Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Simultaneous Parameter and State Estimation of Agro-Hydrological Systems

Song Bo * **Soumya R. Sahoo** * **Xunyuan Yin** * **Jinfeng Liu** *[,1]
Sirish L. Shah *

* *Department of Chemical & Materials Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada.*

Abstract: The Richards equation plays an important role in the study of agro-hydrological systems. It models the water movement in soil in the vadose zone, which is driven by capillary and gravitational forces. Its states (capillary potential) and parameters (hydraulic conductivity, saturated and residual soil moistures and van Genuchten-Mualem parameters) are essential for the accuracy of mathematical modeling, yet difficult to obtain experimentally. In this work, an estimation approach is developed to estimate the parameters and states of the Richards equation simultaneously. Parameter identifiability and sensitivity analysis are used to determine the most important parameters for estimation purpose. Three common estimation schemes (extended Kalman filter, ensemble Kalman filter and moving horizon estimation) are investigated. The estimation performance is compared and analyzed based on extensive simulations.

Keywords: state estimation; parameter estimation; moving horizon estimation; extended kalman filter; ensemble kalman filter; richards equation

1. INTRODUCTION

Water scarcity is becoming serious issue worldwide due to population growth and climate change. According to United Nations statistics (FAO, 2016), approximately 70% of available fresh water is consumed for agricultural activities, with the main consumer being irrigation. Currently, it is still a common practice to use open-loop irrigation, which leads to low average water-use efficiency. Closed-loop irrigation is a promising alternative to reduce water consumption (Mao et al., 2018). In the development of such a closed-loop irrigation system, it is important to have the soil moisture information of the entire field, which is in general very difficult to obtain. One way to overcome this challenge is to estimate the field's soil moisture based on limited sensor measurements. However, this depends on the accuracy of the agro-hydrological model. We aim to develop a systematic parameter and state estimation scheme that can provide accurate estimates of soil moisture.

We consider simultaneous parameter and state estimation based on agro-hydrological systems modeled using the Richards equation. The parameters of Richards equation are related to soil properties. Different approaches have been developed to estimate their values. They were estimated in a soil lab by fitting the soil-water retention curve and hydraulic conductivity curve using collected field data of soil moisture, hydraulic conductivity, and corresponding capillary pressure head (van Genuchten, 1980). However, soil properties may change over time and it would be expensive to take frequent soil samples for lab analysis especially when a big field is considered. As an alternative to direct lab analysis, soil parameters can be estimated indirectly based on the Richards equation and some easily-accessible field measurements by minimizing the difference between measured values and model predicted values. This type of indirect approaches are referred to as inverse estimation (Hwang and Powers, 2003). The above methods can only estimate soil parameters but not soil moisture, and cannot be used for online parameter estimation.

Sequential data assimilation is a widely used approach in estimating soil parameters online. It has the ability to deal with uncertainties in the measurements and the model. Extended Kalman filters (EKF) (Lv et al., 2011), and ensemble Kalman filters (EnKF) (Li and Ren, 2011) are common and widely used algorithms in sequential data assimilation for soil parameter estimation. However, these methods cannot handle constraints on the states or parameters. Constraints on the states and parameters are important information and may be used to significantly improve estimation performance as will be demonstrated in the simulations of this work. To address this issue, the optimization based moving horizon estimation (MHE) method is considered, which is widely used in state estimation of nonlinear systems with explicit constraints taken into account (Rao et al., 2003).

In this work, the investigated system and the formulation of the mathematical model are introduced in Section 2. The estimation methods, MHE, EKF, and EnKF for the augmented system are introduced in Section 3. Section 4 includes the methods of identifiability and sensitivity, used to study the significance of parameters. Section 5 shows the synthetic experimental setup, determination of significant parameters, and comparison of estimation results, followed by concluding remarks in Section 6.

[1] Financial support from the Alberta Innovates and the National Science and Engineering Research Council of Canada is gratefully acknowledged. Corresponding author: Jinfeng Liu. Tel.: +1 780 492 1317. Fax: +1 780 492 2881. Email: jinfeng@ualberta.ca.

