent bias decrease when more information becomes available. For OL, the drought extent bias for the 1<sup>st</sup> outlook was 33.13% for summer 2015, and it decreased to 29.67%

and 24.39% for the 2<sup>nd</sup> and 3<sup>rd</sup> outlook, respectively. For DA, the drought extent bias for the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> outlooks were 28.89%, 23.63%, and 17.13%, respectively. For summer 2015, the drought extent bias for seasonal drought forecasting is the lowest, which is due to the persistence of soil moisture memory. The seasonal drought forecasting provides better results than the seasonal drought outlook. The inter-annual drought forecasting for April-September 2015 shows the highest drought extent bias. This result indicates the large uncertainty associated with the long-term drought forecasting.

-----  
Please place Figure 8 here  
-----

### **3.5 Real Case Study**

When assimilating the real satellite soil moisture data, the systematic biases between the satellite-based and model-based soil moisture cannot be avoided (Reichle and Koster, 2004; Su et al., 2014; Yan et al., 2015; Yilmaz and Crow, 2013). Proper treatment of these systematic biases is important, as the DA algorithm is designed to work with errors that are strictly random (Dee and Da Silva, 1998; Doucet and Johansen, 2011; Evensen, 1994). The most common approach, the cumulative distribution function (CDF)-matching (Reichle and Koster, 2004), was implemented here to rescale the satellite observations to the model's climatology. The CDF-matching approach can correct all the moments of the distribution regardless of its shape. Leaving out the first five years as model spin-up period, the CCI soil moisture products were rescaled from 1984-2014. The AMSR2 soil moisture products were rescaled from 2012-2015.

For real data study, the perturbation errors were the same as the synthetic study except for the model error. The model error is normally distributed with a coefficient of variation of 0.15 (Yan et al., 2015). Since the satellite soil moisture data quality control plays an important role in the final DA performance (Champagne et al., 2011; Yin et al., 2014), before the rescaling procedure, both the CCI and AMSR2 soil moisture dataset are screened by the data provider for larger water bodies (percent of land in pixel < 95%), frozen soils (surface temperature < 273 K), and dense vegetation (vegetation optical depth > 0.8).

Since no “true” drought data exist for real case study, the state drought declarations are used as the references to assess the drought forecasting skill (Shukla et al., 2011). Two case studies are presented here to indicate the added-value of DA for improving drought forecasting skill. (1) In spring 2013, drought declarations were issued for nine counties in the southern Idaho (ID). Three months later, a total of 19 counties in ID issued drought emergence. (2) In winter 2015, the PNW received historically low snowpack conditions. In June 2015, WA Governor declared the statewide drought and OR Governor declared drought emergencies for 19 out of 36 Oregon counties (about 80% of the state’s landmass). As a result, the DA was performed by assimilation of the real satellite surface soil moisture for six-month period until the forecast initialization date.

The drought conditions of the PNW in summer 2013 and spring 2015 can be predicted based on the drought conditions in spring 2013 and winter 2015, with the copula-based seasonal drought forecasting system. The DA can further improve the drought forecasting skill as demonstrated in the synthetic study. Similar to the synthetic study, prior to investigating the seasonal drought forecasting results, it is necessary to compare the initial conditions of OL and DA. In Figure 9, the initial conditions for seasonal drought forecasting beginning on July 1, 2013



(summer forecasting) and April 1, 2015 (spring forecasting) are presented. Each sub-plot contains the basin-averaged daily root-zone soil moisture for the OL, shown as a single value; and the DA, shown as a distribution of values, which represent the probability distribution of initial root-zone soil moisture values. For both seasons, the DA root-zone soil moisture show lower values, which is more consistent with the real drought situations.

-----  
Please place Figure 9 here  
-----

Figure 10 presents the spatial distributions of seasonal drought forecasting probabilities and MLE forecasted droughts for summer 2013 and spring 2015, respectively. The MLE forecasted drought extent (%) over the U.S. portion of the CRB between the OL and DA is shown in Figure 11. Given the initial conditions in spring 2013, the basin-averaged forecasted drought probabilities for summer 2013 were 51.80% and 54.64% for OL and DA, respectively. The forecasting drought probabilities for spring 2015 were 32.86% and 49.58% for OL and DA, respectively. For both cases, the DA suggests a higher probability of drought in summer 2013 and spring 2015, which is more consistent with the state declaration. The MLE forecasted drought extents increased from 64.67% and 25.85% in the OL to 70.57% and 52.83% in the DA for summer 2013 and spring 2015, respectively. Especially for the spring 2015, the OL forecasts underestimated the severe drought conditions for WA and OR. In terms of both forecasted probabilities and MLE forecasted drought extents, the DA is much more consistent with the state declaration. In summary, compared with the OL, the DA improves the drought forecasting skill

for both summer 2013 and spring 2015. These results demonstrate the added-value of DA to facilitate the state drought preparation and declaration, at least three months before the official state drought declaration.

-----  
Please place Figures 10-11 here  
-----

It is also noted that with increasing ensemble size, the computational cost will increase. Therefore, the DA in this study is performed on a high-performance computing infrastructure. All the DA simulations are performed on the Linux Hydra Cluster located at Portland State University (PSU) with 24 nodes, 384 processors. The ensemble size of 200 was selected based on the DA performance accuracy and computational demand. Since we use the PMCMC approach in this study, it requires larger ensemble size than the EnKF. However, similar to the approach developed for the EnKF (Yin et al., 2015), the optimal ensemble size in the PMCMC can be estimated.

#### **4 Conclusions**

In this study, we proposed a probabilistic drought forecasting system, with a combination of dynamical and statistical components. The dynamical hydrologic modeling is coupled with the copula-based statistical forecasting. Moreover, the ensemble data assimilation technique is used to improve state initialization in the copula-based probabilistic forecasting framework by

allowing for uncertainty in the initial condition. This probabilistic drought forecasting system was implemented in the Columbia River Basin, Pacific Northwest. We examine the impacts of assimilating remotely sensed surface soil moisture for quantifying the initial condition uncertainty and their subsequent contributions toward an improved forecasting of agricultural droughts. Results from both synthetic and real case studies suggest that the proposed drought forecasting system significantly improves the seasonal agricultural drought forecasting skills and can facilitate the state drought preparation and declaration. Similar to what has been explained here, other satellite or in-situ data, e.g. precipitation and total water storage, can also be assimilated in the same way to improve the forecasting of meteorological and hydrological droughts.

### **Acknowledgement**

Partial financial support for this project was provided by the National Oceanic and Atmospheric Administration (NOAA) Modeling, Analysis, Predictions, and Projections (MAPP) (Grant No. NA140AR4310234) and National Science Foundation (NSF) Cyber-Innovation for Sustainability Science and Engineering (CyberSEES) (Grant No. CCF-1539605).

### **References**

- Ahmadalipour, A., Moradkhani, H., Yan, H., Zarekarizi, M., 2017. Remote Sensing of Drought: Vegetation, Soil Moisture, and Data Assimilation, in: Lakshmi, V. (Ed.), Remote Sensing of Hydrological Extremes. Springer International Publishing Switzerland, pp. 121–149. doi:10.1007/978-3-319-43744-6\_7
- Beechie, T., Imaki, H., 2014. Predicting natural channel patterns based on landscape and geomorphic controls in the Columbia River basin, USA. *Water Resour. Res.* 50, 39–57.

