Agricultural and Forest Meteorology 321 (2022) 108982

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

An ensemble data assimilation approach to improve farm-scale actual evapotranspiration estimation

Proloy Deb *, Peyman Abbaszadeh, Hamid Moradkhani

Center for Complex Hydrosystems Research, Department of Civil, Construction and Environmental Engineering, The University of Alabama, Tuscaloosa, AL, 35487, USA

ARTICLE INFO

Keywords:
Data assimilation
Priestley-Taylor model
Evapotranspiration
Mobile river basin
Alabama
Particle filter

ABSTRACT

Estimation of actual evapotranspiration (ET) is key to irrigation water application and basin-scale agricultural water demand assessment. While modelers and water managers rely on stand-alone ET estimation model application in their planning and management, several uncertainties including model structure, parameter set, and initial condition exist, cascading in ET calculation leading to inaccurate results. In this study, an ensemble data assimilation approach is employed to explore the benefit of remotely sensed actual ET to improve the simulations of the widely used Priestley-Taylor ET model while accounting for uncertainties. The study is conducted at farm-scale for three different crops (corn, cotton, and soybean) in the Mobile River basin in Deep South United States, which has experienced severe droughts during the cropping seasons in the past. Prior to employing data assimilation, the Priestley-Taylor model is modified for each crop to simulate actual ET instead of reference ET. Following which the model is calibrated over 320,000 farms in the river basin for identifying the optimal parameters. The calibrated model is later used for the Open-Loop simulation, as well as in the development and implementation of data assimilation. The simulated and observed actual ET is used to calculate the Kalman gain and update the model initialization every time step during the assimilation period. The findings of the study showed that assimilating the actual ET observation into the Priestley-Taylor model results in more accurate and reliable model initialization and also posterior ET estimates at farm-scale compared to open-loop simulation. These results highlight, the importance of digital agricultural tools in robust agricultural planning and management and open door for further research.

1. Introduction

The continual surge in global food demand is putting immense pressure on food systems for boosting crop yields. One approach to attain this increasing food demand is by adopting irrigated farming (Pandey, 2019; Zhang et al., 2021). Majority of the agricultural operations in several countries across the world have already shifted from rainfed to irrigated agriculture including South and Southeast Asia, United States (US), China, and European nations (Puy et al., 2021). It is anticipated that by 2050, the irrigated areas across the globe may increase up to 800 Mha (Puy et al., 2020) relative to the current coverage of 367 Mha (Siebert et al., 2005). This projected escalation of irrigated arable land will lead to an immense increase in freshwater demand, yet, with the world approaching towards a water scarce future (Boretti and Rosa, 2019; Falkenmark, 2013), hence, sustainable and judicious water application is paramount. Irrigation water demand is a direct function of crop water consumption, also known as evapotranspiration (ET). Therefore, designing irrigation water projects including regional agricultural water demand and water allocation requires reliable and accurate estimates of actual ET (**Deb et al., 2022**). This is more crucial during droughts and regions experiencing persistent droughts **occurring across the world** such as in western United States (US) and Australia (Deb et al., 2019; Deb and Kiem, 2020; Hatchett et al., 2016; **Xu et al., 2019**). Moreover, it is also a key phenomenon in the water cycle. This is because globally actual ET contributes to returning approximately 60% of moisture into the atmosphere by using half of the solar energy absorbed by the earth's surface (Trenberth et al., 2007).

There are several methods for estimating actual ET including eddy covariance method (Mizutani et al., 1997), lysimeter (Fisher and Allen, 1991), and empirical modeling. While the first two methods are very location-specific and most accurate, the limitation is in their installation that is tedious and expensive. Whereas, models are mathematical representations of the interactions among different variables that drive the ET process. Several models have been developed since the 1940s and are

* Corresponding author.
E-mail address: pdeb1@ua.edu (P. Deb).

https://doi.org/10.1016/j.agrformet.2022.108982
Received 26 January 2022; Received in revised form 26 March 2022; Accepted 29 April 2022

widely used globally. Since actual ET varies with land cover and crops, reference ET is a more generalized variable, which is the ET from a reference crop, i.e., alfalfa (*Medicago sativa*) under given climatic conditions. Majority of the developed models calculate the reference ET and can be converted into actual ET for a specific crop by multiplying the stage-wise crop coefficients. While Food and Agricultural Organization (FAO-56) Penman-Monteith model (Allen et al., 1998) is considered as one of the most sophisticated models in estimating reference ET, it requires numerous forcing data including temperature, humidity, wind speed, solar radiation, soil heat flux, saturation, and actual vapor pressure (Deb et al., 2015; Yadav et al., 2016). Several developing nations still lack a good historical record of these variables. Although at global scale, gridded information is available for these variables, yet they are too coarse to represent the actual magnitude (Sun et al., 2018).

Alternatively, more simplistic models such as the Priestley-Taylor model (solar radiation and temperature-based) (Priestley and Taylor, 1972) and Hargreaves-Samani (temperature-based) (Hargreaves and Samani, 1985) models have also been evaluated in several locations and found equally good as the Penman-Monteith model. For instance, Utset et al. (2004), and Akumaga and Alderman (2019) compared Priestley-Taylor and Penman-Monteith models against observed reference ET in the Mediterranean region, and Oklahoma US, respectively. The results show an insignificant difference in the model simulated ET at daily time step for the earlier study and similar model performance at rainfed farms in humid regions for the latter.

While modelers tend to calibrate these models in estimating reference ET, often times the results are biased or erroneous due to several uncertainties including: 1) forcing data uncertainty due to measurement error and spatial and temporal representativeness of data; 2) model structural uncertainty due to imperfect representation of the system; 3) parameter uncertainty generating from the model conceptualization; and 4) initial or boundary condition uncertainty (Abbaszadeh et al., 2019a; Moradkhani et al., 2019; Pathiraja et al., 2018; Renard et al., 2010; **Xu et al., 2021a**). Therefore, it is well recognized that model simulations are robust within a probabilistic framework that estimates all the above mentioned uncertainties (Moradkhani et al., 2019). Bayesian inference is an approach to quantify this, and several Bayesian methods are well-acknowledged in the agricultural sector, especially crop yield predictions (Abbaszadeh et al., 2022; Besag and Higdon, 1999; Drury et al., 2017; Gavahi et al., 2021; Morrison et al., 2012; Shirley et al., 2020). Data assimilation is a technique that facilitates integration of observations into model simulations to result in the best model estimates. This is done by sequential Bayesian estimation methods within a probabilistic framework that constrains model predictions with the observations at a given time step (Vetra-Carvalho et al., 2018).

The Ensemble Kalman Filter (EnKF) algorithm (Crow and Wood, 2003; Reichle et al., 2002) is the widely used data assimilation technique in hydrology where the ensemble members or "particles" are updated based on linear equations. While this technique assumes linear and Gaussian distributions of the dataset, for non-Gaussian distributions, the Particle Filter (PF) algorithm provides the flexibility to not only generate random replicates of state variables and parameters but also updating the weights associated with these quantities (i.e. posterior) according to their importance in accurately estimating or forecasting the model outputs. Further statistical inferences can be made based on the ensemble of realizations (Moradkhani et al., 2005). Furthermore, in addition to the mean and covariance, a thorough representation of the posterior distribution can be availed for the ensemble of realizations using PF algorithm, and therefore, it is more suited for highly non-linear dynamical systems (Dechant and Moradkhani, 2012; Moradkhani et al., 2019). In this study, PF in conjunction with Variable Variance Multiplier (hereafter PF-VVM) is used to assimilate the actual ET observations into the modified Priestley-Taylor model to improve its predictive skill and explore the benefit of utilizing data assimilation in agricultural planning and management.

Majority of the data assimilation techniques were developed and employed by focusing on improving hydrologic modeling while accounting for uncertainties (Abbaszadeh et al., 2021, 2018; Bahrami et al., 2021; Gavahi et al., 2020; Jafarzadegan et al., 2021; Kim et al., 2021; Li et al., 2015; Schumacher et al., 2018; Xu et al., 2020; Zhang et al., 2021). Similarly, in the agricultural sector, several of these data assimilation algorithms were employed for improving the model predictions. For instance, remote sensing based soil moisture was assimilated into crop models using the EnKF algorithm for improving crop yield (de Wit and van Diepen, 2007; Ines et al., 2013; Zhuo et al., 2019; Ziliani et al., 2022) and agricultural drought prediction (Bolten et al., 2010). Silvestro et al. (2017) assimilated canopy cover into crop models using the EnKF approach for improving crop yield estimation. On the other hand, Huang et al. (2015) and Mokhtari et al. (2018) assimilated Leaf Area Index (LAI) using a sophisticated four-dimensional variations data assimilation algorithm into crop models to improve yield estimation. Also, Li et al. (2014) improved corn yield estimations of coupled WOFOST-HYDRUS-1D model by assimilating LAI using the EnKF data assimilation algorithm.

These studies illustrate that the application of data assimilation algorithms in the agricultural sector is still in its preliminary stage and the primary focus of the studies is crop yield estimation. Another major drawback of the existing studies is employing the EnKF algorithm, which assumes a linear updating rule and a Gaussian distribution of observation and model errors, which may not always be true for the considered variables. Given these knowledge gaps and the importance of actual ET estimation, the objective of this paper is to assimilate observed ET into a widely used ET estimation model (Priestley-Taylor) for improving the actual ET estimates at farm-scale. There are two novelties in this study: 1) data assimilation of satellite-based observed ET into ET estimation models (as no other study has investigated this), and 2) employing PF-VVM of data assimilation in the agricultural sector, which has not been used before. The findings of this study will be useful for regional agricultural water resources planning and management.

2. Study area and datasets

The Mobile River Basin located in the southeast US spans over four states including Tennessee, Georgia, Mississippi, and Alabama (Fig. 1). The river basin comprises an area of 115,200 km^2 and approximately two-thirds of the basin is contained in Alabama. It is the fourth largest river basin in the US and contributes to an average annual discharge of 1760 m^3/s. The average annual temperature and precipitation range from 15 °C to 21 °C and 1270 mm to 1524 mm, respectively. Sandy loam soil dominates the river basin, however, the central region also comprises of silty loam soil (Mitchell et al., 2004). The average annual actual evapotranspiration is approximately 1023 mm across the river basin (Senay et al., 2013). The dominant land cover in the river basin is forest (~60%), and agriculture contributes to 26% of the land cover (Warner et al., 2005). Crop acreage study suggests corn, cotton, and soybean are the most widely grown crops in the basin (Harned et al., 2004). Generally, corn is sown between mid-March and is harvested mid-July, whereas, cotton and soybean are sown in mid-April and mid-May, and harvested in mid-September and mid-October, respectively (Norfleet et al., 1997). Also, it is to be noted that the average farm size across the Mobile River basin is approximately 161 ha (~398 acres) (USDA, 2020).