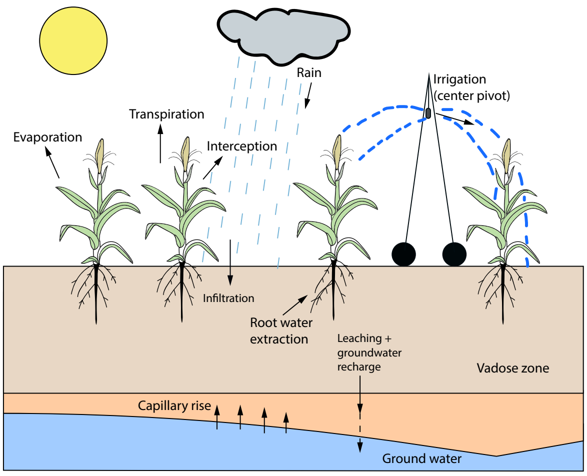

Fig. 1. A schematic diagram of an agro-hydrological system

2. SYSTEM DESCRIPTION AND PROBLEM FORMULATION

An agro-hydrological system shown in Figure 1 (Nahar, 2019) describes the water movements between soil, crop, and atmosphere. In this work, we focus on soil that is above the water table (i.e., soil in the vadose zone), where water movement is mainly driven by capillary and gravitational forces. The water dynamics is modeled using Richards equation shown below (Richards, 1931):

$$c(h)\frac{\partial h}{\partial t} = \frac{\partial}{\partial z}\left[K(h)\left(\frac{\partial h}{\partial z}+1\right)\right] \tag{1}$$

where h (m) is the capillary potential in the unsaturated soil. The value 1 on the right-hand-side denotes the impact of gravitational force on water in the vertical (z) direction. $K(h)$ (m/s) and $c(h)$ $(1/m)$ denote hydraulic conductivity and capillary capacity of the soil, respectively, as follows (van Genuchten, 1980):

$$\begin{aligned} K(h) = {} & K_s\left[(1+(-\alpha h)^n)^{-\left(1-\frac{1}{n}\right)}\right]^{\frac{1}{2}} \\ & \cdot\left[1-\left[1-\left[(1+(-\alpha h)^n)^{-\left(1-\frac{1}{n}\right)}\right]^{\frac{n}{n-1}}\right]^{1-\frac{1}{n}}\right]^2 \end{aligned} \tag{2}$$

$$c(h) = (\theta_s-\theta_r)\,\alpha n\left(1-\frac{1}{n}\right)(-\alpha h)^{n-1}\left[1+(-\alpha h)^n\right]^{-\left(2-\frac{1}{n}\right)} \tag{3}$$

where K_s (m/s), θ_s (m^3/m^3), and θ_r (m^3/m^3) are saturated hydraulic conductivity, saturated and residual soil moisture, respectively. The van Genuchten-Mualem parameters α $(1/m)$ and n characterize the properties of the soil, which are proportional to the inverse of the soil air entry pressure and of soil porosity, respectively. The soil-water retention curve built by van Genuchten (1980) is shown below:

$$\theta(h) = (\theta_s-\theta_r)\left[\frac{1}{1+(-\alpha h)^n}\right]^{1-\frac{1}{n}} + \theta_r \tag{4}$$

In (4), θ (m^3/m^3) denotes volumetric water content in soil. The five parameters θ_s, θ_r, α, n, and K_s determine the properties of a type of soil. The sequential estimation of soil properties is studied based on real-time field measurements: capillary potential h. We consider that the soil properties are spatially and temporally homogeneous.

2.1 Finite difference model development

By applying two-point forward difference scheme and two-point central difference scheme to approximate the derivatives with respect to the temporal and spatial variables, respectively, the discrete-time finite difference model at node i and time instant $k+1$ can be obtained as follows:

$$\begin{aligned} h_i(k+1) = h_i(k) + {} & \frac{\Delta t}{\frac{1}{2}c_i(h(k))(\Delta z_{i-1}+\Delta z_i)} \\ & \cdot\left[K_{i-\frac{1}{2}}(h(k))\left(\frac{h_{i-1}(k)-h_i(k)}{\Delta z_{i-1}}+1\right)\right. \\ & \left. - K_{i+\frac{1}{2}}(h(k))\left(\frac{h_i(k)-h_{i+1}(k)}{\Delta z_i}+1\right)\right] \end{aligned} \tag{5}$$