doi:10.1002/2013WR013629

- Champagne, C., McNairn, H., Berg, A. a., 2011. Monitoring agricultural soil moisture extremes in Canada using passive microwave remote sensing. *Remote Sens. Environ.* 115, 2434–2444. doi:10.1016/j.rse.2011.04.030
- Chen, Y.D., Zhang, Q., Xiao, M., Singh, V.P., Zhang, S., 2015. Probabilistic forecasting of seasonal droughts in the Pearl River basin, China. *Stoch. Environ. Res. Risk Assess.* 1–10. doi:10.1007/s00477-015-1174-6
- Cosens, B.A., Williams, M.K., 2012. Resilience and water governance: Adaptive governance in the Columbia River Basin. *Ecol. Soc.* 17. doi:10.5751/ES-04986-170403
- Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L., Pasteris, P., 2002. A knowledge-based approach to the statistical mapping of climate. *Clim. Res.* 22, 99–113. doi:10.3354/cr022099
- Day, G.N., 1985. Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plan. Manag.* 111, 157–170. doi:10.1061/(ASCE)0733-9496(1985)111:2(157)
- De Lannoy, G.J.M., Reichle, R.H., Arsenault, K.R., Houser, P.R., Kumar, S., Verhoest, N.E.C., Pauwels, V.R.N., 2012. Multiscale assimilation of Advanced Microwave Scanning Radiometer–EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado. *Water Resour. Res.* 48, W01522. doi:10.1029/2011WR010588
- DeChant, C.M., Moradkhani, H., 2015. Analyzing the sensitivity of drought recovery forecasts to land surface initial conditions. *J. Hydrol.* 526, 89–100. doi:10.1016/j.jhydrol.2014.10.021
- DeChant, C.M., Moradkhani, H., 2014. Toward a reliable prediction of seasonal forecast uncertainty: Addressing model and initial condition uncertainty with ensemble data assimilation and sequential Bayesian combination. *J. Hydrol.* 519, 2967–2977. doi:10.1016/j.jhydrol.2014.05.045
- DeChant, C.M., Moradkhani, H., 2012. Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting. *Water Resour. Res.* 48, W04518. doi:10.1029/2011WR011011
- DeChant, C.M., Moradkhani, H., 2011. Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. *Hydrol. Earth Syst. Sci.* 15, 3399–3410. doi:10.5194/hess-15-3399-2011
- Dee, D.P., Da Silva, A.M., 1998. Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.* 124, 269–295. doi:10.1002/qj.49712454512
- Dong, J., Steele-Dunne, S.C., Judge, J., van de Giesen, N., 2015. A particle batch smoother for soil moisture estimation using soil temperature observations. *Adv. Water Resour.* 83, 111–122. doi:10.1016/j.advwatres.2015.05.017
- Dorigo, W.A., Gruber, A., De Jeu, R.A.M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R.M., Kidd, R., 2015. Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sens. Environ.* 162, 380–395. doi:10.1016/j.rse.2014.07.023
- Doucet, A., Johansen, A., 2011. A tutorial on particle filtering and smoothing: fifteen years later, in: *The Oxford Handbook of Nonlinear Filtering*. Oxford University Press, pp. 656–704.
- Duan, Q., Sorooshian, S., Gupta, V.K., 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. *J. Hydrol.* 158, 265–284. doi:10.1016/0022-1694(94)90057-4
- Entekhabi, D., Njoku, E.G., O’Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin,

- J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., Neill, P.E.O., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., 2010. The soil moisture active passive (SMAP) mission. *Proc. IEEE* 98, 704–716. doi:10.1109/JPROC.2010.2043918
- Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* 99, 10143–10162. doi:10.1029/94JC00572
- Farnsworth, R.K., Thompson, E.S., Peck, E.L., 1982. *Evaporation Atlas for the Contiguous 48 United States*. NOAA Technical Report NWS 33.
- Genest, C., Rémillard, B., 2008. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann. l'Institut Henri Poincaré, Probab. Stat.* 44, 1096–1127. doi:10.1214/07-AIHP148
- Gupta, H. V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* 377, 80–91. doi:10.1016/j.jhydrol.2009.08.003
- Hamlet, A.F., Lettenmaier, D.P., 1999. Effects of climate change on hydrology and water resources in the Columbia River basin. *J. Am. Water Resour. Assoc.* 35, 1597–1623. doi:10.1111/j.1752-1688.1999.tb04240.x
- Hay, L.E., Umemoto, M., 2006. *Multiple-objective stepwise calibration using Luca*. US Geological Survey.
- Hayes, M., Svoboda, M., Le Comte, D., Redmond, K.T., Pasteris, P., 2005. Drought monitoring: New tools for the 21st century, in: White, D.A. (Ed.), *Drought and Water Crises: Science, Technology, and Management Issues*. CRC Press, pp. 53–69.
- Howitt, R., Medellín-Azuara, J., MacEwan, D., Lund, J., Sumner, D., 2014. *Economic analysis of the 2014 drought for California agriculture* University of California, Davis.
- Imaoka, K., Kachi, M., Kasahara, M., Ito, N., Nakagawa, K., Oki, T., 2010. Instrument performance and calibration of AMSR-E and AMSR2. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* 38, 13–16.
- Joe, H., 1997. *Multivariate models and dependence concepts*, Analysis. doi:Export Date 21 May 2013
- Keyantash, J., Dracup, J.A., 2002. The quantification of drought: An evaluation of drought indices. *Bull. Am. Meteorol. Soc.* 83, 1167–1180. doi:10.1175/1520-0477(2002)083<1191:TQODAE>2.3.CO;2
- Koster, R.D., Mahanama, S.P.P., Yamada, T.J., Balsamo, G., Berg, A.A., Boissarie, M., Dirmeyer, P.A., Doblas-Reyes, F.J., Drewitt, G., Gordon, C.T., Guo, Z., Jeong, J.-H., Lawrence, D.M., Lee, W.-S., Li, Z., Luo, L., Malyshev, S., Merryfield, W.J., Seneviratne, S.I., Stanelle, T., van den Hurk, B.J.J.M., Vitart, F., Wood, E.F., 2010. Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophys. Res. Lett.* 37, L02402. doi:10.1029/2009GL041677
- Kumar, S. V., Harrison, K.W., Peters-Lidard, C.D., Santanello, J. a., Kirschbaum, D., 2014a. Assessing the impact of L-band observations on drought and flood risk estimation: A