Several datasets were required to pursue the study which are reported in Table 1. It is to be noted that due to the unavailability of AmeriFlux tower sites within the river basin, the actual ET used in this study is a satellite-derived dataset from the operational Simplified Surface Energy Balance (SSEBop) by the United States Geological Survey (USGS). The dataset is derived by blending the ET fractions from the Moderate Resolution Imaging Spectroradiometer (MODIS) imagery and reference ET from a thermal index approach. This dataset has been widely used across the US and has been reported to perform well when compared to *in situ* measurements, especially in croplands (Chen et al.,

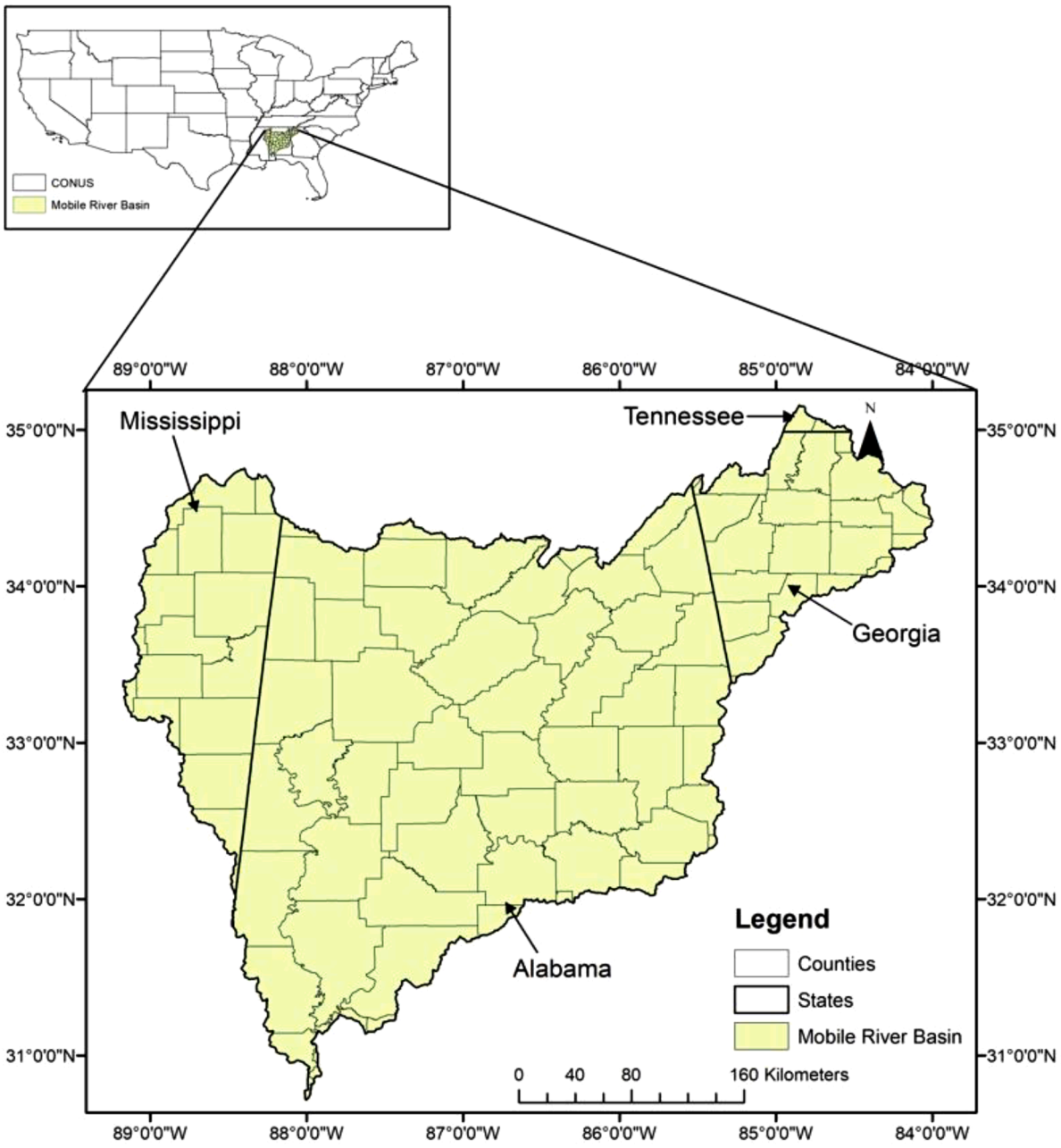

Fig. 1. Location of Mobile River Basin within southeast US.

2016). Also, it is worth noting that the cropland data layer which comprises of the annual agricultural croplands across the Contiguous US was used in this study for the year 2018.

3. Methodology

This section first describes the steps taken to extract the forcing variables at farm-scale from the cropland data layer. Following this, the ET estimation model is summarized, i.e. Priestley-Taylor equation, and its modifications in this study, then the calibration/validation of the model. This is followed by the steps of data assimilation of the Priestley-Taylor model. **A Bayesian data assimilation approach was employed where the prior and the posterior probability distributions of the variables were obtained, and the predictive uncertainty was characterized. A schematic data assimilation framework employed in this study is shown in Fig. 2 and further details on the approach are given in Section 3.4.**

3.1. Extraction of input variables at farm-scale

Since in this study the analysis was conducted at farm-scale, all the forcing variables were required at the same spatial scale. To do this, first, the cropland data layer (which was in raster format) was converted to polygons, where each polygon represented a farm. This resulted in over 224 million polygons across the river basin comprising of 59 crops. Since the target crops were corn, cotton, and soybean, the polygons corresponding to these three crops were extracted and stored. Thereafter, the forcing variables (incoming shortwave radiation, minimum and maximum temperature, and minimum and maximum relative humidity) of different spatial resolutions (as in Table 1) were regridded to the spatial resolution of the output variable, i.e., observed actual ET of 900 m. **The regridding was done by using the resampling tool in ArcPy**

Table 1
The datasets used in this study with their spatial and temporal resolutions and their sources.

Data	Spatial resolution	Temporal resolution	Source
Observed actual ET	900 m	Daily	SSEBop, USGS (Senay et al., 2013)
Incoming shortwave radiation	4 km	Daily	gridMET (Abatzoglou, 2013)
Maximum temperature	4 km	Daily	gridMET (Abatzoglou, 2013)
Minimum temperature	4 km	Daily	gridMET (Abatzoglou, 2013)
Maximum relative humidity	4 km	Daily	gridMET (Abatzoglou, 2013)
Minimum relative humidity	4 km	Daily	gridMET (Abatzoglou, 2013)
Elevation	30 m	-	National Elevation Dataset, USGS
Cropland data layer	30 m	-	United States Department of Agriculture (USDA) (Johnson, 2019)

using the bilinear interpolation since it accounts for the weighted distance average of the nearest four grids in assigning the value. It is worth noting that since the average farm size is 161 ha in the river basin, each pixel (**810,000 m^2 or** 81 ha) of the regridded variables was well within the farm size.

3.2. Priestley-Taylor ET model

The Priestley-Taylor model (Priestley and Taylor, 1972) was employed for the calculation of reference ET in this study since it was identified to be the best performing model at a daily time step in a suit of six models within the southeast US (Lu et al., 2005). The basic model structure is given in Eq. (1).

$$ET = \alpha \times \frac{\Delta \times (R_n - G)}{\lambda \times (\Delta + \gamma)} \quad (1)$$

where ET is reference ET in mm/day, α is model parameter used in model calibration, Δ is slope of saturated vapor pressure, R_n is net solar radiation in W/m^2, G is soil heat flux in W/m^2 and can be considered as zero at a daily time step (Purdy et al., 2016), λ is latent heat of vaporization of water which is constant (2.26 MJ/kg), and γ is the psychrometric constant.

The variables Δ, Rn, and γ need further calculation and the basic equations are provided here in Eqs. (2), (3), and (5) respectively.

$$\Delta = \frac{2504 \times exp^{\left(\frac{17.27 \times T}{T+237.2}\right)}}{(T + 237.2)^2}. \quad (2)$$

where T is the average temperature in °C calculated from the minimum and maximum temperatures.

$$R_n = R_{ns} - R_{nl} \quad (3)$$

where R_{ns} and R_{nl} are net shortwave and longwave radiation in W/m^2. R_{ns} is calculated as (1- α) $\times$ R_s, where R_s is the incoming shortwave radiation, and R_{nl} is calculated as in Eq. (4).

$$R_{nl} = \sigma \times \left\{ \frac{(T_{max} + 273.2)^4 + (T_{min} + 273.2)^4}{2} \times (0.34 - 0.14\sqrt{e_a}) \times \left(1.35 \times \frac{R_s}{(R_{so} - 0.35)}\right) \right\} \quad (4)$$

where σ is Stefan-Boltzman constant (5.67 $\times$ 10^{-8} W/m^2/K^4), T_{max} and T_{min} are maximum and minimum temperature in °C, e_a is actual vapor pressure in kPa, R_s and R_{so} are intermediate variables which are calculated based on different forcing variables and more details can be found in De Bruin and Keijman (1979).

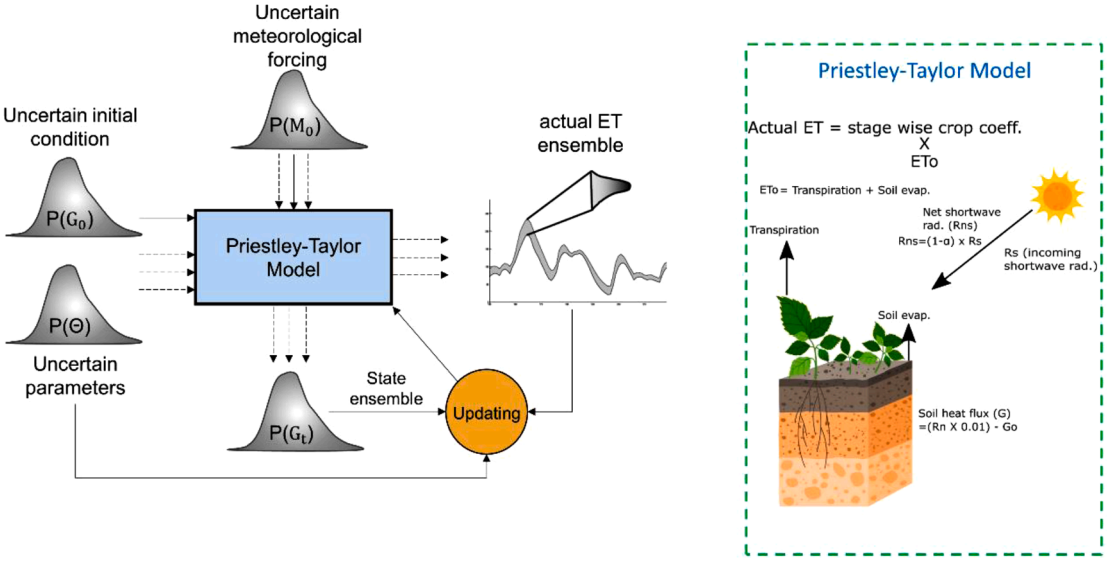

Fig. 2. A schematic of the framework of the study. The box in the right-hand side displays a schematic of the Priestley-Taylor model.

$$\gamma = 0.00163 \times \frac{P}{2.264} \tag{5}$$

where P is an intermediate variable and calculated as in Eq. (6).

$$P = 101.3 \times \left\{\frac{(293 - 0.0065) \times h_{Geo}}{293}\right\}^{5.26} \tag{6}$$

where h_{Geo} is the elevation in m.

It is to be noted that the original Priestley-Taylor model simulates the reference ET, however, due to the requirement of the study, the model was modified by multiplying the reference ET with stage wise crop coefficients so that the final outcome is actual ET (modified Priestley-Taylor model called from here on). The crop coefficients were derived from Allen et al. (1998) for corn, cotton, and soybean crops for the four crop stages, initial, development, middle, and late. Additionally, the modified Priestley-Taylor model is static in nature, i.e. for the forcing data at a given day, the model simulates the reference ET for the corresponding day and there is no relationship in the inherent processes or calculations. This may not be true at locations with low annual precipitation or during low-precipitation months since the net soil heat flux may be close to zero during the start of the simulation, however, during low precipitation months, the flux can accumulate into higher values. Therefore, to rectify this, in this study incoming shortwave radiation that drives the soil heat flux was considered as a state variable, meaning the past day's condition of the incoming shortwave radiation will affect the present day, and thus the rest of the model calculation. To account for this in this study, from literature one percent of the net solar radiation is assumed to be absorbed by the soil at any given day (Fig. 1 from Trenberth et al., 2009). Also, to relate the antecedent soil heat conditions with the following day Eq. (7) was employed in this study.

$$G = 0.01 \times R_n - G_0 \tag{7}$$

where G and R_n are the present day absorbed soil heat and net solar radiation in W/m^2, and G_0 is the past day's absorbed soil heat in W/m^2. For the initialization of the model run G_0 was considered as 0.

3.3. Modified Priestley-Taylor model calibration

The modified model was first calibrated for each polygon for all three crops using a global optimization technique, the Shuffled Complex Evolution algorithm (Duan et al., 1994). This algorithm first stochastically distributes a sample of points within the parameter space using the lower and upper parameter bounds. Then each sample point is treated as a member of population and a unique genetic information is assigned to each individual. Thereby the population tends to attain an optimum value by altering the genetic information among themselves (parameter values). Here the optimum value is the user defined objective function corresponding to the model simulated and observed values. The initial sample is also sub-divided into several smaller samples (called complexes). Theoretically, each complex combines with another to produce offspring using the simplex procedure of Nelder and Mead (1965). Since the fitness of an individual contributes to the likelihood of an offspring, the older points are replaced by the younger ones. This proceeds to a global optimum which is assisted by (a) probability that the newer points are created automatically within the parameter space, and (b) a continuous combination of points into new complexes. Further details on the Shuffled Complex Evolution algorithm and its use in model calibration can be found in Duan et al. (1994, 1993).

The modified Priestley-Taylor model was calibrated for α parameter against the observed actual ET at each polygon. Generally, α =1.26 is used for the parameter in model simulations (Priestley and Taylor, 1972; Stewart and Rouse, 1977), however, it can vary widely across different land use types and climatic zones. For instance, for humid conditions, the parameter can be lower than one, whereas, in arid conditions, it can be more than two. Therefore, in this study to define the parameter space, the lower and upper bounds were set as 0.50 and 2.0. Several objective functions were used for the model calibration which are discussed in detail in Section 4. The model was calibrated for all three crops for the cropping season of 2018 for each crop.