where $k \in [0, N_t] \subset \mathbb{Z}$ and $i \in [1, N_x] \subset \mathbb{Z}$, representing time and position indexes, respectively. N_t and N_x are the total number of time instants and states investigated. Δt and Δz represent the discretization step sizes in the temporal and spatial domains. The hydraulic conductivity, for example, $K_{i-\frac{1}{2}}$, is linearized explicitly as $K_{i-\frac{1}{2}}(h) = K(\frac{h_{i-1}+h_i}{2})$ and $c_i(h(k))$ is defined as $c(h_i(k))$.

The Neumann boundary condition is utilized to characterize the top and bottom boundaries of the system and are shown below, respectively:

$$\left.\frac{\partial h(k)}{\partial z}\right|_T = -1 - \frac{q_T(k)}{K(h(k))} \tag{6}$$

$$\left.\frac{\partial (h(k)+z)}{\partial z}\right|_B = 1 \tag{7}$$

where the subscripts T and B represent the top and bottom boundary conditions, respectively. The q_T (m/s) is the irrigation rate which is considered as the input of the system and free drainage boundary condition is applied at the bottom.

For the sake of simplicity, the compact form of the model is obtained by combining N_x of (5) for all spatial nodes and the boundary conditions, (6) and (7). It is shown below:

$$x(k+1) = F(x(k), u(k), p(k)) + \omega_x(k) \tag{8}$$

where $x(k) \in \mathbb{X} \subset \mathbb{R}^{N_x}$ represents the state vector containing N_x capillary pressure values for corresponding spatial nodes, at the defined time instant k. $p(k) \in \mathbb{P} \subset \mathbb{R}^{N_p}$, represents the parameter vector containing the parameters to be estimated. $u(k) \in \mathbb{U} \subset \mathbb{R}^{N_u}$, $\omega_x(k) \in \mathbb{W}_x \subset \mathbb{R}^{N_{\omega_x}}$ denote the input and the model disturbances, respectively.

The general output function, with the measurement noise taken into account, is shown below:

$$y(k) = G(x(k), p(k)) + \nu(k) \tag{9}$$

where $y(k) \in \mathbb{Y} \subset \mathbb{R}^{N_y}$ and $\nu(k) \in \mathbb{V} \subset \mathbb{R}^{N_\nu}$ denote the measurement vector and measurement noise. The tensiometers are used to measure the water potential h in the soil, $G(\cdot)$ in (9) represents a matrix indicating which states are measured by the tensiometers.

Furthermore, in order to estimate the states and parameters simultaneously, the parameter vector is augmented at the end of the state vector and treated as a part of the augmented state vector, $X = [x, p]^T$. An estimation of the augmented state vector X brings the benefit to estimate the states and parameters at the same time. The augmented model can be constructed by augmenting (8) with the following equation:

$$p(k+1) = p(k) + \omega_p(k) \tag{10}$$

where $\omega_p(k) \in \mathbb{W}_p \subset \mathbb{R}^{N_{\omega_p}}$. When the parameter vector p is assumed to be constant, ω_p is equal to 0.

The augmented model and output function used for simultaneous parameter and state estimation are shown below:

$$X(k+1) = F_a(X(k), u(k)) + \omega_a(k) \tag{11a}$$
$$y(k) = G_a(X(k)) + \nu(k) \tag{11b}$$

where $X(k) \in \mathbb{X}_a \subset \mathbb{R}^{N_x+N_p}$, $\omega_a(k) \in \mathbb{W}_a \subset \mathbb{R}^{N_w+N_p}$, and the subscript a of $F(\cdot)$ and $G(\cdot)$ denotes the augmentation.

3. ESTIMATION METHODS

In this work, three common estimation schemes, MHE, EKF, and EnKF are applied to the augmented model to estimate the states and parameters. The design of these methods are detailed next.