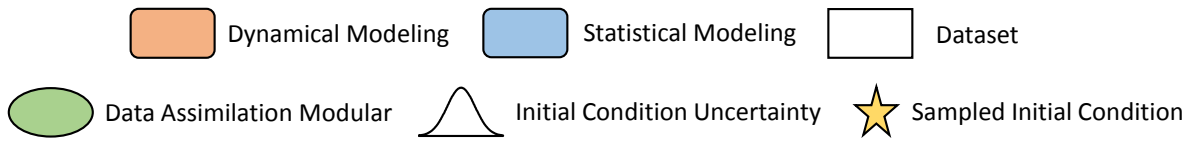
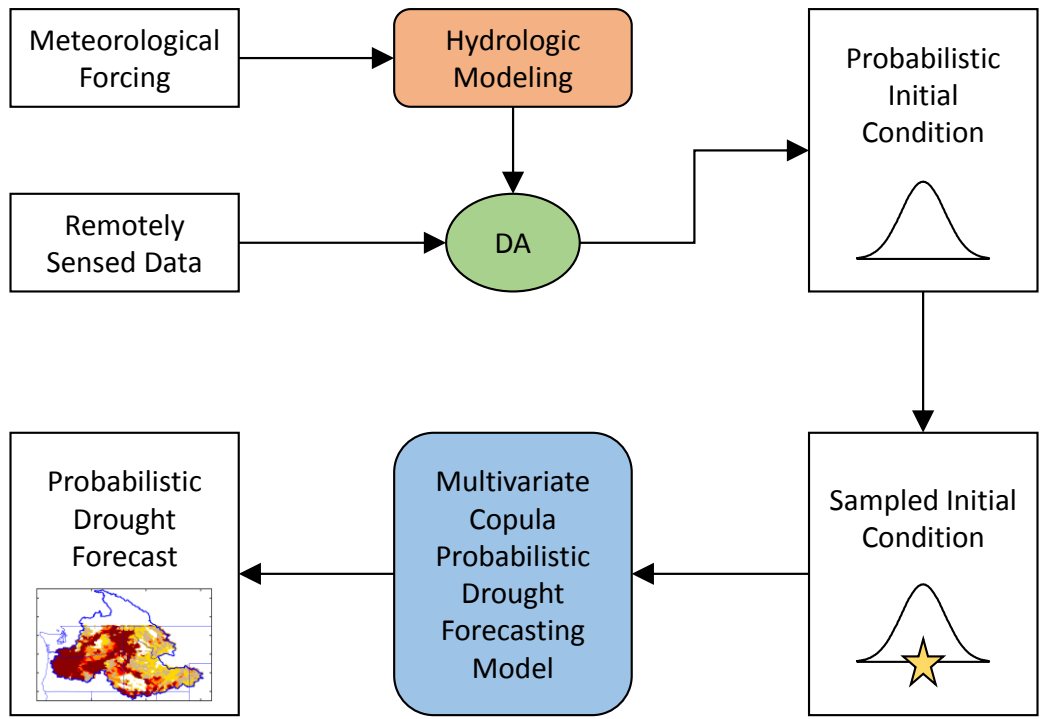
- decision-theoretic approach in an OSSE environment. *J. Hydrometeorol.* 15, 2140–2156. doi:10.1175/JHM-D-13-0204.1
- Kumar, S. V., Peters-Lidard, C.D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K.R., Xia, Y., Ek, M., Riggs, G., Livneh, B., Cosh, M., 2014b. Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation. *J. Hydrometeorol.* 15, 2446–2469. doi:10.1175/JHM-D-13-0132.1
- Kumar, S. V., Reichle, R.H., Harrison, K.W., Peters-Lidard, C.D., Yatheendradas, S., Santanello, J.A., 2012. A comparison of methods for a priori bias correction in soil moisture data assimilation. *Water Resour. Res.* 48. doi:10.1029/2010WR010261
- Kumar, S. V., Reichle, R.H., Koster, R.D., Crow, W.T., Peters-Lidard, C.D., 2009. Role of subsurface physics in the assimilation of surface soil moisture observations. *J. Hydrometeorol.* 10, 1534–1547. doi:10.1175/2009JHM1134.1
- Lavers, D., Luo, L., Wood, E.F., 2009. A multiple model assessment of seasonal climate forecast skill for applications. *Geophys. Res. Lett.* 36, L23711. doi:10.1029/2009GL041365
- Leavesley, G.H., Lichty, R.W., Thoutman, B.M., Saindon, L.G., 1983. Precipitation-runoff modeling system: User's manual. US Geological Survey Colorado, CO.
- Li, H., Luo, L., Wood, E.F., Schaake, J., 2009. The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *J. Geophys. Res.* 114, D04114. doi:10.1029/2008JD010969
- Liu, Q., Reichle, R.H., Bindlish, R., Cosh, M.H., Crow, W.T., de Jeu, R., De Lannoy, G.J.M., Huffman, G.J., Jackson, T.J., 2011. The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system. *J. Hydrometeorol.* 12, 750–765. doi:10.1175/JHM-D-10-05000.1
- Liu, Y.Y., Dorigo, W.A., Parinussa, R.M., De Jeu, R.A.M., Wagner, W., McCabe, M.F., Evans, J.P., Van Dijk, A.I.J.M., 2012. Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sens. Environ.* 123, 280–297. doi:10.1016/j.rse.2012.03.014
- Luo, L., Wood, E.F., 2007. Monitoring and predicting the 2007 U.S. drought. *Geophys. Res. Lett.* 34, L22702. doi:10.1029/2007GL031673
- Madadgar, S., Moradkhani, H., 2016. Copula function and drought, in: *Handbook of Drought and Water Scarcity, Vol. 1: Principles of Drought and Water Scarcity.* Francis and Taylor.
- Madadgar, S., Moradkhani, H., 2014. Spatio-temporal drought forecasting within Bayesian networks. *J. Hydrol.* 512, 134–146. doi:10.1016/j.jhydrol.2014.02.039
- Madadgar, S., Moradkhani, H., 2013. A Bayesian framework for probabilistic seasonal drought forecasting. *J. Hydrometeorol.* 14, 1685–1705. doi:10.1175/JHM-D-13-010.1
- Mao, Y., Nijssen, B., Lettenmaier, D.P., 2015. Is climate change implicated in the 2013–2014 California drought? A hydrologic perspective. *Geophys. Res. Lett.* 42, 2805–2813. doi:10.1002/2015GL063456
- Markstrom, S.L., Niswonger, R.G., Regan, R.S., Prudic, D.E., Barlow, P.M., 2008. GSFLOW-coupled ground-water and surface-water FLOW model based on the integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005). US Geological Survey.
- Markstrom, S.L., Regan, R.S., Hay, L.E., Viger, R.J., Webb, R.M., Payn, R.A., LaFontaine, J.H., 2015. PRMS-IV, the precipitation-runoff modeling system, version 4. US Geological Survey.
- Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier,

- D.P., Stouffer, R.J., 2008. Stationarity is dead: whither water management? *Science* (80-. ). 319, 573–574. doi:10.1126/science.1151915
- Mishra, A.K., Desai, V.R., 2005. Drought forecasting using stochastic models. *Stoch. Environ. Res. Risk Assess.* 19, 326–339. doi:10.1007/s00477-005-0238-4
- Mishra, A.K., Singh, V.P., 2011. Drought modeling – A review. *J. Hydrol.* 403, 157–175. doi:10.1016/j.jhydrol.2011.03.049
- Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. *J. Hydrol.* 391, 202–216. doi:10.1016/j.jhydrol.2010.07.012
- Moradkhani, H., 2008. Hydrologic remote sensing and land surface data assimilation. *Sensors* 8, 2986–3004. doi:10.3390/s8052986
- Moradkhani, H., Dechant, C.M., Sorooshian, S., 2012. Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. *Water Resour. Res.* 48, W12520. doi:10.1029/2012WR012144
- Moradkhani, H., Hsu, K.-L., Gupta, H., Sorooshian, S., 2005. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resour. Res.* 41, W05012. doi:10.1029/2004WR003604
- NCDC, 2012. Billion-dollar weather/climate events. [WWW Document]. NOAA.
- Omerik, J.M., Bailey, R.G., 1997. Distinguishing between watersheds and ecoregions. *J. Am. Water Resour. Assoc.* 33, 935–949. doi:10.1111/j.1752-1688.1997.tb04115.x
- Owe, M., de Jeu, R., Holmes, T., 2008. Multisensor historical climatology of satellite-derived global land surface moisture. *J. Geophys. Res. Earth Surf.* 113. doi:10.1029/2007JF000769
- Pan, M., Wood, E.F., 2010. Impact of accuracy, spatial availability, and revisit time of satellite-derived surface soil moisture in a multiscale ensemble data assimilation system. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 3, 49–56. doi:10.1109/JSTARS.2010.2040585
- Plaza, D.A., De Keyser, R., De Lannoy, G.J.M., Giustarini, L., Matgen, P., Pauwels, V.R.N., 2012. The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter. *Hydrol. Earth Syst. Sci.* 16, 375–390. doi:10.5194/hess-16-375-2012s
- Reichle, R.H., Koster, R.D., 2004. Bias reduction in short records of satellite soil moisture. *Geophys. Res. Lett.* 31, L19501. doi:10.1029/2004GL020938
- Reichle, R.H., Koster, R.D., Liu, P., Mahanama, S.P.P., Njoku, E.G., Owe, M., 2007. Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *J. Geophys. Res.* 112, D09108. doi:10.1029/2006JD008033
- Reichle, R.H., Kumar, S. V., Mahanama, S.P.P., Koster, R.D., Liu, Q., 2010. Assimilation of satellite-derived skin temperature observations into land surface models. *J. Hydrometeorol.* 11, 1103–1122. doi:10.1175/2010JHM1262.1
- Ross, T., Lott, N., 2003. A climatology of 1980-2003 extreme weather and climate events. National Oceanic and Atmospheric Administration.
- Sheffield, J., Wood, E.F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., Ogallo, L., 2014. A drought monitoring and forecasting system for sub-sahara african water resources and food security. *Bull. Am. Meteorol. Soc.* 95, 861–882. doi:10.1175/BAMS-D-12-00124.1
- Shukla, S., Lettenmaier, D.P., 2011. Seasonal hydrologic prediction in the United States: Understanding the role of initial hydrologic conditions and seasonal climate forecast skill.