3.4. Data assimilation

The basics of the data assimilation approach, i.e., PF-VVM used in this study is described here. Following previous work by Moradkhani et al. (2005 and 2019), the differential equations that illustrate the generic nonlinear dynamic system in a hydrologic system are described as follows:

$$x_t = f(x_{t-1},\ u_t,\ \theta) + \omega_t \tag{8}$$

$$y_t = h(x_t) + \upsilon_t \tag{9}$$

where $x_t \in \mathbb{R}^n$ is a vector of the uncertain state variables (incoming shortwave radiation in this study) at time step t, u_t is the uncertain forcing inputs (which are minimum and maximum temperature, and minimum and maximum relative humidity), $\theta \in \mathbb{R}^d$ is a vector of model parameters (α in this study), $y_t \in \mathbb{R}^m$ is a vector of observation data (observed actual ET), ω_t represents the model errors due to the imperfect model, υ_t is the measurement error. In most cases, ω_t and υ_t were assumed as white noises with mean zero and covariance Q_t and R_t, respectively. Furthermore, the two noises ω_t and υ_t were assumed to be independent. Because Moradkhani et al. (2005) and Moradkhani et al. (2019) provided a complete literature of the sequential Bayesian filtering formalism, this study just provides a brief overview. The posterior distribution of the state variables at time t, according to Bayes' Law, is as follows:

$$p(x_t|y_{1:t}) = p(x_t|y_{1:t-1}, y_t) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{\int^p (y_t|x_t)p(x_t|y_{1:t-1})dx_t} \tag{10}$$

$$p(x_t|y_{1:t-1}) = \int^p (x_t,\ x_{t-1}\ |y_{1:t-1})dx_{t-1} = \int^p (x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1} \tag{11}$$

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

$$p(y_{1:t}) = p(y_1)\prod^p (y_t|y_{1:t-1}) \tag{12}$$

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

$$p(y_t|y_{1:t-1}) = \int^p (y_t,\ x_t\ |y_{1:t-1})dx_t = \int^p (y_t|x_t)p(x_t|y_{1:t-1})dx_t \tag{13}$$

Only in special cases, such as the linear system with Gaussian assumption of noises in the system (i.e., the Kalman filter), an analytic solution is available for Eq. (10). In most cases, a group of random samples were used to approximate the posterior distribution.

$$p(x_t|y_{1:t}) \approx \sum_{i=1}^{N} w^{i+}\delta(x_t - x_t^i) \tag{13}$$

where w^{i+} is the posterior weight of the i-th particle, δ is the Dirac delta function, and N is the number of particles. The normalized weights were calculated using:

$$w^{i+} = \frac{w^{i-}.p(y_t|x_t^i,\ \theta_t^i)}{\sum_{i=1}^{N} w^{i-}.p(y_t|x_t^i,\ \theta_t^i)} \tag{14}$$

where w^{i-} is the prior particle weights, and the $p(y_t|x_t^i,\ \theta_t^i)$ was be computed from the likelihood $L(y_t|x_t^i, \theta_t^i)$. Generally, a Gaussian likeli-

hood was used to estimate $L(y_t|x_t^i,\theta_t^i)$:

$$L(y_t|x_t^i,\theta_t^i)=\frac{1}{\sqrt{(2\pi)^m|R_t|}}exp\left[-\frac{1}{2}(y_t-h(x_t^i))^T R_t^{-1}(y_t-h(x_t^i))\right] \quad (15)$$

A resampling procedure is needed to acquire approximate samples from $p(x_t|y_{1:t})$. To resample particles with a probability greater than the uniform probability, the sequential importance resampling (SIR) algorithm was used. The particles with lower weights were eliminated using this procedure, while the particles with higher weights were kept. After application of the SIR algorithm, all the particle weights were set equal to $1/N$. Particles with large weights were more likely to be drawn several times during resampling, resulting in a loss of particle diversity. A perturbation of the resampled parameters is recommended to avoid sample impoverishment. Then, a proposal distribution was formed to generate proposed parameters $\theta_t^{i,p}$:

$$\theta_t^{i,p}=\theta_t^{i+}+\varepsilon_t^i \qquad \varepsilon_t^i\sim N\left[0,s_t Var(\theta_t^{i-})\right] \quad (16)$$

where θ_t^{i+} is the parameters after SIR, $Var(\theta_t^{i-})$ is the variance of the prior parameters at the current time step, and s_t is a small tuning time-variant parameter. Since the optimal tuning factor s_t was unknown in a sequential framework, it was beneficial to treat the s_t as a time-variant variable and estimate it automatically. Moradkhani et al. (2012) proposed an approach which automatically finds the most fitting tuning factor s_t in Eq. (16). The tuning factor is self-adaptive based on the spread of the previous states. The procedure of VVM includes:

$$\widehat{\varepsilon}_t=\left|E(y_t^{'})-y_t\right| \quad (17)$$

$$ub_t=\begin{cases} y_t^{'75}-E(y_t^{'}) & E(y_t^{'})<y_t \\ E(y_t^{'})-y_t^{'25} & E(y_t^{'})>y_t \end{cases} \quad (18)$$

$$er_t=\tau\left[\text{median}\left(\frac{\varepsilon_{(t-lag):t}^{'}}{ub_{(t-lag):t}}\right)-1\right]+1 \quad (19)$$

$$s_t=er_t.E\left[s_{(t-lag):t}\right] \quad (20)$$

where $E(y_t^{'})$ is the forecast expected value, y_t is the observation, $y_t^{'25}$ and $y_t^{'75}$ are the 25th and 75th forecast percentiles, respectively; τ is the smoothing value and was set to 0.5 as per recommendation of Abbaszadeh et al. (2018).

Prior to employing the data assimilation, an open-loop simulation of the modified Priestley-Taylor model was done for the cropping periods of the three crops at all polygons within the study basin. For both open-loop and data assimilation scenarios, an ensemble size of 100 was used. The ensemble mean of actual ET simulations was derived with and without using data assimilation for all the polygons using different performance measures explained in the following section. It should be noted that the ensemble members were generated by introducing noise of 10% (percentage of error) to the forcing variables (minimum and maximum relative humidity, and minimum and maximum temperature), and the state variable (net solar radiation; noise was introduced to the daily values calculated by the model during the calibration). Please note that in the data assimilation scenario, the model parameters were used which were calibrated before (and used in the open-loop simulation), and added noise to them to ensure avoiding the sample impoverishment during the assimilation process. This was done according to Eq. (16). In other words, the calibrated model parameter was used as an initial guess (*priori*) for the dynamic Priestley-Taylor model and then sequentially updated it in conjunction with model state-variable (i.e., net solar radiation) during the assimilation period. In this study, white noise was added with mean zero and variance equal to 10 percent of the actual value.

4. Performance measures

In this study, two deterministic performance measures were used to assess the model performance during model calibration (Mean Absolute Error (Moriasi et al., 2007), and Pearson correlation coefficient (Pearson, 1895)), and additional two performance measures (Root Mean Square Error (Moriasi et al., 2015), and Kling-Gupta Efficiency (Gupta et al., 2009)) were used to evaluate the data assimilation of the model. Furthermore, a probabilistic measure (*Reliability* (Renard et al., 2010)) was also used to investigate the probabilistic actual ET simulation by the data assimilation of the modified Priestley-Taylor model. These measures are defined in Eqs. (21)–(25).

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|y_i-x_i| \quad (21)$$

$$r=\frac{\sum^{(x_i-\overline{x})}(y_i-\overline{y})}{\sqrt{\sum^{(x_i-\overline{x})^2}\times\sum^{(y_i-\overline{y})^2}}} \quad (22)$$

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(y_i-x_i)^2}{n}} \quad (23)$$

$$KGE=1-\sqrt{\left[\left(\frac{Cov_{x,y}}{\sigma_x\sigma_y}\right)-1\right]^2+\left[\left(\frac{\sigma_x}{\sigma_y}\right)-1\right]^2+\left[\left(\frac{\mu_x}{\mu_y}\right)-1\right]^2} \quad (24)$$

$$Reliability=1-\frac{2}{T}\sum_{t=1}^{T}\left|\frac{Z_t}{T}-U_t\right| \quad (25)$$

where *MAE* is Mean Absolute Error in mm/day, n is the total number of observations, y_i is observed value of ith observation, x_i is model simulated value of ith observation. r is Pearson correlation coefficient, $\overline{y}$ and $\overline{x}$ represent the mean observed and model simulated values. *RMSE* is Root Mean Square Error in mm/day. *KGE* represents Kling-Gupta efficiency, $Cov_{x,y}$ denotes the covariance calculated between the model simulated and observed values; σ_x, and σ_y represent the standard deviations of the model simulated and observed values respectively. μ_x and μ_y denote the average of the model simulations and observed values respectively. Z_t represents the score corresponding to quantile of each observation time step (T), and U_t represents the uniform distribution between 0 and 1. While lower value (close to zero) is ideal for *MAE*, and *RMSE*, higher value (close to 1) is ideal for *r*, *KGE*, and *Reliability*. In this study *MAE* was employed as the objective function for the calibration of the modified Priestley-Taylor model, whereas, for model performance evaluation *MAE* and Pearson correlation coefficient was employed. Similarly, for evaluating the performance of the open-loop simulation and data assimilation of the modified Priestley-Taylor model, *KGE*, *RMSE*, and *MAE* were used. For evaluation of the probabilistic simulation of the model at specific locations, *Reliability* was employed.

5. Results and discussion

As mentioned earlier, the overarching goal of this study is to improve the actual ET estimation of the Priestley-Taylor model at farm-scale by employing the data assimilation technique. In order to achieve this, the first step is to calibrate the model at farm-scale and the performance of the model without considering any uncertainty in the forcing data and model initialization. The results are shown in Fig. 3. The model performance was evaluated using two deterministic performance measures, i.e., *MAE* and Pearson correlation coefficient. The results indicate that the *MAE* for corn ranges from 0.62 mm/day to 2.45 mm/day across the Mobile River basin, with higher values in the west-central and the south-central region. Lower magnitudes are noted for the western, northern, and eastern farms (Fig. 3(a)). For cotton and soybean, the *MAE* ranges

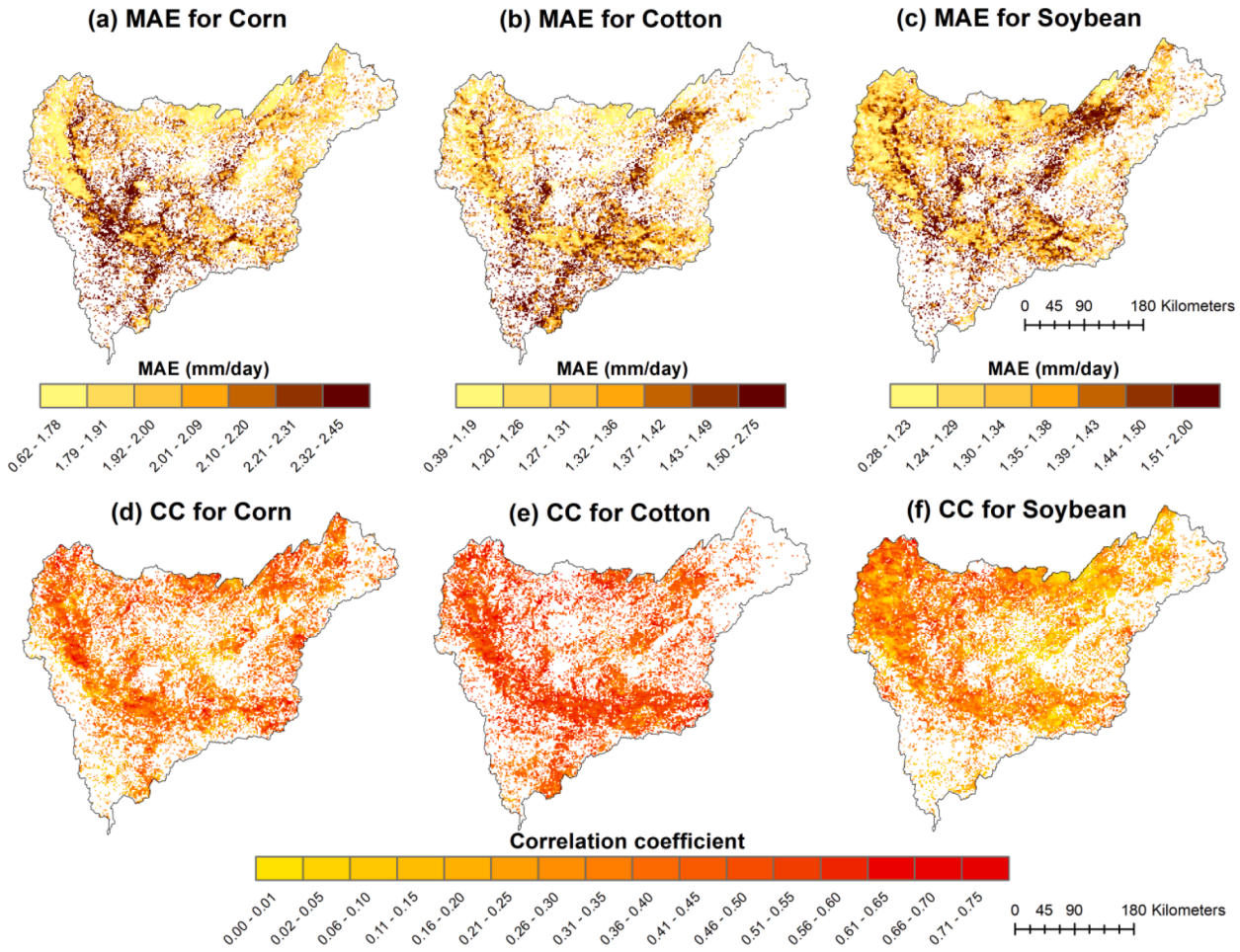

Fig. 3. Spatial variability of *MAE* and Pearson correlation coefficient calculated for modified Priestley-Taylor model during its calibration (in 2018) for corn, cotton, and soybean at Mobile River basin. Note: each dot represents a farm; CC stands for Pearson correlation coefficient.