3.1 Moving horizon estimation

MHE is an online optimization based estimation method (Rao et al., 2003). The MHE optimization problem used in this work is formulated as follows:

$$\min_{\substack{\hat{X}(k-N),\cdots,\hat{X}(k),\\ \hat{\omega}_a(k-N),\cdots,\hat{\omega}_a(k-1)}} V(\hat{X}(k-N)) + \sum_{j=k-N}^{k-1} \|\hat{\omega}_a(j)\|^2_{Q^{-1}} + \sum_{j=k-N}^{k} \|\hat{v}(j)\|^2_{R^{-1}} \tag{12a}$$

$$\text{s.t. } \hat{X}(j+1) = F_a(\hat{X}(j), u(j)) + \hat{\omega}_a(j), \quad j \in [k-N, k-1] \subset \mathbb{Z} \tag{12b}$$
$$y(j) = G_a(\hat{X}(j)) + \hat{\nu}(j), \quad j \in [k-N, k] \subset \mathbb{Z} \tag{12c}$$
$$V(k-N) = \left\| \hat{X}(k-N) - \bar{X}(k-N) \right\|^2_{P^{-1}} \tag{12d}$$
$$\bar{X}(k-N) = \hat{X}(k-N|k-N) \tag{12e}$$
$$\hat{X}(j) \in \mathbb{X}_a, \ \hat{v}(j) \in \mathbb{V}, \quad j \in [k-N, k] \subset \mathbb{Z} \tag{12f}$$
$$\hat{\omega}_a(j) \in \mathbb{W}_a, \quad j \in [k-N, k-1] \subset \mathbb{Z} \tag{12g}$$

In the MHE optimization, the objective is to minimize the distance between the predicted and observed measurements which is measured by the term $\|\hat{\nu}\|^2_{R^{-1}}$ as shown in (12a). The caret sign ˆ indicates that the variable is estimated. In addition, the model uncertainty or the process disturbance is taken into account and represented by $\|\hat{\omega}_a\|^2_{Q^{-1}}$. The arrival cost, V summarizes the information from the initial state of the model up to the beginning of the estimation window of the MHE. N denotes the length of the estimation window. After each optimization, only the last estimated state within the estimation window is used. $\hat{X}$ and $\hat{\omega}_a$ within the moving window are the decision variables of the optimization problem. The terms $\hat{\omega}_a$ and $\hat{\nu}$ obey the process constraints of (12b) and (12c), respectively. And $\bar{X}$ follows the definition of (12e). $\hat{X}(k-N|k-N)$ represents the estimated state $\hat{X}$ at time instant $k-N$, which is estimated at time instant $k-N$. Matrices P, Q, R are positive definite matrices and they are the covariance matrices of state uncertainty, process noise ω_a, and measurement noise ν, respectively. In addition, MHE takes into account constraints on the states, parameters, and model uncertainties as expressed in (12f) and (12g).

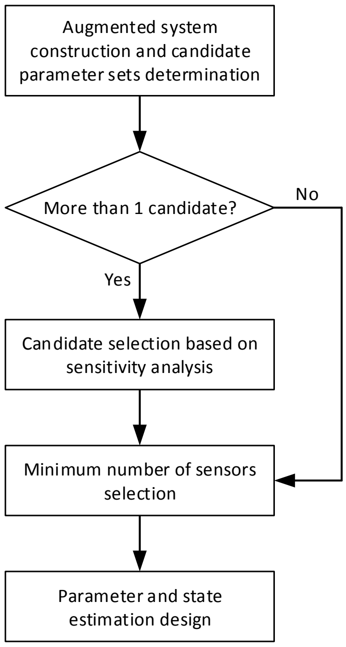

Fig. 2. A flowchart of the procedure to determine the significant parameters and number of sensors

3.2 Extended Kalman filter and ensemble Kalman filter

EKF is a common method used for state estimation of nonlinear systems based on successive linearization of the nonlinear system. It involves a prediction step and an update step. The algorithm can be found in Lv et al. (2011).

The EnKF is a method developed by Evensen (1994) based on Monte Carlo method. An ensemble of trajectories of the system is generated based on the priori probability distribution of the case. A practical implementation scheme is introduced by Gillijns et al. (2006), which estimated the probability distribution based on the information embedded within ensembles, instead of propagation of the state covariance matrix P. This scheme is used in this work.