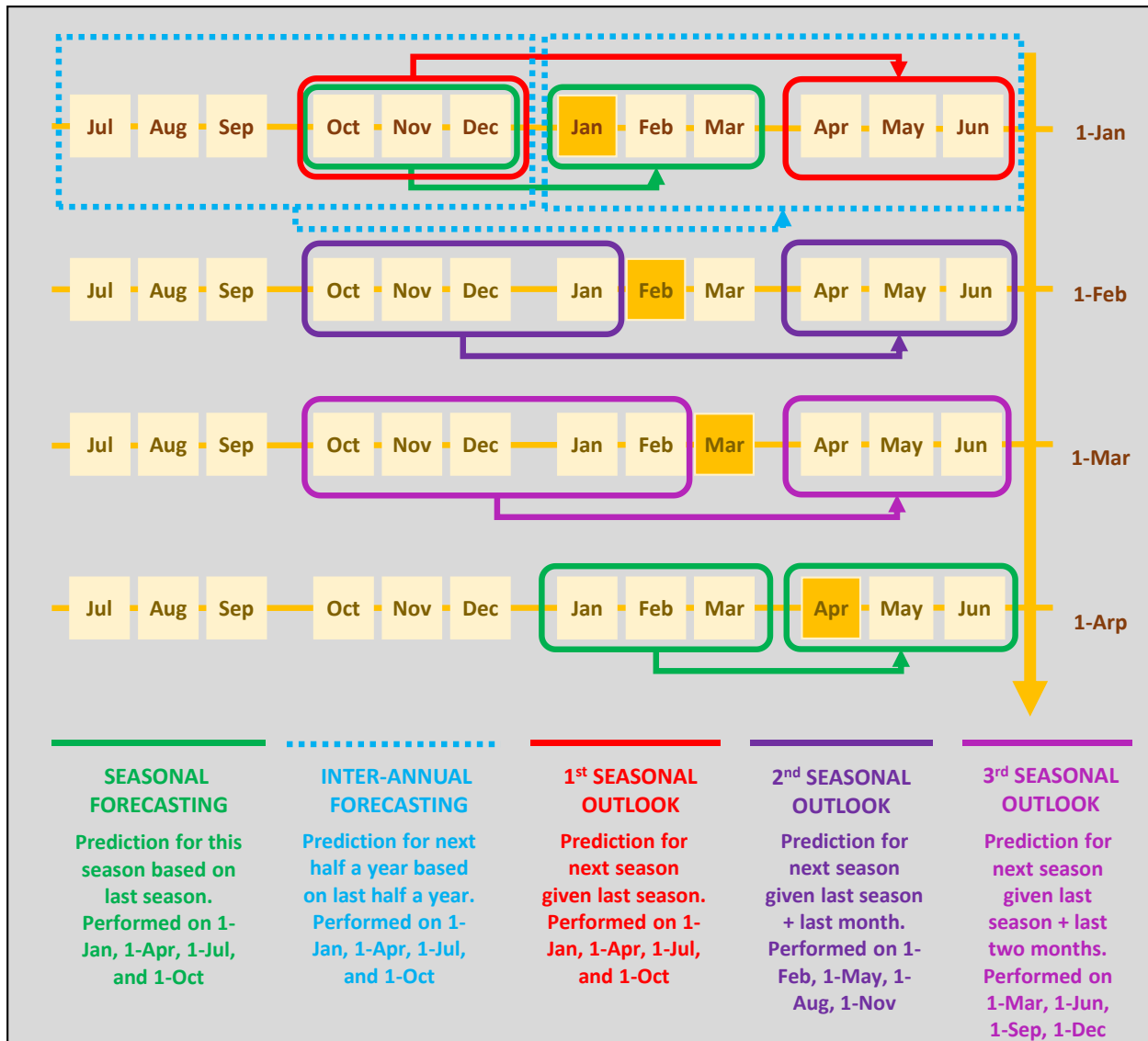
- Hydrol. Earth Syst. Sci. 15, 3529–3538. doi:10.5194/hess-15-3529-2011
- Shukla, S., Steinemann, A.C., Lettenmaier, D.P., 2011. Drought monitoring for Washington State: indicators and applications. *J. Hydrometeorol.* 12, 66–83. doi:10.1175/2010JHM1307.1
- Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. *Publ. l'Institut Stat. l'Université Paris 8*, 229–231.
- Stainforth, D.A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D.J., Kettleborough, J. a, Knight, S., Martin, A., Murphy, J.M., Piani, C., Sexton, D., Smith, L. a, Spicer, R. a, Thorpe, a J., Allen, M.R., 2005. Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* 433, 403–406. doi:10.1038/nature03301
- Steinemann, A.C., 2006. Using climate forecasts for drought management. *J. Appl. Meteorol. Climatol.* 45, 1353–1361. doi:10.1175/JAM2401.1
- Su, C.H., Ryu, D., Crow, W.T., Western, A.W., 2014. Beyond triple collocation: Applications to soil moisture monitoring. *J. Geophys. Res. Atmos.* 119, 6419–6439. doi:10.1002/2013JD021043
- Svoboda, M., LeCompte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., Stephens, S., 2002. The drought monitor. *Bull. Am. Meteorol. Soc.* 83, 1181–1190. doi:10.1175/1520-0477(2002)083<1181:TDM>2.3.CO;2
- Van Loon, A.F., 2015. Hydrological drought explained. *Wiley Interdiscip. Rev. Water* 2, 359–392. doi:10.1002/wat2.1085
- Viger, R.J., 2014. Preliminary spatial parameters for PRMS based on the Geospatial Fabric, NLCD2001 and SSURGO. *US Geol. Surv.* doi:10.5066/F7WM1BF7
- Wang, A., Lettenmaier, D.P., Sheffield, J., 2011. Soil moisture drought in China, 1950-2006. *J. Clim.* 24, 3257–3271. doi:10.1175/2011JCLI3733.1
- Washington Department of Agriculture, 2015. Interim report: 2015 drought and agriculture.
- WCRP, 2010. Drought predictability and prediction in a changing climate: assessing current capabilities, World Climate Research Programme. Barcelona, Spain.
- Wood, A.W., Hopson, T., Newman, A., Brekke, L., Arnold, J., Clark, M., 2016. Quantifying streamflow forecast skill elasticity to initial condition and climate prediction skill. *J. Hydrometeorol.* 17, 651–668. doi:10.1175/JHM-D-14-0213.1
- Wood, A.W., Lettenmaier, D.P., 2008. An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophys. Res. Lett.* 35, L14401. doi:10.1029/2008GL034648
- Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., 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. *J. Geophys. Res. Atmos.* 117, D03109. doi:10.1029/2011JD016048
- Xia, Y., Sheffield, J., Ek, M.B., Dong, J., Chaney, N., Wei, H., Meng, J., Wood, E.F., 2014. Evaluation of multi-model simulated soil moisture in NLDAS-2. *J. Hydrol.* 512, 107–125. doi:10.1016/j.jhydrol.2014.02.027
- Yan, H., 2016. From Drought Monitoring to Forecasting: A Combined Dynamical-Statistical Modeling Framework. Ph.D. Dissertation, Portland State University.
- Yan, H., DeChant, C.M., Moradkhani, H., 2015. Improving Soil Moisture Profile Prediction With the Particle Filter-Markov Chain Monte Carlo Method. *IEEE Trans. Geosci. Remote*