from 0.39 mm/day to 2.75 mm/day, and 0.28 mm/day to 2.00 mm/day, respectively (Fig. 3(b) and (c)). For both crops (cotton and soybean), a higher magnitude of *MAE*, within the range of 1.32 mm/day to 2.75 mm/day for cotton, and 1.35 mm/day to 2.00 mm/day for soybean are observed for the southern, central, and the central-east farms. A relatively higher number of farms with low *MAE* are observed in the eastern region for both crops.

The Pearson correlation coefficient for the model calibration ranges from 0 to 0.55, 0 to 0.75, and 0 to 0.40 for corn, cotton, and soybean crops, respectively (Fig. 3(d) to (e)). Lower magnitude of the correlation coefficient (within the range of 0 to 0.30) is more dominant across the river basin for corn and soybean crops. On the other hand, for cotton, a relatively higher magnitude 0.36 to 0.65 is noted to be dominant (Fig. 3 (e)). It is worth noting that in the majority of farms, a lower *MAE* corresponds to a higher Pearson correlation coefficient. Overall these results indicate that in a majority of the farms model performance is unsatisfactory solely based on the optimization of the model parameter α. Similar results of poor to moderate performance of the Priestley-Taylor model were also reported by Islam and Alam (2021) and Yang et al. (2016) in humid and semi-humid regions of Bangladesh and China, respectively. One of the primary causes of the poor model performance is due to the fact that the model was developed for saturated and open water body conditions (De Bruin and Keijman, 1979) by eliminating the product of wind function and vapor pressure deficit from the Penman-Monteith equation. Since croplands do not correspond to these ideal conditions as assumed in the model and are water/energy-limited (Javadian et al., 2020), the model performance declines.

Fig. 4 displays the heterogeneity in *MAE* calculated for the modified Priestley-Taylor model during calibration at three randomly selected sites with corn, cotton, and soybean farmlands. Clearly, at the selected locations, *MAE* ranges from 2.021 mm/day to 2.220 mm/day, 1.151 mm/day to 1.350 mm/day, and 1.271 to 1.530 mm/day for corn, cotton, and soybean crops respectively. This indicates that at the selected sites the best performance for the model is obtained for the cotton crop, followed by soybean, and corn, yet overall, the performance is quite poor. This is because error within the range of 0 – 25% with respect to the observed actual ET for a given period is considered acceptable (Melton et al., 2021), whereas in this study at several locations the *MAE* $>$ 45%. It is to be noted that these results are obtained for the best parameter value optimized using the Shuffle Complex Evolution algorithm and minor departure from the optimal value results in worsening the model performance.

As mentioned in the last paragraph of Section 3.4, before conducting the data assimilation on the modified Priestley-Taylor model, an open-loop simulation analysis was performed. The mean ensembles for the open-loop simulation are presented in the form of a heat scatter plot in the top row of Fig. 5. As per results, the model performs similarly for the three crops with close values for the deterministic performance measures. In the case of corn, *KGE, RMSE*, and *MAE* are observed to be 0.47, 1.20 mm/day, and 0.97 mm/day respectively. Similarly, for cotton and soybean the values are noted to be 0.46, 1.28 mm/day, and 1.05 mm/day, and 0.46, 1.17 mm/day, and 0.95 mm/day respectively. Although, these performance measures indicate superior performance over the model calibration results (as presented in Figs. 3 and 4), yet the results are moderate (Melton et al., 2021; Moriasi et al., 2007). Also, the darker red color in case of soybean is due to the larger number of pixels (farms)

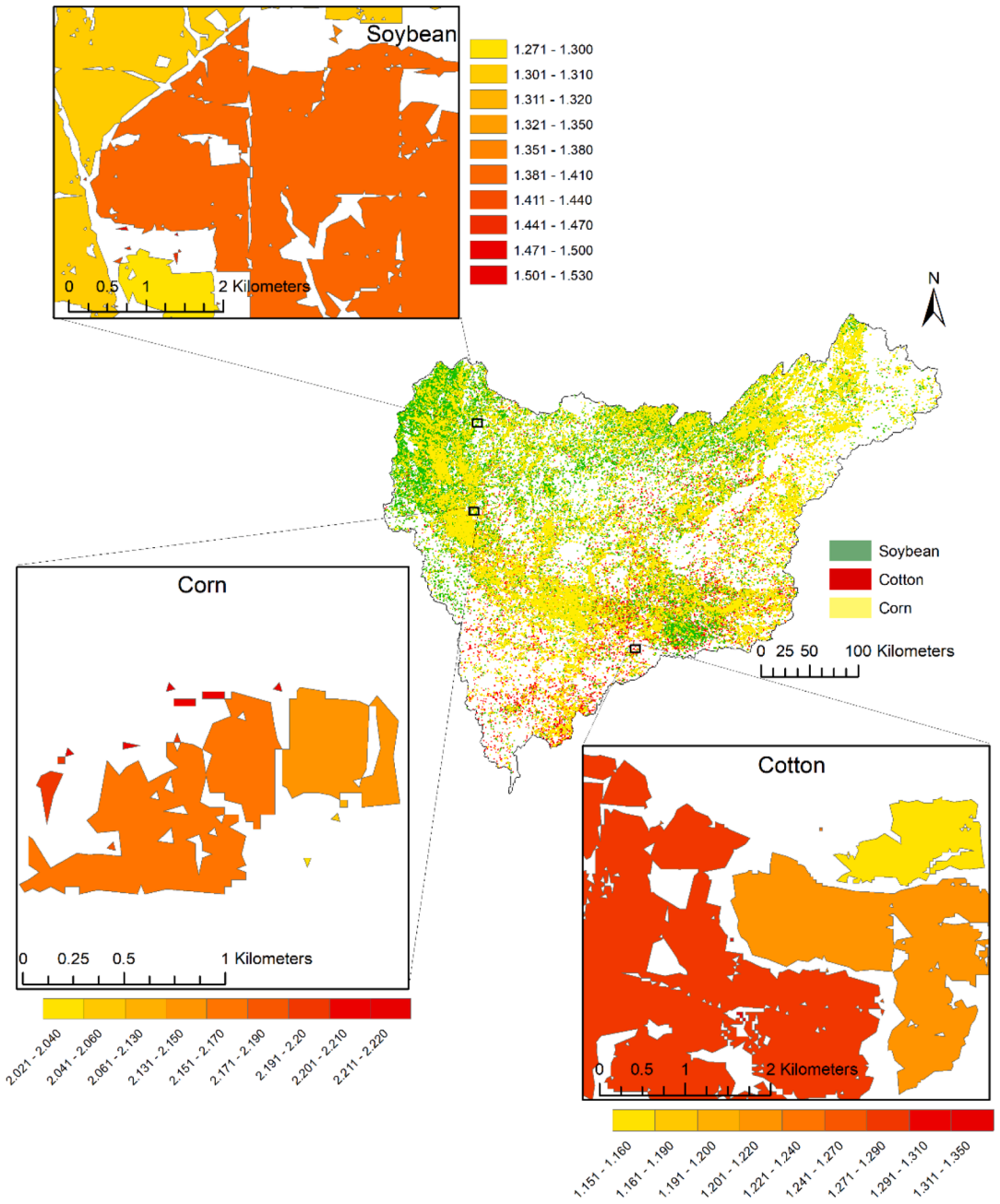

Fig. 4. *MAE* calculated for modified Priestley-Taylor model calibration at three randomly selected locations growing corn, cotton, and soybean crops within the study area. Note: each polygons in the zoomed frames represents a farm.

in the case of soybean relative to corn and cotton cropped farms. Here, it is noted that although open-loop simulations represent the prior distributions of the outputs (actual ET here), the model run is dependent on the initial conditions which are updated based on the observed values (observed actual ET here) from the previous time step (Abbaszadeh et al., 2020; Jafarzadegan et al., 2021). Similar improvements in ET estimation and water balance estimation of models for open-loop simulation over model calibration were also obtained by Khaki et al. (2020) and Tian et al. (2021) respectively.

The bottom row of Fig. 5 illustrates the assimilated actual ET vs observed actual ET for the three crops and the performance measures. A significant improvement in the actual ET estimates can be seen here compared to the open-loop simulations with the concentration of bins near the diagonal 1:1 line for all three crops. Moreover, the calculated *KGE, RMSE*, and *MAE* are noted to be 0.82, 0.59 mm/day, and 0.49 mm/day for corn, 0.84, 0.59 mm/day, and 0.50 mm/day for cotton, and 0.82, 0.58 mm/day, and 0.48 mm/day for soybean respectively. The improved performance is resulting from the posterior distributions of the actual ET in the data assimilation which accounts for the prior distributions from the open-loop simulation and the updated observations. Given that the posterior distribution of the actual ET is generated for each polygon, resulting in the spatial representation of actual ET data assimilation, overall, higher performance is noticed for the entire basin compared to other studies where soil moisture is assimilated for streamflow prediction such as in Jadidoleslam et al. (2021).

Fig. 6 illustrates the spatial variability of *MAE* and Pearson correlation coefficient calculated for the assimilated actual ET compared to the observed actual ET at the study basin. The results indicate that for a majority of the polygons (farms), *MAE* is in the range of 0.448 – 0.501 mm/day, 0.460 – 0.510 mm/day, and 0.466 – 0.503 mm/day for corn,