4. PROPOSED PROCEDURE TO DETERMINE SIGNIFICANT PARAMETERS AND NUMBER OF SENSORS

In reality, it is nearly impossible to measure all states, and the parameters can not be determined easily. First, according to Sahoo et al. (2019), it states that the original system of (8) is observable using limited number of measurements. However, for this work the augmented system (11) is studied, for this case it is necessary to ensure that the parameters are also identifiable. The proposed procedure to check the identifiability of the parameters, to select appropriate parameters for estimation and to determine the minimum number of sensors is shown in Figure 2. The key steps are explained below.

4.1 Determine candidate parameter sets for estimation

After augmenting the original nonlinear system with the parameters, the entire system may not be observable. In order to determine which parameters can and should be estimated online, we resort to observability analysis. In this step, we assume that all states are measured. This ensures that the observability analysis results depend only on the parameters. If the augmented system is not observable, then the unobservability is caused by the augmentation of the parameters in the state vector.

When checking the observability of the augmented system, the system is started with all the parameters augmented. If the augmented system is not observable, then one of the parameters is removed from the augmented system. If there are N_p parameters, then there are N_p different ways to remove the one parameter. All these N_p cases are considered. If after removing one parameter and upon finding that the new augmented system is observable, we continue to the next step to determine which parameter set to estimate (described in the next subsection). If we can still not find an observable augmented system after removing one parameter, we continue to remove two parameters from the original augmented system. Again, all the possible cases should be considered. If we can still not find an observable system, we continue to remove three parameters from the original augmented system. This continues until we find at least a system that is observable.

When checking the observability, we propose to use the Popov-Belevitch-Hautus (PBH) observability theory. Since the augmented system is a nonlinear system, it should be linearized before PBH can be applied. Instead of linearizing the system at one point, it can be linearized at different points along typical operating trajectories as used in Nahar et al. (2019).

Note that the observability analysis described in this step may generate more than one candidate parameter sets that can be estimated through augmentation of the original agro-hydrological system.

4.2 Sensitivity analysis

If there is only one candidate parameter set from the previous step, we can continue with the candidate and move to the next subsection to find the minimum number of sensors. However, if there are more than one candidates, we need to determine which parameter set to choose. Sensitivity analysis is proposed to determine the importance of these parameters and pick the set containing the most important parameters for further analysis.

The sensitivity analysis measures how the outputs respond when there is a change in one parameter. The sensitivity matrix $S_y(k)$ shown in Stigter et al. (2017) contains the information about, at time instant k, how each output is affected by $X(0)$ which is constituted of the initial state $x(0)$ and the parameters p.

The detailed steps to derive the sensitivity matrix is inspired by Stigter et al. (2017). Once the sensitivity matrix is obtained, the relative importance of different parameters can be determined. Specifically, we can exam the magnitudes of the elements in the sensitivity matrix. Each parameter corresponds to one column in the sensitivity matrix. We can use, for example, the summation of the absolute values of the elements of each column to compare the relative importance of parameters. A bigger value implies a more important parameter in terms of its impact on the outputs. Among all the candidate parameter sets, we keep the parameter set with the highest sensitivity values.

4.3 Minimum number of sensors

After the parameter set to be estimated is determined, the original system is augmented with the parameters, as illustrated in Sahoo et al. (2019), we can use the maximum multiplicity theory (Yuan et al., 2013) to determine the minimum number of sensors required to ensure the observability of the entire system. Then, state estimation techniques can be used to estimate the states and parameters simultaneously.

5. SIMULATION RESULTS AND DISCUSSION

5.1 System description

In this work, a total length (L) of 67 cm loam soil column is investigated. It is equally partitioned into 32 compartments. Correspondingly, Richards equation is spatially discretized into 32 states (N_x) in the z-direction, with each state centered at the corresponding compartment. At the surface of the soil, the irrigation, q_T, is performed at the rate of 2.50 cm/day, from 12:00 PM to 4:00 PM daily. At the bottom, the free drainage boundary condition is considered. The soil column has the homogeneous initial condition ($x(0)$) of -0.514 m capillary pressure head and the parameters of the soil are shown in Table 1 (Carsel and Parrish, 1988).