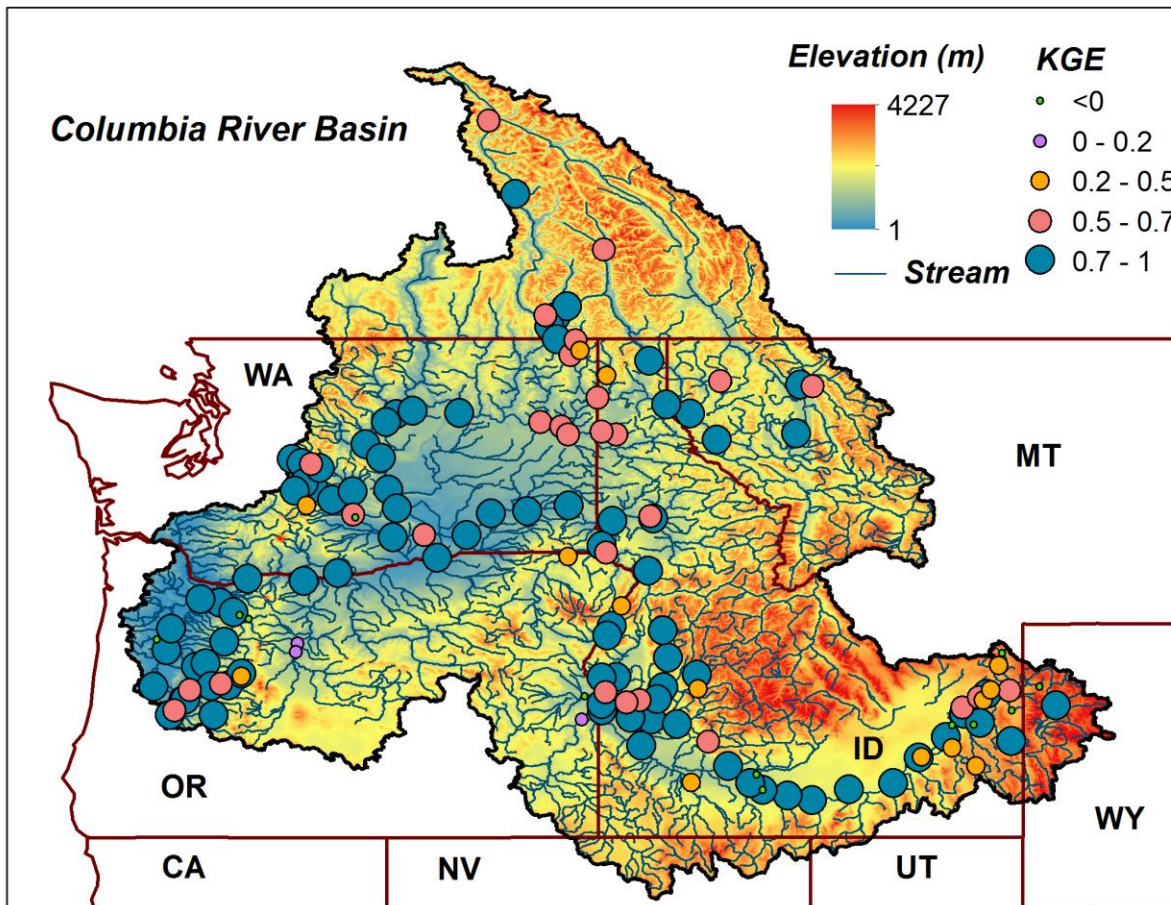
- Sens. 53, 6134–6147. doi:10.1109/TGRS.2015.2432067
- Yan, H., Edwards, F.G., 2013. Effects of Land Use Change on Hydrologic Response at a Watershed Scale, Arkansas. *J. Hydrol. Eng.* 18, 1779–1785. doi:10.1061/(ASCE)HE.1943-5584.0000743
- Yan, H., Moradkhani, H., 2016. Combined assimilation of streamflow and satellite soil moisture with the particle filter and geostatistical modeling. *Adv. Water Resour.* 94, 364–378. doi:10.1016/j.advwatres.2016.06.002
- Yilmaz, M.T., Crow, W.T., 2013. The optimality of potential rescaling approaches in land data assimilation. *J. Hydrometeorol.* 14, 650–660. doi:10.1175/JHM-D-12-052.1
- Yin, J., Zhan, X., Zheng, Y., Hain, C.R., Liu, J., Fang, L., 2015. Optimal ensemble size of ensemble Kalman filter in sequential soil moisture data assimilation. *Geophys. Res. Lett.* 42, 6710–6715. doi:10.1002/2015GL063366
- Yin, J., Zhan, X., Zheng, Y., Liu, J., Hain, C.R., Fang, L., 2014. Impact of quality control of satellite soil moisture data on their assimilation into land surface model. *Geophys. Res. Lett.* 41, 7159–7166. doi:10.1002/2014GL060659
- Yossef, N.C., Winsemius, H., Weerts, A., Van Beek, R., Bierkens, M.F.P., 2013. Skill of a global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resour. Res.* 49, 4687–4699. doi:10.1002/wrcr.20350
- Yuan, X., 2016. An experimental seasonal hydrological forecasting system over the Yellow River basin - Part 2: The added value from climate forecast models. *Hydrol. Earth Syst. Sci.* 20, 2453–2466. doi:10.5194/hess-20-2453-2016
- Yuan, X., Ma, F., Wang, L., Zheng, Z., Ma, Z., Ye, A., Peng, S., 2016. 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. doi:10.5194/hess-20-2437-2016
- Yuan, X., Wood, E.F., Luo, L., Pan, M., 2011. A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophys. Res. Lett.* 38, L13402. doi:10.1029/2011GL047792
- Yuan, X., Wood, E.F., Roundy, J.K., Pan, M., 2013. CFSv2-Based seasonal hydroclimatic forecasts over the conterminous United States. *J. Clim.* 26, 4828–4847. doi:10.1175/JCLI-D-12-00683.1



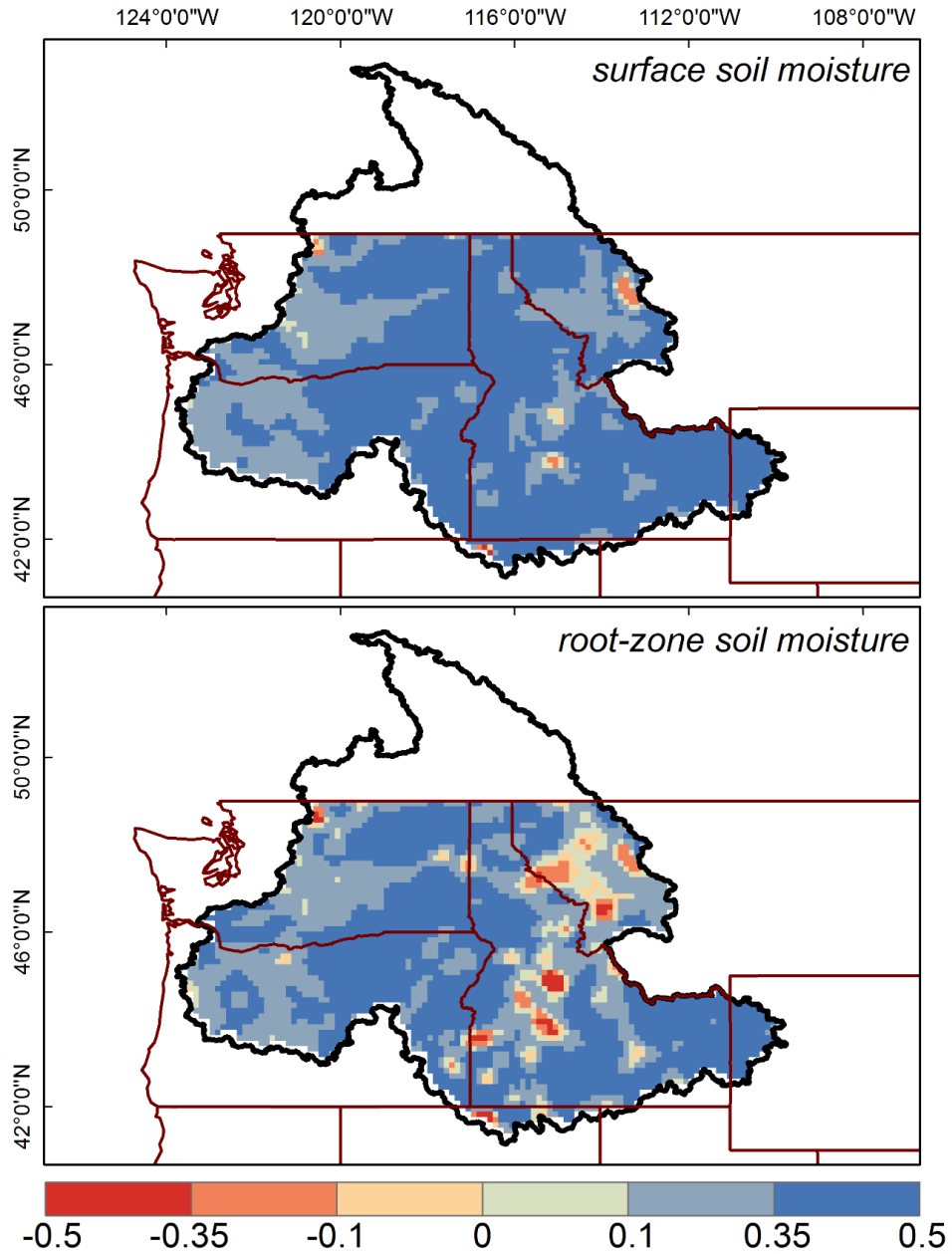
**Figure 1.** The framework of the proposed dynamical-statistical probabilistic drought forecasting system. For each forecasting initialization date, the data assimilation (DA) technique is used to account for the initial condition uncertainty, in terms of a probability density function (PDF). An updated initial condition is then sampled from the PDF and input into the multivariate copula drought forecasting model to generate the probabilistic drought forecasts.