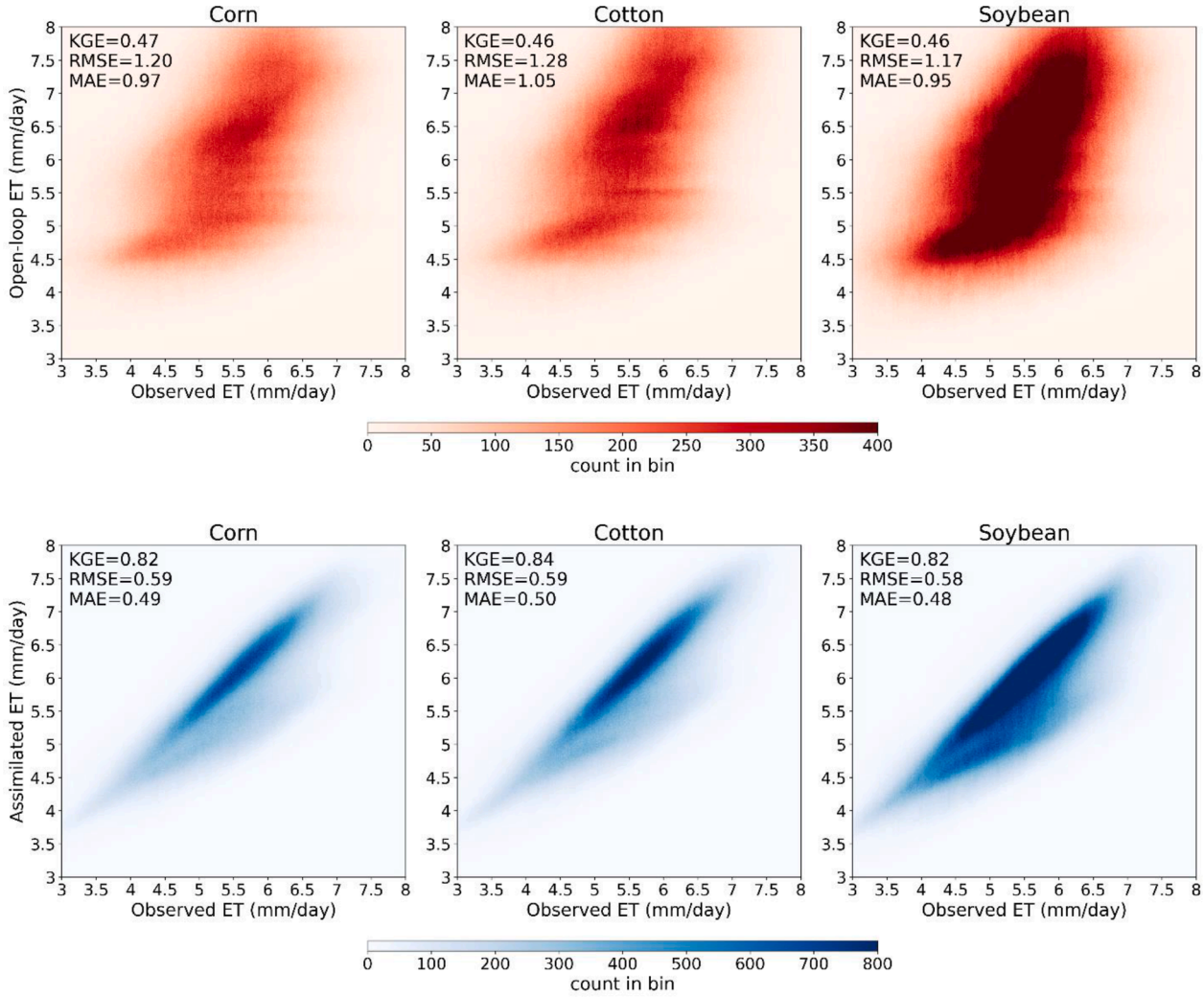

Fig. 5. Heatscatter plot of the observed actual ET and open-loop simulation of the modified Priestley-Taylor model (top row), and assimilated ET (bottom row) for the three crops across the study area. Note: The color intensity represents the density of the points in scatterplot.

cotton, and soybean respectively. Additionally, the Pearson correlation coefficient ranges from 0.823 – 0.907, 0.834 – 0.859, and 0.813 – 0.840 for the corresponding crops. A significant improvement in these performance measures can be seen when compared to the ones obtained for the model calibration results as shown in Fig. 3. Similarly, spatially improved results were also reported for soil moisture using data assimilation techniques in other studies such as Abbaszadeh et al. (2019b) and Seo et al. (2021).

To better display the farm-scale performance of the data assimilation, the calculated *MAE*s for the three different crops are zoomed in and are presented in Fig. 7. The *MAE* for corn farms ranges from 0.424 to 0.531. Similarly, for cotton and soybean, it ranges from 0.453 to 0.555 and 0.423 to 0.544 respectively. This implies a significant improvement over the *MAE* calculated for the model calibration results obtained in Fig. 4, where the corresponding *MAE* ranged from 2.021 to 2.220, 1.151 to 1.350, and 1.271 to 1.530. Given that these *MAE* ranges are well within the recommended range of 0 – 25% of the observed actual ET for a given period, the findings suggest that assimilating the actual ET in the model simulation significantly improves the model results at the farm-scale.

Evaluation of the data assimilation involves satisfactory performance for both deterministic and probabilistic performance measures. Therefore, in this study, three deterministic measures (*KGE, RMSE*, and *MAE*) and a probabilistic measure (*Reliability*) were used to compare the observed actual ET and the assimilated actual ET from the modified Priestley-Taylor model. Moreover, the *Reliability* was also calculated for the open-loop model simulation. The results in Fig. 8 are presented for three locations with the largest polygons of Figs. 4 and 7. Here the time-series plots indicate that the ensemble means of the open-loop simulation overestimate (dashed green line) the observed actual ET at all three locations for the three crops, especially during the peak ET months (i.e., May – mid-June for corn, July – August for cotton, and July – September for soybean). The results also indicate a significant improvement in the assimilated actual ET with *KGE, RMSE*, and *MAE* in the range of 0.69 to 0.78, 0.53 mm/day to 0.57 mm/day, and 0.44 to 0.50 mm/day, respectively for the three crops.

The right column in Fig. 8 illustrates Q-Q plots for comparing the *Reliability* of open-loop simulation and the assimilated actual ET at the three locations (mentioned above) for all three crops. An ideal Q-Q plot should have an S-shape and *Reliability* index should be 1. Among the three crops, the best results are obtained for the cotton crop, followed by corn, and soybean in case of open-loop simulated and assimilated actual ET. Overall, the *Reliability* is observed to be 0.95, 0.78, and 0.67 for cotton, corn, and soybean, respectively, and these values reflect the satisfactory performance of the probabilistic performance measure for the assimilated ET for all three crops (Gavahi et al., 2020; Renard et al., 2010). A closer look into the time series for all three crops indicates that the assimilated actual ET results are slightly overestimated relative to

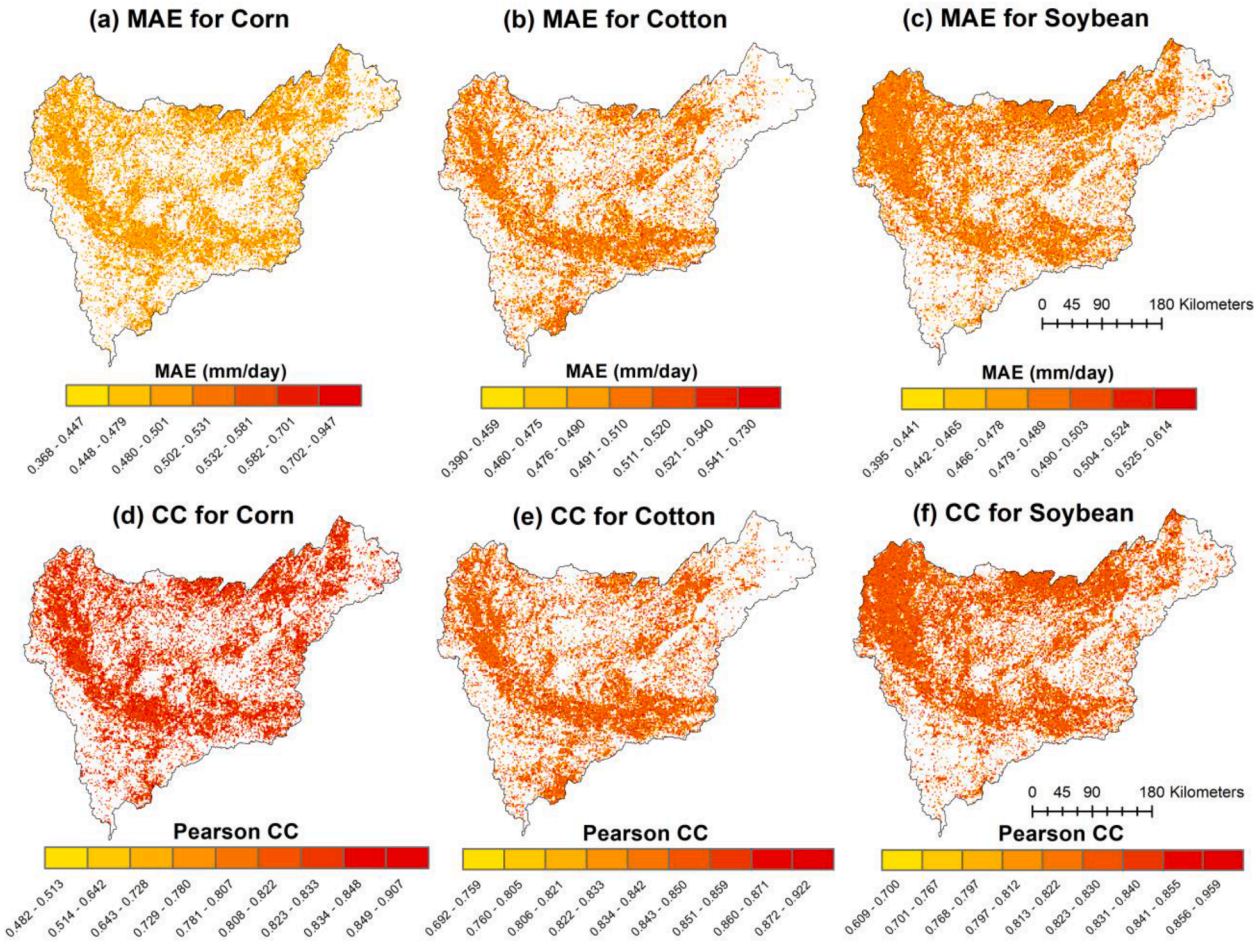

Fig. 6. Spatial variability of *MAE* and Pearson correlation coefficient calculated for the assimilated actual ET (in 2018) for corn, cotton, and soybean at Mobile River basin. Note: each dot represents a farm; CC stands for correlation coefficient.

the observed actual ET which is likely due to the systemic bias associated with model structure. Since this research is a proof-of-concept study, some limitations are acknowledged by the authors, however, it is to be noted that using a robust and superior model such as Penman-Monteith model can likely minimize these biases in results.

For operational irrigation management and planning robust water requirement estimates are crucial, especially during drought periods. For calculating these water requirements, generally reference ET simulation models are employed and converted into the crop water requirement using crop coefficients (Darouich et al., 2021; Pereira et al., 2020; Tang et al., 2018) or employ crop simulation models (Amiri et al., 2022; Moseki et al., 2019). While employing a model in actual ET simulation, performing model calibration is meaningful and useful, since the optimal parameter value(s) can be comparable among different locations/sites. However, often times the calculation of actual ET and crop water requirement at regional scale faces shortfall due to several assumptions including the same cultivar of crops grown throughout the region, lack of high-quality weather data, and model structure associated uncertainties. All of these cascade in the irrigation water requirement simulation leading to significant bias in estimation (Lu et al., 2021). This can create a critical challenge for policymakers/water managers due to the incorrect amount of water diversion to agricultural farms. For such conditions, data assimilation can enhance the estimation of actual ET significantly and thus the crop water requirement.

It is worth mentioning that in this study only the advantage of observed actual ET was employed in the data assimilation. These estimates can be further enhanced by using satellite-based fine resolution soil moisture (Abbaszadeh et al., 2019b; Yin et al., 2020) and LAI (Houborg and McCabe, 2018). Moreover, this study only employed a radiation-based reference ET estimation model (Priestley-Taylor model), which ignores the components of wind speed and the vapor pressure deficit. These variables can have a significant influence in the humid, and tropical regions, and therefore, data assimilation using the PF-VVM on a more robust model such as the Penman-Monteith model can be more suitable. To the best of the authors' knowledge, this is the first attempt to predict the farm-scale actual ET using data assimilation technique in conjunction with the Priestley-Taylor model.

Overall, the data assimilation using the PF-VVM on the modified Priestley-Taylor model is readily transferrable to other locations as well, with some minor changes in the model calibration process, forcing variables, and the observed actual ET data. Also, the cropping season (day of year) including date when crops are planted through harvested, and the elevation data are also required which are location specific and are needed to be altered. Additionally, it is important to highlight that the recently developed open ET dataset (Melton et al., 2021) which has a spatial resolution of 30 m can further improve the actual ET estimates of this study, however, the dataset is only limited to the western US region. While finer scale ET data (~30 m) is not available globally for irrigation and water resources planning and management, our results demonstrate that accounting for uncertainties associated with the forcing data, state variable, and model parameter using data assimilation can result in superior actual ET estimates even with coarser resolutions relative to stand-alone model run which is generally employed for water resources planning and management.