5.2 Determination of significant parameters and number of sensors

The augmented system (11) is utilized to achieve simultaneous parameter and state estimation. First without knowing the observability of the augmented system, all 5 parameters (K_s, θ_s, θ_r, α, and n) are augmented; that is, $N_p = 5$. All 32 states are assumed to be measured. A 10-day state trajectory, without considering the process and measurement noise, is used in the rest of the subsection for selecting appropriate parameters for estimation and determining the minimum number of sensors. It is assumed that the measurements are available every 1 hour.

Following the procedure as discussed in Section 4.1, we apply the PBH observability test on the augmented system to check the identifiability of the parameters. The test is conducted every sampling time, which requires the system to be linearized accordingly. According to the results, the augmented system is not observable. This implies that it is impossible to identify the 5 parameters simultaneously. In order to look for an observable system, parameters are removed from the augmented system. We start with removing only 1 of the parameters and this results in 5 different augmented systems with each one augmented with 4 parameters. Then, the observability of the 5 augmented systems is checked. It was found that 2

Table 1. Initial states and parameters of the investigated loam soil column and initial guesses used in filters and estimator

	$x(0)$ (m)	K_s (m/s)	θ_s (m^3/m^3)	α $(1/m)$	n	θ_r (m^3/m^3)
System	-0.514	2.89×10^{-6}	0.430	3.60	1.56	0.0780
Initial guess	-0.617	3.18×10^{-6}	0.387	3.24	1.72	0.0780

of the 5 systems are observable. In these two subsystems, either θ_s or θ_r is removed. Since observable systems are found, we proceed to the next step to determine the final parameter set.

The significance of θ_s and θ_r is compared based on the sensitivity analysis described in Section 4.2. Sensitivity analysis is conducted based on the original augmented system with all the parameters. By comparing the summation of the absolute values of the elements of each column of the normalized sensitivity matrices S_N, it can be found that the summation corresponding to the column $\frac{\partial y_i}{\partial \theta_s}$ (98600) is much bigger than the one for $\frac{\partial y_i}{\partial \theta_r}$ (17900). Based on this, θ_s is considered as a more important parameter because it has more impact on the output than θ_r. Therefore, the parameter set containing θ_r is removed and the final parameter set will be used in the remaining analysis is $\{K_s, \theta_s, \alpha, n\}$.

When the set of parameters is determined, we determine the minimum number of sensors (measurements) needed to ensure the observability of the augmented system with 4 parameters. Following the method described in Section 4.3, the maximum multiplicity method is conducted, and it can be found that the minimum number of sensors is 4.

5.3 Simultaneous parameter and state estimation

According to the minimum number of sensors found above, 4 tensiometers (N_y) are installed at 7.30 cm, 24.1 cm, 40.8 cm, and 57.6 cm below the surface, which measure the 4^{th}, 12^{th}, 20^{th}, and 28^{th} states, respectively. In the simulations, the actual parameter values used are shown in Table 1 and they are assumed to be homogeneous temporally. Process and measurement noises (ω_x and ν) are considered in the simulations and they have zero mean and standard deviations 3×10^{-6} m and 8×10^{-3} m, respectively.

In the design of the state and parameter filters (EKF, EnKF) and estimator (MHE), the model augmented with 4 parameters (K_s, θ_s, α, and n) is used. The initial guesses of the initial states and parameters in the filters and estimator are listed in Table 1.

For the EKF and EnKF, the weighting matrices Q and R are designed as the auto-covariance matrices of ω_x and ν with the standard deviations mentioned before. However, the diagonal elements of Q corresponding to augmented parameters are 0, because the parameters are assumed to be constant. In simulations, 10^{-20} is used to approximate the value 0 and to ensure the positive definiteness of the matrix. The diagonal elements of P corresponding to the states are configured as the square of 3×10^{-3} and those of parameters are configured as the square of 3×10^{-2}. For the designed EnKF, 100 ensembles are used.