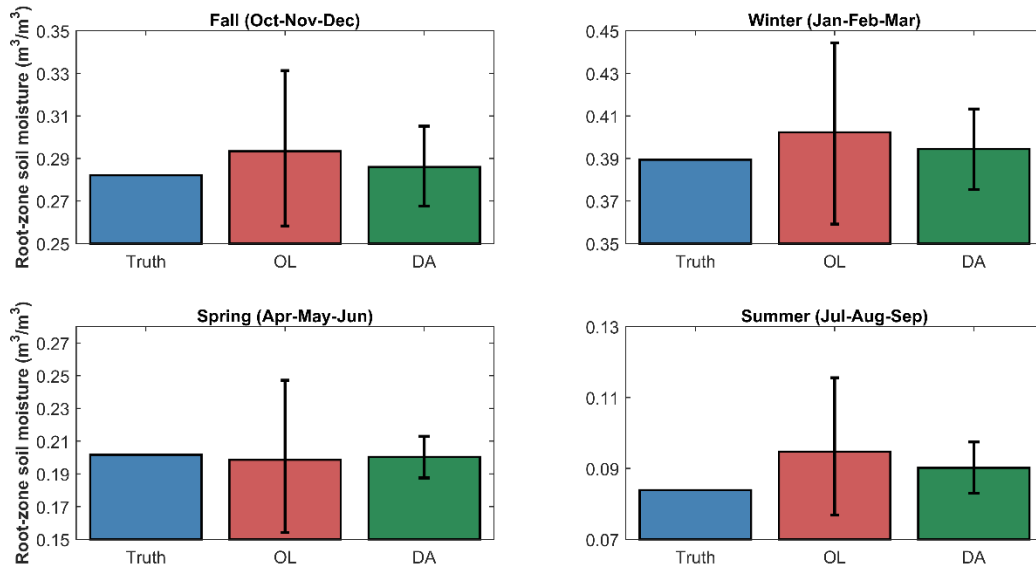
**Figure 2.** The seasonal and inter-annual drought forecasting framework.



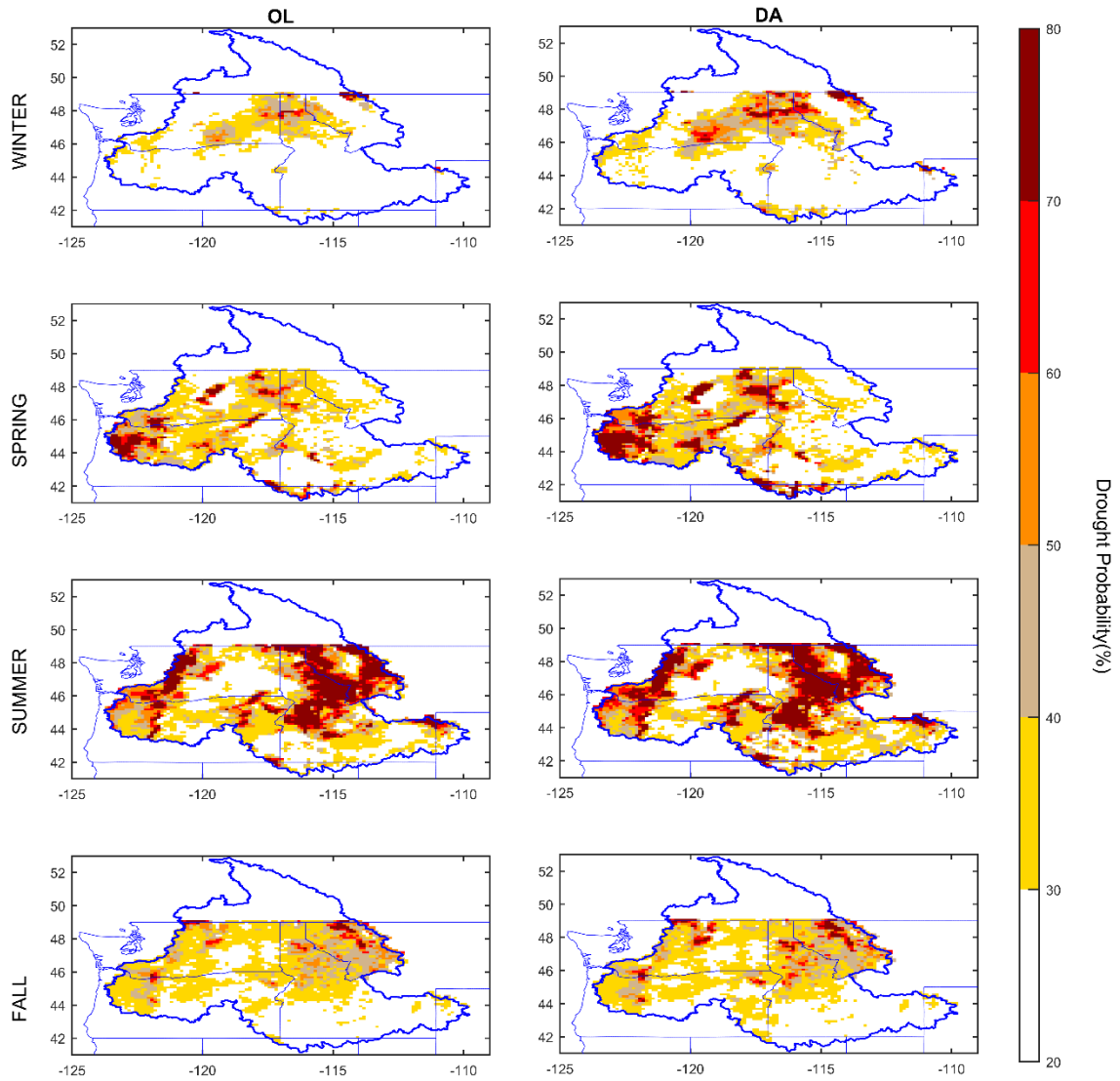
**Figure 3.** The location of the Columbia River Basin and the Kling-Gupta efficiency (KGE) values for the 146 No Regulation No Irrigation (NRNI) streamflow gauges.