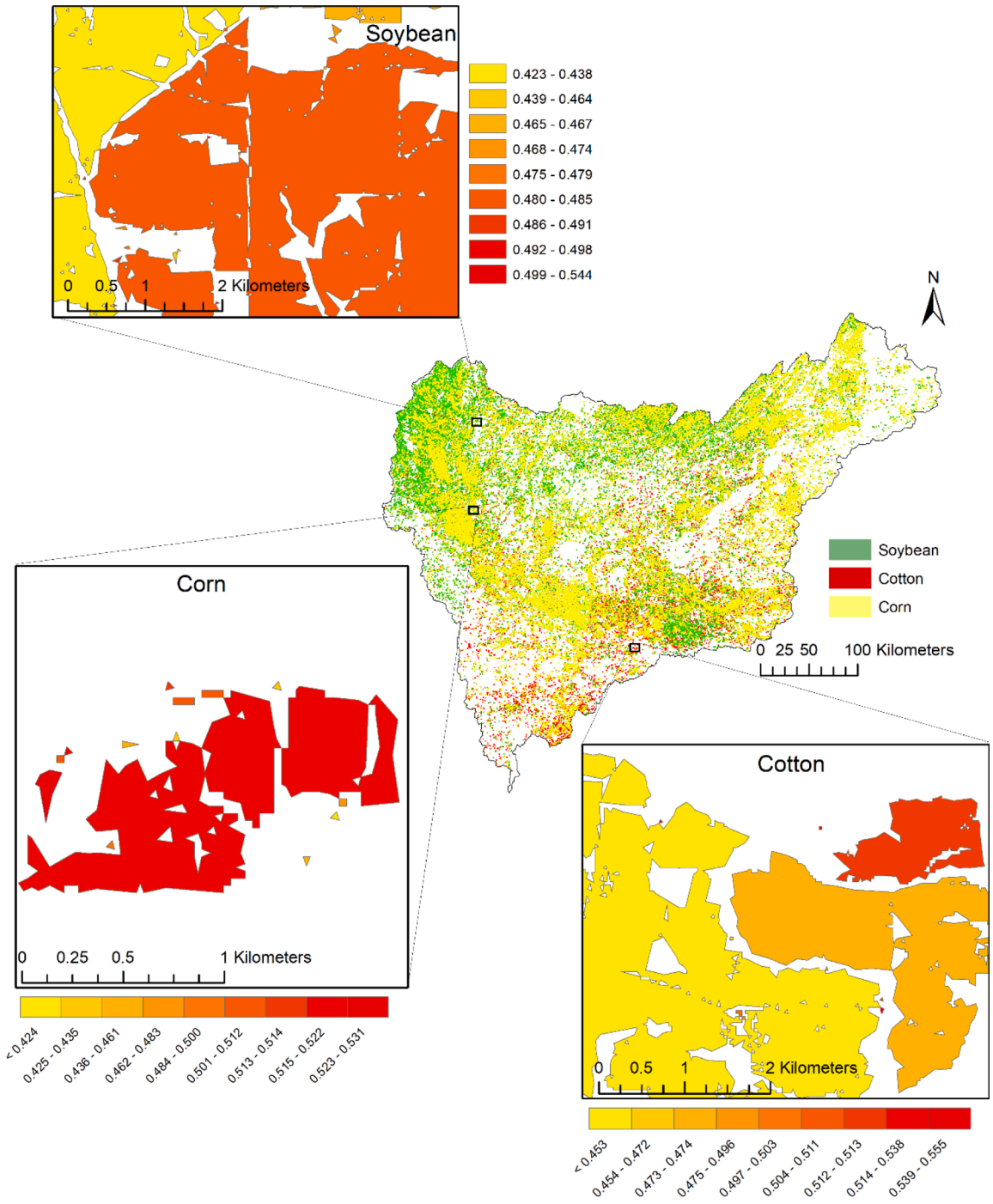

Fig. 7. *MAE* calculated for assimilated actual ET at three randomly selected locations growing corn, cotton, and soybean crops within the study area. Note: these three locations correspond to the same ones reported in Fig. 4. Also, each polygon in the zoomed frames represents a farm.

6. Conclusions

A novel approach of farm-scale actual ET estimation is developed in this study by employing PF-VVM algorithm integrated with the modified Priestley-Taylor model for corn, cotton, and soybean crops at the Mobile River basin in Deep South. The approach uses calibration of the model to identify optimal model parameter, followed by perturbation of the forcing data and state variable to perform the open-loop model run. To implement DA, the observed and model simulated actual ET were used at each time step to calculate the Kalman gain and use it to update the model initialization and generate posterior ET estimates during the assimilation period. The results of the assimilated actual ET, stand-alone model simulated actual ET during calibration and the observed ET were compared at three selected farms and over 320,000 farms within the river basin. Overall, the **major outcomes of the study can be summarized as:**

- **For stand-alone modified Priestley-Taylor model calibration *MAE* ranges from 0.62 to 2.45 mm/day, 0.39 to 2.75 mm/day, and 0.28 to 2.00 mm/day for corn, cotton, and soybean, respectively. Similarly, for the corresponding crops, the Pearson correlation coefficient ranges from 0 to 0.55, 0 to 0.75, and 0 to 0.40, respectively.**

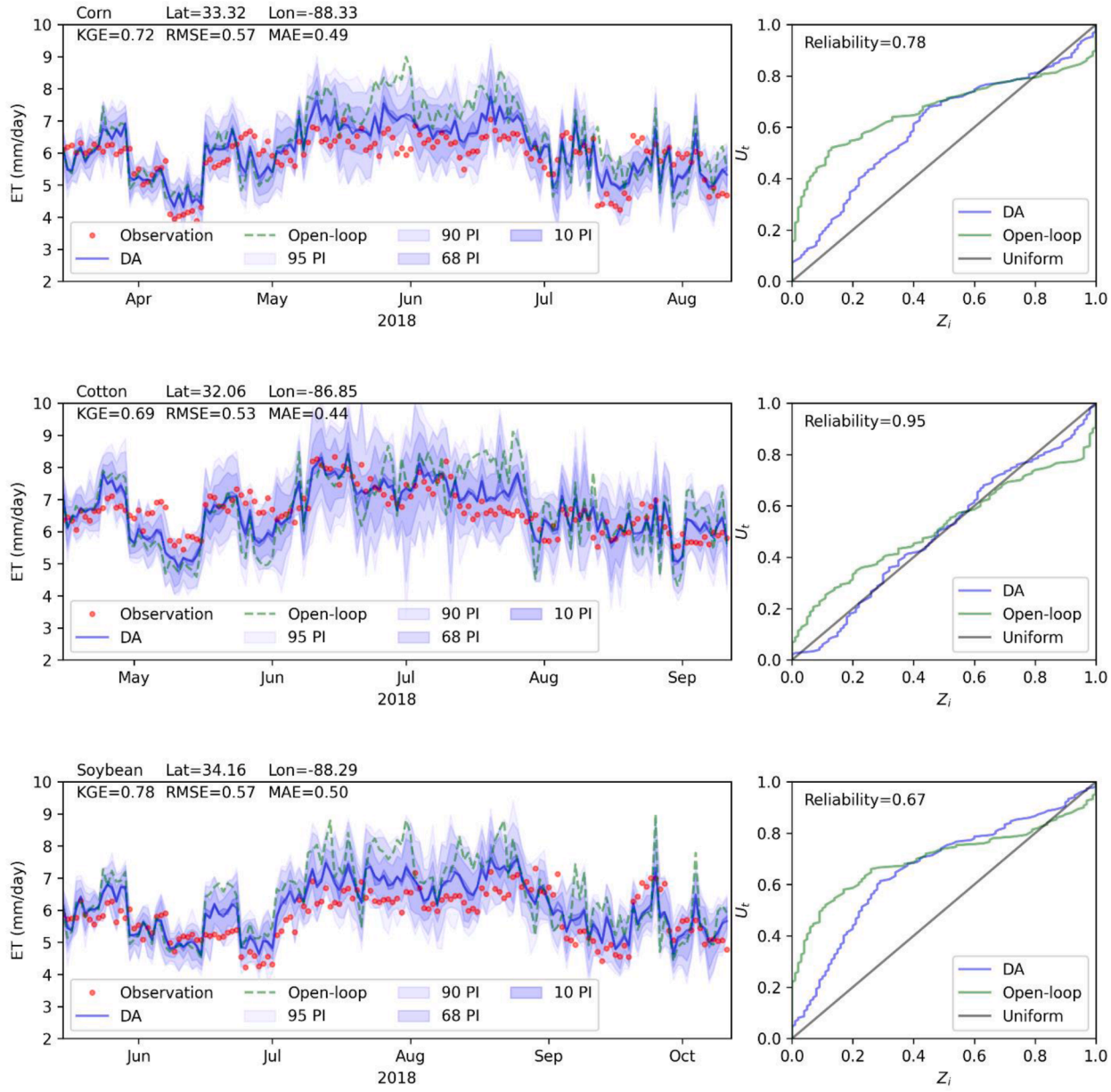

Fig. 8. The left column shows the time series of observed actual ET vs simulated actual ET with and without data assimilation. The right column shows the Q-Q plot of open-loop simulation vs data assimilation. Note: these results correspond to the same locations as pointed in Figs. 4 and 7 for corn, cotton, and soybean crops.

- **For open-loop model simulation, *KGE*, and *RMSE* ranges between 0.46 to 0.47, and 1.17 to 1.28 mm/day, respectively for the three crops across the study area. On the other hand for the assimilated actual ET, *KGE* and *RMSE* ranges from 0.82 to 0.84, and 0.48 to 0.50 mm/day, respectively.**
- **The calculated reliabilities for assimilated actual ET at three selected farms are noted to be 0.78, 0.95, and 0.67 for corn, cotton, and soybean, respectively.**
- **Overall, these results demonstrate that assimilated actual ET is significantly improved over the stand-alone modified Priestley-Taylor model simulated actual ET.**

More efficient and reliable actual ET estimates at farm-scale can enable improved strategies for agricultural interventions including irrigation and as an enhancement in the farm net profitability. Although, there are a number of areas of the development of the proposed approach including sequential assimilation of finer resolution observed ET data, using *in situ* data in the analysis, and employing more sophisticated data assimilation algorithms such as four-dimensional variational data assimilation algorithm and sequential filtering approximations based on three-dimensional variational data assimilation algorithms. Furthermore, the proposed approach can also be employed in forecasting the actual ET at farm-scale which this study did not account for given that this study is a proof-of-concept study, and other studies demonstrated that forecast skills decline with increase with lead-time Abbaszadeh et al., 2018; Hatfield et al., 2018; Stern et al., 2015). In general, this study illustrates farm-scale estimation of actual ET by modified Priestley-Taylor model can result in deceptive actual ET estimates even using a global parameter optimization algorithm. This is decisive especially during droughts when appropriate diversion of irrigation water is crucial and stand-alone model application (which water managers and modelers employ) can lead to inaccurate water allocation. The findings of this study will be beneficial in regional water resource planning and management, especially during drought periods and in data scarce-regions. This is salient given that droughts are expected to be more severe and frequent under projected climate change (Stewart et al.,

2020; Xu et al., 2021b).

CRediT authorship contribution statement

Proloy Deb: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft. **Peyman Abbaszadeh:** Conceptualization, Formal analysis, Methodology, Software, Validation, Visualization, Writing – review & editing. **Hamid Moradkhani:** Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Validation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that there is no competing interest associated with this paper.

Acknowledgements

Financial support was provided by the National Science Foundation, Grant 1856054.

References

Abatzoglou, J.T., 2013. Development of gridded surface meteorological data for ecological applications and modelling. Int. J. Climatol. 33, 121–131. https://doi.org/10.1002/JOC.3413.

Abbaszadeh, P., Gavahi, K., Alipour, A., Deb, P., Moradkhani, H., 2022. Bayesian multi-modeling of deep neural nets for probabilistic crop yield prediction. Agric. For. Meteorol. 314, 108773 https://doi.org/10.1016/J.AGRFORMET.2021.108773.

Abbaszadeh, P., Gavahi, K., Moradkhani, H., 2020. Multivariate remotely sensed and in-situ data assimilation for enhancing community WRF-Hydro model forecasting. Adv. Water Resour. 145 https://doi.org/10.1016/J.ADVWATRES.2020.103721.

Abbaszadeh, P., Moradkhani, H., Daescu, D.N., 2019a. The quest for model uncertainty quantification: a hybrid ensemble and variational data assimilation framework. Water Resour. Res. 55, 2407–2431. https://doi.org/10.1029/2018WR023629.

Abbaszadeh, P., Moradkhani, H., Gavahi, K., Kumar, S., Hain, C., Zhan, X., Duan, Q., Peters-Lidard, C., Karimiziarani, S., 2021. High-resolution SMAP satellite soil moisture product: exploring the opportunities. Bull. Am. Meteorol. Soc. 102, 309–315. https://doi.org/10.1175/bams-d-21-0016.1.

Abbaszadeh, P., Moradkhani, H., Yan, H., 2018. Enhancing hydrologic data assimilation by evolutionary Particle Filter and Markov Chain Monte Carlo. Adv. Water Resour. 111, 192–204. https://doi.org/10.1016/J.ADVWATRES.2017.11.011.