For the design of MHE, the estimation window size is selected to be 8 hours. The weighting matrices P, Q, and R retain the same ratio with respect to those used in EKF and EnKF, but with a much bigger magnitude to ensure the numerical stability of the associated optimization problem. In addition, the P matrix is constant for all the optimizations. The constraints of the states, parameters, and the model uncertainty are listed in Table 2. The upper and lower bounds of the term $\hat{\omega}_p$ are 0 because the parameters are constant.

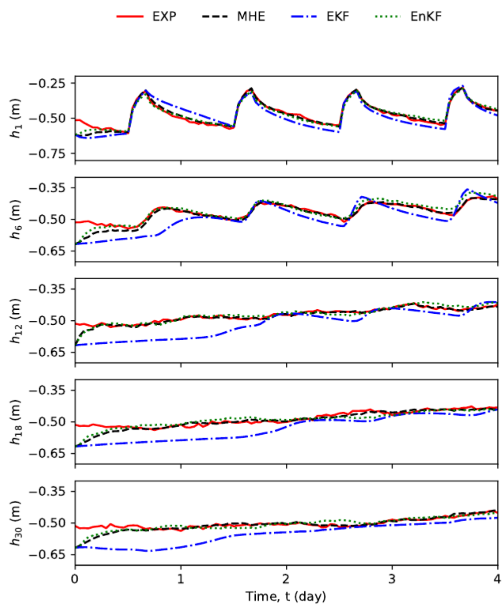

Fig. 3. Selected trajectories of the process state and estimated states using MHE, EKF, and EnKF

The root mean square errors (RMSEs) in terms of states and parameter are used to evaluate the performance of the MHE, EKF, and EnKF. They are shown below:

$$RMSE_x(k) = \sqrt{\frac{\sum_{i=1}^{N_x} (\hat{x}_i(k) - x_i(k))^2}{N_x}} \tag{13}$$

$$RMSE_p(k) = \sqrt{\frac{\sum_{i=1}^{N_p} (\hat{p}_i(k) - p_i(k))^2}{N_p}} \tag{14}$$

First, we performed simulations assuming that the parameter θ_r is known and is the same as the value used in the actual system. Figures 3 and 4 show some representative estimated states and all the parameters using MHE, EKF, and EnKF, which are also compared with their true values. Figure 3 shows the state trajectories of the top node, a few middle nodes, and one bottom node. From the figure, it can be seen that the top node has more dynamics because it takes time for irrigated water to pass from the upper part and to the lower part. In terms of state estimation performance, from Figure 3, it can be seen that MHE and EnKF give very much more accurate state estimates than the EKF. Note that from Figure 3, it can also be seen that the estimates of the 12^{th} state (h_{12}) converge faster than the other estimates. This is because it is a sensor node.

Table 2. Lower and upper bounds used in MHE

	$\hat{x}$ (m)	$\hat{K}_s$ (m/s)	$\hat{\theta}_s$ (m^3/m^3)	$\hat{\alpha}$ $(1/m))$	$\hat{n}$	$\hat{\omega}_x$	$\hat{\omega}_p$
Lower bounds	-1.00	2.31×10^{-6}	0.344	2.88	1.25	$-\infty$	0.00
Upper bounds	-1.00×10^{-4}	3.47×10^{-6}	0.516	4.32	1.87	∞	0.00

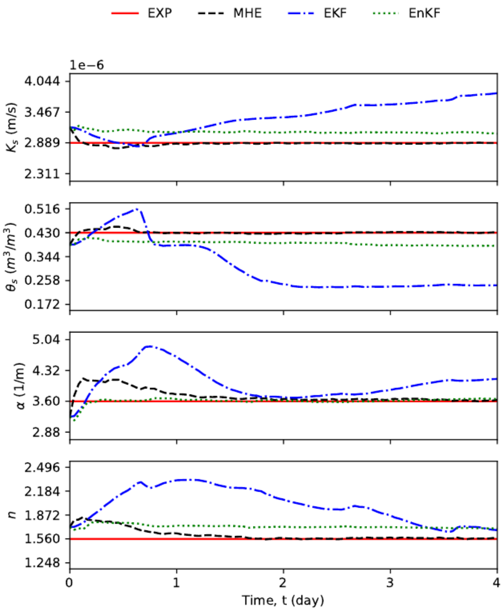

Fig. 4. Trajectories of estimated parameters using MHE, EKF, and EnKF, compared with their actual values

In terms of parameter estimation, Figure 4 shows the results. From the figure, it can be seen that only MHE is capable of estimating the parameters, whereas those estimated by EKF and EnKF diverge from their true values. This may be because of the constraints used in MHE. These constraints provide more useful information to MHE in addition to the measurements.