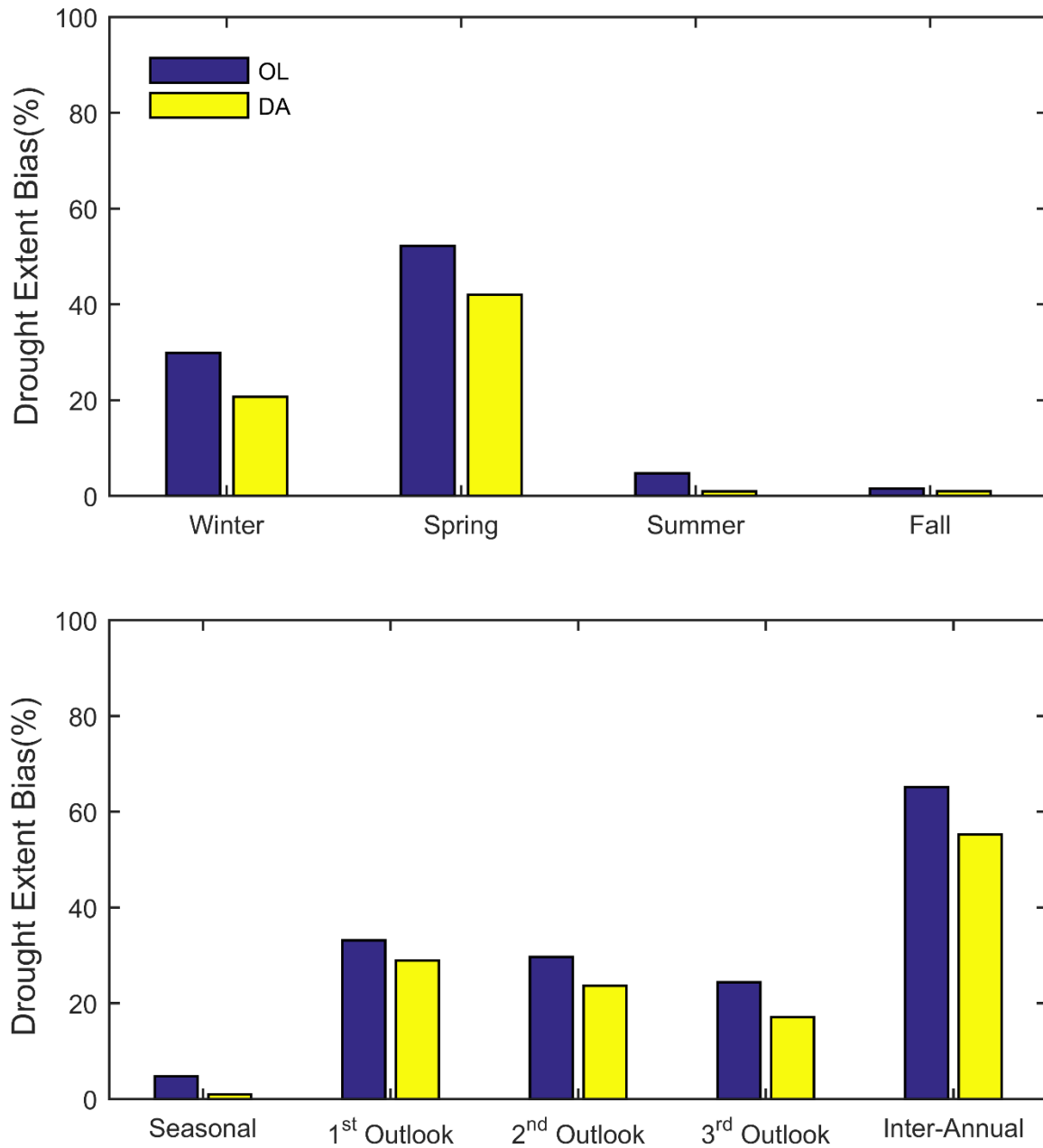
**Figure 4.** The normalized information contribution (NIC) value. The positive value indicates that the DA improves soil moisture prediction against OL; negative value indicates the degradation over the OL.



**Figure 5.** Comparison of the basin-averaged daily root-zone soil moisture ( $\text{m}^3/\text{m}^3$ ) by the open-loop (OL) and data assimilation (DA) for fall 2014 and winter/spring/summer 2015 across the CRB. The error bars show the 95% prediction intervals.

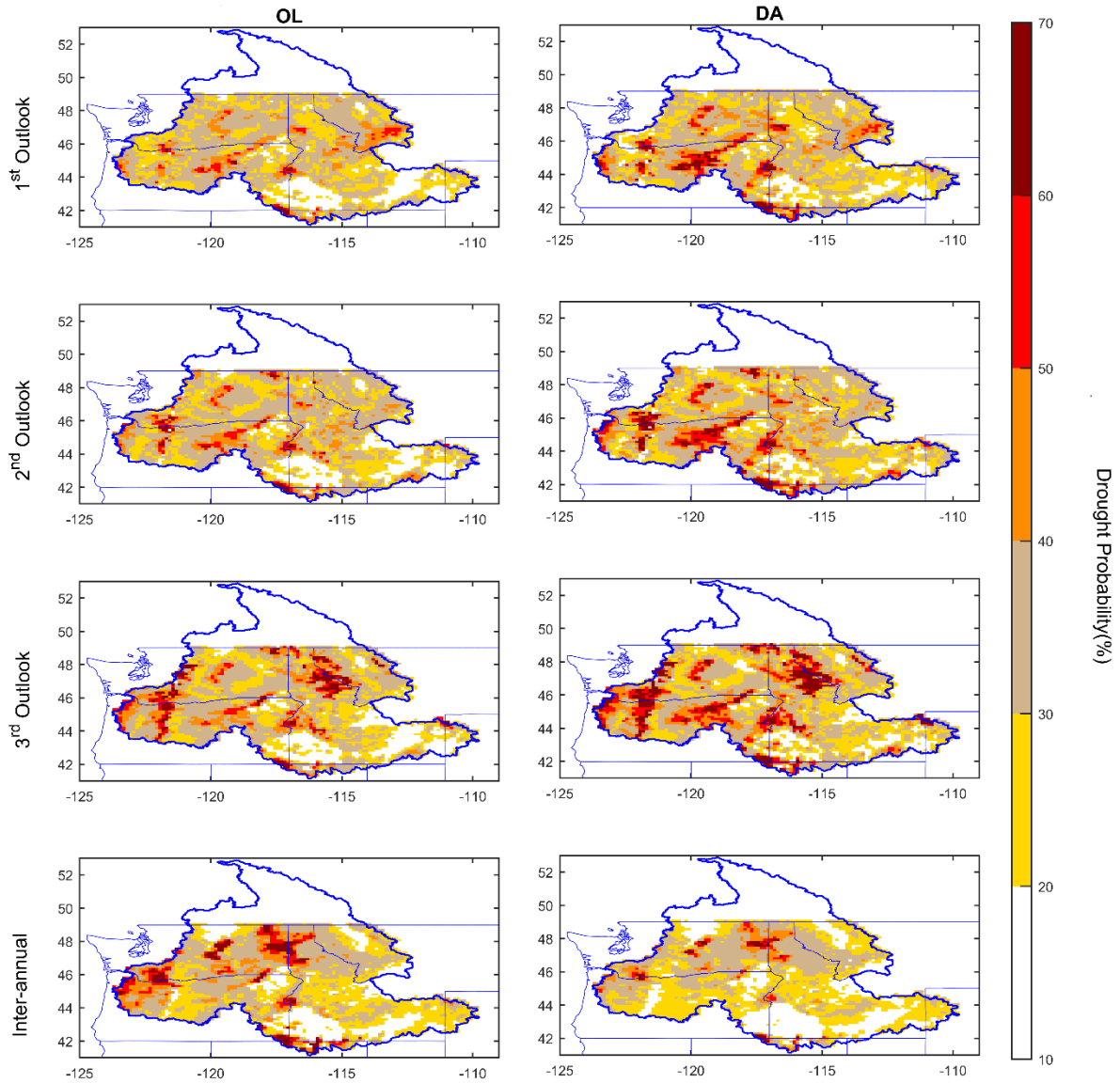


**Figure 6.** Seasonal probabilistic drought forecasting for both OL and DA for different seasons in 2015 given the drought status in each of the previous seasons.