Abbaszadeh, P., Moradkhani, H., Zhan, X., 2019b. Downscaling SMAP radiometer soil moisture over the CONUS using an ensemble learning method. Water Resour. Res. 55, 324–344. https://doi.org/10.1029/2018WR023354.

Akumaga, U., Alderman, P.D., 2019. Comparison of Penman–Monteith and Priestley-Taylor evapotranspiration methods for crop modeling in Oklahoma. Agron. J. 111, 1171–1180. https://doi.org/10.2134/AGRONJ2018.10.0694.

Allen, R., Pereira, L., Raes, D., Smith, M., 1998. Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. Rome, Italy.

Amiri, E., Irmak, S., Araji, H.A., 2022. Assessment of CERES-Maize model in simulating maize growth, yield and soil water content under rainfed, limited and full irrigation. Agric. Water Manag. 259, 107271 https://doi.org/10.1016/J.AGWAT.2021.107271.

Bahrami, A., Goïta, K., Magagi, R., Davison, B., Razavi, S., Elshamy, M., Princz, D., 2021. Data assimilation of satellite-based terrestrial water storage changes into a hydrology land-surface model. J. Hydrol. 597, 125744 https://doi.org/10.1016/J.JHYDROL.2020.125744.

Besag, J., Higdon, D., 1999. Bayesian analysis of agricultural field experiments. J. R. Stat. Soc. Ser. B (Statistical Methodol. 61, 691–746. https://doi.org/10.1111/1467-9868.00201.

Bolten, J.D., Crow, W.T., Jackson, T.J., Zhan, X., Reynolds, C.A., 2010. Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 57–66. https://doi.org/10.1109/JSTARS.2009.2037163.

Boretti, A., Rosa, L., 2019. Reassessing the projections of the World Water Development Report. NPJ Clean Water 2019 21 (2), 1–6. https://doi.org/10.1038/s41545-019-0039-9.

Chen, M., Senay, G.B., Singh, R.K., Verdin, J.P., 2016. Uncertainty analysis of the operational simplified surface energy balance (SSEBop) model at multiple flux tower sites. J. Hydrol. 536, 384–399. https://doi.org/10.1016/J.JHYDROL.2016.02.026.

Crow, W.T., Wood, E.F., 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: a case study based on ESTAR measurements during SGP97. Adv. Water Resour. 26, 137–149. https://doi.org/10.1016/S0309-1708(02)00088-X.

Darouich, H., Karfoul, R., Ramos, T.B., Moustafa, A., Shaheen, B., Pereira, L.S., 2021. Crop water requirements and crop coefficients for jute mallow (Corchorus olitorius L.) using the SIMDualKc model and assessing irrigation strategies for the Syrian Akkar region. Agric. Water Manag. 255, 107038 https://doi.org/10.1016/J.AGWAT.2021.107038.

De Bruin, H.A.R., Keijman, J.Q., 1979. The Priestley-Taylor evaporation model applied to a large shallow lake in the Netherlands. J. Appl. Meteorol. 18, 898–903.

de Wit, A.J.W., van Diepen, C.A., 2007. Crop model data assimilation with the Ensemble Kalman filter for improving regional crop yield forecasts. Agric. For. Meteorol. 146, 38–56. https://doi.org/10.1016/J.AGRFORMET.2007.05.004.

Deb, P., Kiem, A.S., 2020. Evaluation of rainfall–runoff model performance under non-stationary hydroclimatic conditions. Hydrol. Sci. J. 65, 1667–1684. https://doi.org/10.1080/02626667.2020.1754420.

Deb, P., Kiem, A.S., Willgoose, G., 2019. Mechanisms influencing non-stationarity in rainfall-runoff relationships in southeast Australia. J. Hydrol. 571, 749–764. https://doi.org/10.1016/j.jhydrol.2019.02.025.

Deb, P., Moradkhani, H., Han, X., Abbaszadeh, P., Xu, L., 2022. Assessing irrigation mitigating drought impacts on crop yields with an integrated modeling framework. J. Hydrol. 127760 https://doi.org/10.1016/J.JHYDROL.2022.127760.

Deb, P., Shrestha, S., Babel, M.S., 2015. Forecasting climate change impacts and evaluation of adaptation options for maize cropping in the hilly terrain of Himalayas: Sikkim, India. Theor. Appl. Climatol. 121, 649–667. https://doi.org/10.1007/s00704-014-1262-4.

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 https://doi.org/10.1029/2011WR011011.

Drury, B., Valverde-Rebaza, J., Moura, M.F., de Andrade Lopes, A., 2017. A survey of the applications of Bayesian networks in agriculture. Eng. Appl. Artif. Intell. 65, 29–42. https://doi.org/10.1016/J.ENGAPPAI.2017.07.003.

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. https://doi.org/10.1016/0022-1694(94)90057-4.

Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76, 501–521. https://doi.org/10.1007/BF00939380.

Falkenmark, M., 2013. Growing water scarcity in agriculture: future challenge to global water security. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 371 https://doi.org/10.1098/RSTA.2012.0410.

Fisher, D., Allen, R., 1991. Direct load cell-based weighing lysimeter system. In: Allen, R., Howell, T., Pruitt, W. (Eds.), Proceedings of the International Symposium on Lysimetry: Lysimeters for Evapotranspiration and Environmental Measurements. American Society of Civil Engineers, Honolulu, Hawaii.

Gavahi, K., Abbaszadeh, P., Moradkhani, H., 2021. DeepYield: a combined convolutional neural network with long short-term memory for crop yield forecasting. Expert Syst. Appl. 184, 115511 https://doi.org/10.1016/J.ESWA.2021.115511.

Gavahi, K., Abbaszadeh, P., Moradkhani, H., Zhan, X., Hain, C., 2020. Multivariate assimilation of remotely sensed soil moisture and evapotranspiration for drought monitoring. J. Hydrometeorol. 21, 2293–2308. https://doi.org/10.1175/JHM-D-20-0057.1.

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. https://doi.org/10.1016/J.JHYDROL.2009.08.003.

Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1, 96–99. https://doi.org/10.13031/2013.26773.

Harned, D.A., Atkins, J.B., Harvill, J.S., 2004. Nutrient mass balance-accounting for nutrient inputs from atmospheric deposition. J. Am. Water Resour. Assoc. 03070, 765–793.

Hatchett, B.J., Boyle, D.P., Garner, C.B., Kaplan, M.L., Putnam, A.E., Bassett, S.D., 2016. Magnitude and frequency of wet years under a megadrought climate in the western Great Basin, USA. Quat. Sci. Rev. 152, 197–202. https://doi.org/10.1016/J.QUASCIREV.2016.09.017.

Hatfield, S., Subramanian, A., Palmer, T., Düben, P., 2018. Improving weather forecast skill through reduced-precision data assimilation. Mon. Weather Rev. 146, 49–62. https://doi.org/10.1175/MWR-D-17-0132.1.

Houborg, R., McCabe, M.F., 2018. Daily retrieval of NDVI and LAI at 3 m resolution via the fusion of CubeSat, Landsat, and MODIS Data. Remote Sens 10, 890. https://doi.org/10.3390/RS10060890, 2018Page10, 890.

Huang, J., Tian, L., Liang, S., Ma, H., Becker-Reshef, I., Huang, Y., Su, W., Zhang, X., Zhu, D., Wu, W., 2015. Improving winter wheat yield estimation by assimilation of the leaf area index from Landsat TM and MODIS data into the WOFOST model. Agric. For. Meteorol. 204, 106–121. https://doi.org/10.1016/J.AGRFORMET.2015.02.001.

Ines, A.V.M., Das, N.N., Hansen, J.W., Njoku, E.G., 2013. Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. Remote Sens. Environ. 138, 149–164. https://doi.org/10.1016/J.RSE.2013.07.018.

Islam, S., Alam, A.K.M.R., 2021. Performance evaluation of FAO Penman-Monteith and best alternative models for estimating reference evapotranspiration in Bangladesh. Heliyon 7. https://doi.org/10.1016/J.HELIYON.2021.E07487.

Jadidoleslam, N., Mantilla, R., Krajewski, W.F., 2021. Data assimilation of satellite-based soil moisture into a distributed hydrological model for streamflow predictions. Hydrol 8, 52. https://doi.org/10.3390/HYDROLOGY8010052, 2021Page852.

Jafarzadegan, K., Abbaszadeh, P., Moradkhani, H., 2021. Sequential data assimilation for real-time probabilistic flood inundation mapping. Hydrol. Earth Syst. Sci. 25, 4995–5011. https://doi.org/10.5194/HESS-25-4995-2021.

Javadian, M., Behrangi, A., Smith, W.K., Fisher, J.B., 2020. Global trends in evapotranspiration dominated by increases across large cropland regions. Remote Sens. 12, 1221. https://doi.org/10.3390/RS12071221, 2020Page12, 1221.