The average $RMSE_p$ over investigated time, associated with the MHE, EnKF, and EKF are 0.0270, 0.0789, and 0.261, respectively. The average $RMSE_x$, associated with the MHE, EnKF, and EKF are 0.0513, 0.0110, and 0.131, respectively. These values further confirm that the MHE and EnKF have better performance than EKF in estimation of the states, and the MHE outperforms both EnKF and EKF in parameter estimation.

6. CONCLUSIONS

In this work, we have investigated simultaneous parameter and state estimation using MHE, EKF, and EnKF applied to an agro-hydrological system. A procedure was proposed to find the appropriate parameter set for estimation based on observability of the augmented system and the sensitivity of the outputs to the parameters. Our method recommends to consider four parameters with respect to hydraulic conductivity, saturated soil moisture, and van Genuchten-Mualem parameters in estimation. The minimum number of sensors was determined based on the maximum multiplicity theory. Simulation results showed that MHE can provide the better parameter and state estimation performance as compared to EKF and EnKF.

REFERENCES

Carsel, R.F. and Parrish, R.S. (1988). Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, 24(5), 755–769.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5), 10143.

FAO (2016). Aquastat main database. food and agriculture organization of the united nations (fao). website accessed on [31/10/2019].

Gillijns, S., Mendoza, O., Chandrasekar, J., De Moor, B., Bernstein, D., and Ridley, A. (2006). What is the ensemble Kalman filter and how well does it work? In *2006 American Control Conference*, 6 pp. IEEE, Minneapolis, MN, USA.

Hwang, S. and Powers, S.E. (2003). Estimating unique soil hydraulic parameters for sandy media from multi-step outflow experiments. *Advances in Water Resources*, 26(4), 445–456.

Li, C. and Ren, L. (2011). Estimation of unsaturated soil hydraulic parameters using the ensemble Kalman filter. *Vadose Zone Journal*, 10(4), 1205.

Lv, H., Yu, Z., Zhu, Y., Drake, S., Hao, Z., and Sudicky, E.A. (2011). Dual state-parameter estimation of root zone soil moisture by optimal parameter estimation and extended Kalman filter data assimilation. *Advances in Water Resources*, 34(3), 395–406.

Mao, Y., Liu, S., Nahar, J., Liu, J., and Ding, F. (2018). Soil moisture regulation of agro-hydrological systems using zone model predictive control. *Computers and Electronics in Agriculture*, 154, 239 – 247.

Nahar, J. (2019). *Closed-loop irrigation scheduling and control*. Ph.D. thesis, University of Alberta.

Nahar, J., Liu, J., and Shah, S.L. (2019). Parameter and state estimation of an agro-hydrological system based on system observability analysis. *Computers & Chemical Engineering*, 121, 450 – 464.

Rao, C., Rawlings, J., and Mayne, D. (2003). Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations. *IEEE Transactions on Automatic Control*, 48(2), 246–258.

Richards, L.A. (1931). Capillary conduction of liquids through porous mediums. *Physics*, 1(5), 318–333.

Sahoo, S.R., Yin, X., and Liu, J. (2019). Optimal sensor placement for agro-hydrological systems. *AIChE Journal*, 0(0), e16795.

Stigter, J., Joubert, D., and Molenaar, J. (2017). Observability of complex systems: Finding the gap. *Scientific reports*, 7(1), 1–9.

van Genuchten, M.T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44(5), 892.

Yuan, Z., Zhao, C., Di, Z., Wang, W.X., and Lai, Y.C. (2013). Exact controllability of complex networks. *Nature communications*, 4, 2447.