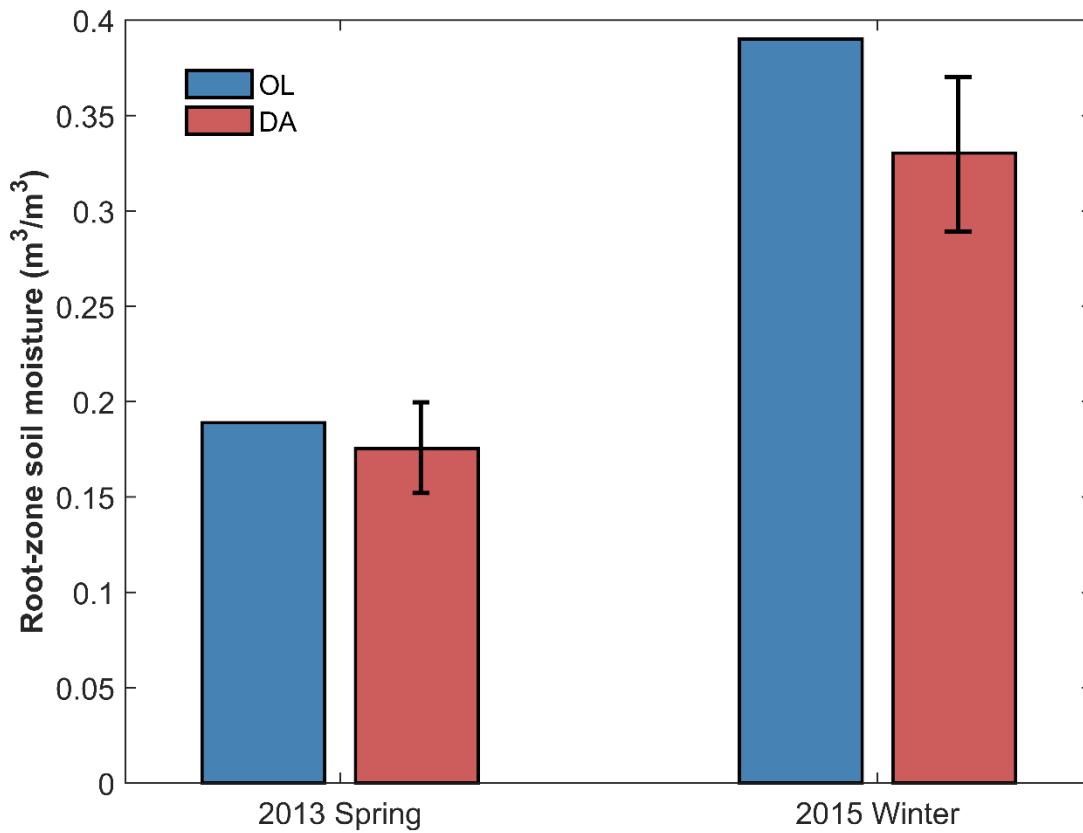


**Figure 7.** The absolute drought extent bias between the OL/DA and synthetic truth for the seasonal forecasting for winter/spring/summer/fall 2015 (top-panel), seasonal outlook for summer 2015 and inter-annual forecasting for April-September 2015 (bottom-panel).

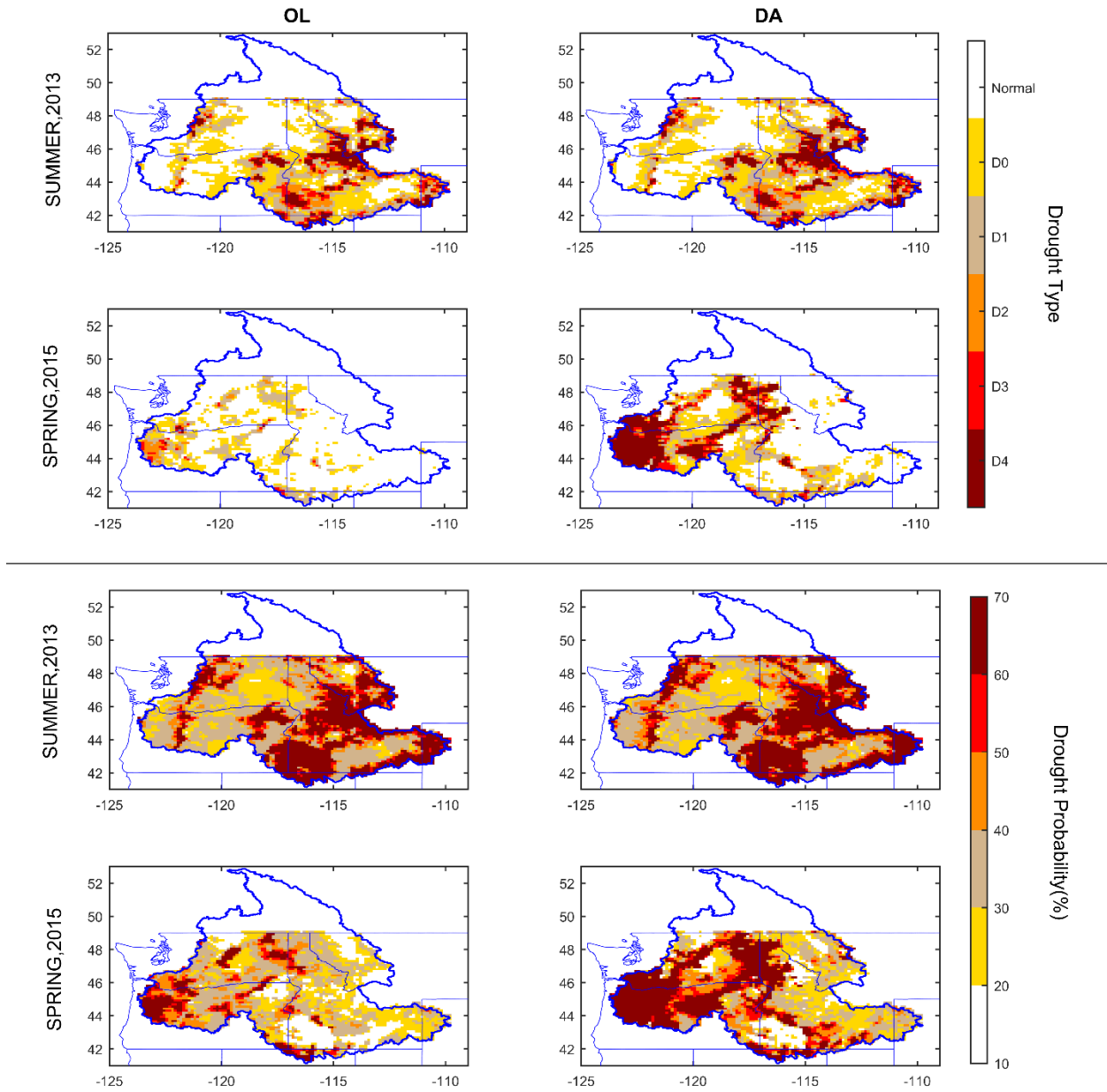




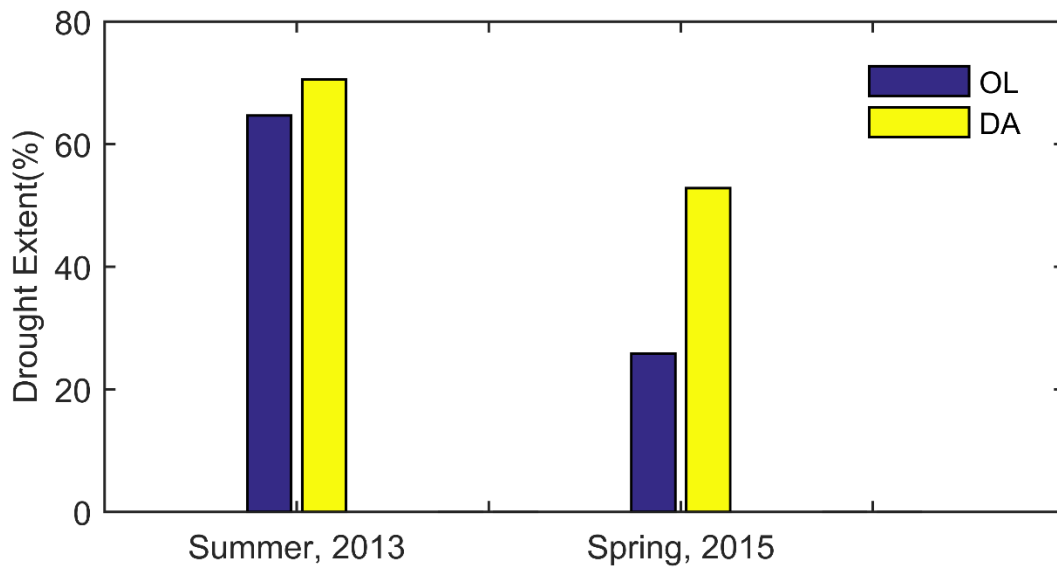
**Figure 8.** The 1st, 2nd, and 3rd probabilistic seasonal outlook results between OL and DA for summer 2015 (according to the framework shown in Figure 2). The inter-annual forecasting for spring and summer 2015 given the drought status in fall 2014 and winter 2015.



**Figure 9.** Comparison of the basin-averaged daily root-zone soil moisture ( $\text{m}^3/\text{m}^3$ ) by the open-loop (OL) and data assimilation (DA) for spring 2013 and winter 2015 across the CRB. The error bars show the 95% prediction intervals.



**Figure 10.** The seasonal probabilistic drought forecasts for summer 2013 and spring 2015 given the drought status in spring 2013 and winter 2015, respectively. The top-panel shows the forecasted drought areas based on MLE and the bottom-panel indicates the forecasted drought probabilities.



**Figure 11.** The forecasted drought extent between the OL and DA based on MLE across the CRB.