Johnson, D.M., 2019. Using the Landsat archive to map crop cover history across the United States. Remote Sens. Environ. 232, 111286 https://doi.org/10.1016/J.RSE.2019.111286.
Khaki, M., Hendricks Franssen, H.J., Han, S.C., 2020. Multi-mission satellite remote sensing data for improving land hydrological models via data assimilation. Sci. Rep. 101 (10), 1–23. https://doi.org/10.1038/s41598-020-75710-5, 2020.
Kim, S., Shen, H., Noh, S., Seo, D.J., Welles, E., Pelgrim, E., Weerts, A., Lyons, E., Philips, B., 2021. High-resolution modeling and prediction of urban floods using WRF-Hydro and data assimilation. J. Hydrol. 598, 126236 https://doi.org/10.1016/J.JHYDROL.2021.126236.
Li, Y., Ryu, D., Western, A.W., Wang, Q.J., 2015. Assimilation of stream discharge for flood forecasting: Updating a semidistributed model with an integrated data assimilation scheme. Water Resour. Res. 51, 3238–3258. https://doi.org/10.1002/2014WR016667.
Li, Y., Zhou, Q., Zhou, J., Zhang, G., Chen, C., Wang, J., 2014. Assimilating remote sensing information into a coupled hydrology-crop growth model to estimate regional maize yield in arid regions. Ecol. Modell. 291, 15–27. https://doi.org/10.1016/J.ECOLMODEL.2014.07.013.
Lu, J., Sun, G., McNulty, S.G., Amatya, D.M., 2005. A comparison of six potential evapotranspiration methods for regional use in the southeastern United States. JAWRA J. Am. Water Resour. Assoc. 41, 621–633. https://doi.org/10.1111/J.1752-1688.2005.TB03759.X.
Lu, Y., Chibarabada, T.P., Ziliani, M.G., Onema, J.M.K., McCabe, M.F., Sheffield, J., 2021. Assimilation of soil moisture and canopy cover data improves maize simulation using an under-calibrated crop model. Agric. Water Manag. 252, 106884 https://doi.org/10.1016/J.AGWAT.2021.106884.
Melton, F.S., Huntington, J., Grimm, R., Herring, J., Hall, M., Rollison, D., Erickson, T., Allen, R., Anderson, M., Fisher, J.B., Kilic, A., Senay, G.B., Volk, J., Hain, C., Johnson, L., Ruhoff, A., Blankenau, P., Bromley, M., Carrara, W., Daudert, B., Doherty, C., Dunkerly, C., Friedrichs, M., Guzman, A., Halverson, G., Hansen, J., Harding, J., Kang, Y., Ketchum, D., Minor, B., Morton, C., Ortega-Salazar, S., Ott, T., Ozdogan, M., ReVelle, P.M., Schull, M., Wang, C., Yang, Y., Anderson, R.G., 2021. OpenET: filling a critical data gap in water management for the Western United States. JAWRA J. Am. Water Resour. Assoc. 1–24. https://doi.org/10.1111/1752-1688.12956.
Mitchell, K.E., Lohmann, D., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Cosgrove, B.A., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Lettenmaier, D.P., Marshall, C.H., Entin, J.K., Pan, M., Shi, W., Koren, V., Meng, J., Ramsay, B.H., Bailey, A.A., 2004. 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 https://doi.org/10.1029/2003JD003823.
Mizutani, K., Yamanoi, K., Ikeda, T., Watanabe, T., 1997. Applicability of the eddy correlation method to measure sensible heat transfer to forest under rainfall conditions. Agric. For. Meteorol. 86, 193–203. https://doi.org/10.1016/S0168-1923(97)00012-9.
Mokhtari, A., Noory, H., Vazifedoust, M., 2018. Improving crop yield estimation by assimilating LAI and inputting satellite-based surface incoming solar radiation into SWAP model. Agric. For. Meteorol. 250–251, 159–170. https://doi.org/10.1016/J.AGRFORMET.2017.12.250.
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 https://doi.org/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, 1–17. https://doi.org/10.1029/2004WR003604.
Moradkhani, H., Nearing, G.S., Abbaszadeh, P., Pathiraja, S., 2019. Fundamentals of data assimilation and theoretical advances. In: Duan, Q., Pappenberger, F., Wood, A., Cloke, H.L., Schaake, J.C. (Eds.), Handbook of Hydrometeorological Ensemble Forecasting. Springer, Berlin, Heidelberg, Berlin, Germany, pp. 675–699. https://doi.org/10.1007/978-3-642-39925-1_30.
Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50, 885–900. https://doi.org/10.13031/2013.23153.
Moriasi, D.N., Gitau, M.W., Pai, N., Daggupati, P., 2015. Hydrologic and water quality models: performance measures and evaluation criteria. Trans. ASABE 58, 1763–1785. https://doi.org/10.13031/TRANS.58.10715.
Morrison, K.T., Nelson, T.A., Nathoo, F.S., Ostry, A.S., 2012. Application of Bayesian spatial smoothing models to assess agricultural self-sufficiency, 26, pp. 1213–1229. https://doi.org/10.1080/13658816.2011.633491 https://doi.org/10.1080/13658816.2011.633491.
Moseki, O., Murray-Hudson, M., Kashe, K., 2019. Crop water and irrigation requirements of Jatropha curcas L. in semi-arid conditions of Botswana: applying the CROPWAT model. Agric. Water Manag. 225, 105754 https://doi.org/10.1016/J.AGWAT.2019.105754.
Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. Comput. J. 7, 308–313.
Norfleet, M.L., Reeves, D.W., Burmester, C.H., Mina, B., Monks, C.D., 1997. Optimal plantingdates for cotton in the Tennessee valley of North Alabama. In: Proceedings of the Beltwide Cotton Conference. National Cotton Council, Memphis, TN, pp. 644–647.
Pandey, S., 2019. The Role of Irrigation for Food Security and Sustainability. In: Ferranti, P., Berry, E.M., Anderson, J.R. (Eds.), Encyclopedia of Food Security and Sustainability. Elsevier, Amsterdam, pp. 142–146. https://doi.org/10.1016/B978-0-08-100596-5.22173-2.
Pathiraja, S., Moradkhani, H., Marshall, L., Sharma, A., Geenens, G., 2018. Data-driven model uncertainty estimation in hydrologic data assimilation. Water Resour. Res. 54, 1252–1280. https://doi.org/10.1002/2018WR022627.
Pearson, K., 1895. Notes on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London. Royal Society of London, London, UK, pp. 240–242.
Pereira, L.S., Paredes, P., Jovanovic, N., 2020. Soil water balance models for determining crop water and irrigation requirements and irrigation scheduling focusing on the FAO56 method and the dual Kc approach. Agric. Water Manag. 241, 106357 https://doi.org/10.1016/J.AGWAT.2020.106357.
Priestley, C.H., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large scale parameters. Mon. Weather Rev. 100, 81–92.
Purdy, A.J., Fisher, J.B., Goulden, M.L., Famiglietti, J.S., 2016. Ground heat flux: an analytical review of 6 models evaluated at 88 sites and globally. J. Geophys. Res. Biogeosci. 121, 3045–3059. https://doi.org/10.1002/2016JG003591.
Puy, A., Borgonovo, E., Lo Piano, S., Levin, S.A., Saltelli, A., 2021. Irrigated areas drive irrigation water withdrawals. Nat. Commun. 2021 121 (12), 1–12. https://doi.org/10.1038/s41467-021-24508-8.
Puy, A., Lo Piano, S., Saltelli, A., 2020. Current models underestimate future irrigated areas. Geophys. Res. Lett. 47 https://doi.org/10.1029/2020GL087360.
Reichle, R.H., McLaughlin, D.B., Entekhabi, D., 2002. Hydrologic data assimilation with the ensemble kalman filter. Mon. Weather Rev. 130, 103–114. https://doi.org/10.1175/1520-0493(2002)130.
Renard, B., Kavetski, D., Kuczera, G., Thyer, M., Franks, S.W., 2010. Understanding predictive uncertainty in hydrologic modeling: the challenge of identifying input and structural errors. Water Resour. Res. 46, 5521. https://doi.org/10.1029/2009WR008328.
Schumacher, M., Forootan, E., van Dijk, A.I.J.M., Müller Schmied, H., Crosbie, R.S., Kusche, J., Döll, P., 2018. Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. Remote Sens. Environ. 204, 212–228. https://doi.org/10.1016/J.RSE.2017.10.029.
Senay, G.B., Bohms, S., Singh, R.K., Gowda, P.H., Velpuri, N.M., Alemu, H., Verdin, J.P., 2013. Operational evapotranspiration mapping using remote sensing and weather datasets: a new parameterization for the SSEB approach. JAWRA J. Am. Water Resour. Assoc. 49, 577–591. https://doi.org/10.1111/JAWR.12057.
Seo, E., Lee, M.I., Reichle, R.H., 2021. Assimilation of SMAP and ASCAT soil moisture retrievals into the JULES land surface model using the Local Ensemble Transform Kalman Filter. Remote Sens. Environ. 253, 112222 https://doi.org/10.1016/J.RSE.2020.112222.
Shirley, R., Pope, E., Bartlett, M., Oliver, S., Quadrianto, N., Hurley, P., Duivenvoorden, S., Rooney, P., Barrett, A.B., Kent, C., Bacon, J., 2020. An empirical, Bayesian approach to modelling crop yield: Maize in USA. Environ. Res. Commun. 2, 025002 https://doi.org/10.1088/2515-7620/AB67F0.
Siebert, S., Döll, P., Hoogeveen, J., Faures, J.-M., Frenken, K., Feick, S., 2005. Development and validation of the global map of irrigation areas. Hydrol. Earth Syst. Sci. 9, 535–547.
Silvestro, P.C., Pignatti, S., Pascucci, S., Yang, H., Li, Z., Yang, G., Huang, W., Casa, R., 2017. Estimating wheat yield in china at the field and district scale from the assimilation of satellite data into the aquacrop and simple algorithm for yield (SAFY) models. Remote Sens. 9, 509. https://doi.org/10.3390/RS9050509, 2017Page9, 509.
Stern, Harvey, Davidson, N.E., Stern, H, Swanston, C., Streets, E., 2015. Trends in the skill of weather prediction at lead times of 1–14 days. Q. J. R. Meteorol. Soc. 141, 2726–2736. https://doi.org/10.1002/QJ.2559.
Stewart, I.T., Rogers, J., Graham, A., 2020. Water security under severe drought and climate change: disparate impacts of the recent severe drought on environmental flows and water supplies in Central California. J. Hydrol. X 7, 100054. https://doi.org/10.1016/J.HYDROA.2020.100054.
Stewart, R.B., Rouse, W.R., 1977. Substantiation of the Priestley and Taylor parameter α = 1.26 for potential evaporation in high latitudes in. J. Appl. Meteorol. Climatol. 16 (6), 649–650. Issue1977J. Appl. Meteorol. Climatol. 16.
Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.L., 2018. A review of global precipitation data sets: data sources, estimation, and intercomparisons. Rev. Geophys. 56, 79–107. https://doi.org/10.1002/2017RG000574.
Tang, Q., Feng, G., Fisher, D., Zhang, H., Ouyang, Y., Adeli, A., Jenkins, J., 2018. Rain water deficit and irrigation demand of major crops in the Mississippi delta. Trans. ASABE 61, 927–935. https://doi.org/10.13031/trans.12397.
Tian, S., Renzullo, L.J., Pipunic, R.C., Lerat, J., Sharples, W., Donnelly, C., 2021. Satellite soil moisture data assimilation for improved operational continental water balance prediction. Hydrol. Earth Syst. Sci. 25, 4567–4584. https://doi.org/10.5194/HESS-25-4567-2021.
Trenberth, K.E., Fasullo, J.T., Kiehl, J., 2009. Climate and earth's energy budget. Bull. Am. Meteorol. Soc. 90, 311–323. https://doi.org/10.1175/2008BAMS2634.1.
Trenberth, K.E., Smith, L., Qian, T., Dai, A., Fasullo, J., 2007. Estimates of the global water budget and its annual cycle using observational and model data. J. Hydrometeorol. 8, 758–769. https://doi.org/10.1175/JHM600.1.
USDA, N., 2020. Farms and Land in Farms 2019 Summary. Washington DC.
Utset, A., Farré, I., Martínez-Cob, A., Cavero, J., 2004. Comparing Penman–Monteith and Priestley–Taylor approaches as reference-evapotranspiration inputs for modeling maize water-use under Mediterranean conditions. Agric. Water Manag. 66, 205–219. https://doi.org/10.1016/J.AGWAT.2003.12.003.
Vetra-Carvalho, S., van Leeuwen, P.J., Nerger, L., Barth, A., Altaf, M.U., Brasseur, P., Kirchgessner, P., Beckers, J.M., 2018. State-of-the-art stochastic data assimilation

methods for high-dimensional non-Gaussian problems. Tellus A Dyn. Meteorol. Oceanogr. 70, 1–43. https://doi.org/10.1080/16000870.2018.1445364.

Warner, K.A., Bonzongo, J.C.J., Roden, E.E., Ward, G.M., Green, A.C., Chaubey, I., Lyons, W.B., Arrington, D.A., 2005. Effect of watershed parameters on mercury distribution in different environmental compartments in the Mobile Alabama River Basin. USA. Sci. Total Environ. 347, 187–207. https://doi.org/10.1016/j.scitotenv.2004.12.011.

Xu, L., Abbaszadeh, P., Moradkhani, H., Chen, N., Zhang, X., 2020. Continental drought monitoring using satellite soil moisture, data assimilation and an integrated drought index. Remote Sens. Environ. 250, 112028 https://doi.org/10.1016/j.rse.2020.112028.

Xu, L., Chen, N., Chen, Z., Zhang, C., Yu, H., 2021a. Spatiotemporal forecasting in earth system science: methods, uncertainties, predictability and future directions. Earth-Sci. Rev. 222, 103828 https://doi.org/10.1016/J.EARSCIREV.2021.103828.

Xu, L., Chen, N., Yang, C., Zhang, C., Yu, H., 2021b. A parametric multivariate drought index for drought monitoring and assessment under climate change. Agric. For. Meteorol. 310, 108657 https://doi.org/10.1016/J.AGRFORMET.2021.108657.

Xu, L., Chen, N., Zhang, X., 2019. Global drought trends under 1.5 and 2 °C warming. Int. J. Climatol. 39, 2375–2385. https://doi.org/10.1002/JOC.5958.

Yadav, S., Deb, P., Kumar, S., Pandey, V., Pandey, P.K., 2016. Trends in major and minor meteorological variables and their influence on reference evapotranspiration for mid Himalayan region at east Sikkim, India. J. Mt. Sci. 13 https://doi.org/10.1007/s11629-014-3238-3.

Yang, Z., Zhang, Q., Yang, Y., Hao, X., Zhang, H., 2016. Evaluation of evapotranspiration models over semi-arid and semi-humid areas of China. Hydrol. Process. 30, 4292–4313. https://doi.org/10.1002/HYP.10824.

Yin, J., Zhan, X., Liu, J., Moradkhani, H., Fang, L., Walker, J.P., 2020. Near-real-time one-kilometre soil moisture active passive soil moisture data product. Hydrol. Process. 34, 4083–4096. https://doi.org/10.1002/HYP.13857.

Zhang, C., Abbaszadeh, P., Xu, L., Moradkhani, H., Duan, Q., Gong, W., 2021. A combined optimization-assimilation framework to enhance the predictive skill of community land model. Water Resour. Res. 57, e2021WR029879 https://doi.org/10.1029/2021WR029879.

Zhuo, W., Huang, J., Li, L., Zhang, X., Ma, H., Gao, X., Huang, H., Xu, B., Xiao, X., 2019. Assimilating Soil Moisture Retrieved from Sentinel-1 and Sentinel-2 Data into WOFOST model to improve winter wheat yield estimation. Remote Sens. 11, 1618. https://doi.org/10.3390/RS11131618, 2019Page11, 1618.

Ziliani, M.G., Altaf, M.U., Aragon, B., Houborg, R., Franz, T.E., Lu, Y., Sheffield, J., Hoteit, I., McCabe, M.F., 2022. Early season prediction of within-field crop yield variability by assimilating CubeSat data into a crop model. Agric. For. Meteorol. 313, 108736 https://doi.org/10.1016/J.AGRFORMET.2021.108736.