CONTAM 03 Embedded Networked Sensing in Soils: A Stochastic Data Assimilation Approach to Network Design and Real-time State and Parameter Estimation

CONTAM 03.1 People

- Principal Investigators: Steve Margulis, UCLA
- Faculty: Steve Margulis
- Graduate Students: Che-Chuan Wu

CONTAM 03.2 Overview

Data assimilation (DA) provides a framework to estimate the true states of an environmental system by systematically merging multiple sources of information concerning the system (e.g. embedded measurements, remote sensing observations, physically based process models, etc.) and has been widely used in many fields, such as atmospheric science and hydrology. In this project, the main goal is using an ensemble DA scheme (i.e. the ensemble Kalman filter (EnKF)) to merge soil state observations from embedded sensor networks deployed in the Palmdale wastewater reuse testbed site into a coupled unsaturated water flow and solute transport model to provide sufficient, reliable soil state (moisture, temperature and nitrate concentration) and flux estimates for maximizing irrigation rates while preventing groundwater pollution. Based on previous results, the EnKF can provide a significant improvement in soil state estimates while assimilating measurements into the coupled model. However, the estimates of water and nitrate fluxes are not well characterized when uncertain (and erroneous) soil hydraulic parameters are involved. This implies that parameter estimates are necessary to obtain reliable flux estimations. In this report, we summarize research tasks designed to test the approach. The first task is to attempt to estimate these time-invariant parameters through the EnKF scheme for a homogeneous soil profile and test if flux estimates are subsequently improved. The second task is an application of the DA framework to the stratified soil system at the Palmdale study site. For layered soil media, the correlations between states at observed and unobserved locations are reduced due to the different soil parameters. Thus, we also simulate the evolution of soil state by assimilating measurements at different depths to see if better estimates could be obtained while deeper observations are available.

CONTAM 03.3 Approach: Ensemble Data Assimilation Methodology

Due to the complex and nonlinear processes in hydrologic and solute transport models, an ensemble data assimilation approach (i.e. the EnKF) is chosen in this project. The basic concept of the EnKF involves using a Monte Carlo approach to describe how the conditional probability density of the system state evolves over time (between measurements) and how it changes when new measurements are incorporated. In our previous work, the unsaturated water flow and solute transport models were developed and incorporated into the EnKF framework and were shown to provide reasonable estimates of soil states. In this report, we focus on the application of the EnKF to estimate time-invariant parameters in a homogeneous soil medium, and soil state and flux estimates in a stratified soil system.

CONTAM 03.4 Systems/Experiments

Fluxes (e.g. evapotranspiration, water and nitrate flux, etc.) are important variables in environmental systems, especially in this application of wastewater reuse via irrigation. For example, evapotranspiration affects the moisture near soil surface, and nitrate flux estimates tell us the possible source of groundwater contamination. But fluxes are difficult to estimate directly with sensors. In general, flux estimates are usually evaluated with modeling approaches and are functions of the underlying soil states and parameters that characterize the soil. Accurate estimates of the states are not necessarily sufficient to provide accurate estimates of the fluxes. To accurately estimate fluxes, the parameters used in the models should also be well characterized. The approach used for this problem uses state augmentation, whereby the model parameters are treated as time-invariant state variables that are augmented to the original state vector and then estimated with the EnKF framework.

Experiment 1: Parameter estimates in homogeneous soil media with the EnKF approach
In our previous work, we found that the EnKF can significantly improve soil state estimates in homogeneous soil media. However, when we consider uncertainties in model parameters, the EnKF does not accurately estimate water and nitrate fluxes (despite the accurate estimates of soil states). The problem is due to parameter errors. In this experiment, we try to update four time-invariant parameters through the EnKF framework, i.e. saturated hydraulic conductivity, first-order decay rates for organic-N, ammonium, and nitrate, since they play important roles in the processes of water flow and solute transformation. Figure 1 shows soil moisture evolution in the "Truth", open-loop, and EnKF without/with parameter estimation. We can see that the EnKF without parameter estimation can capture the dynamics of soil moisture at measurement times (as expected). Due to parameter errors, the states tend to drift away from the true conditions between measurements. When we consider simultaneous parameter estimation, the soil moisture state estimates improve over those without parameter estimation.

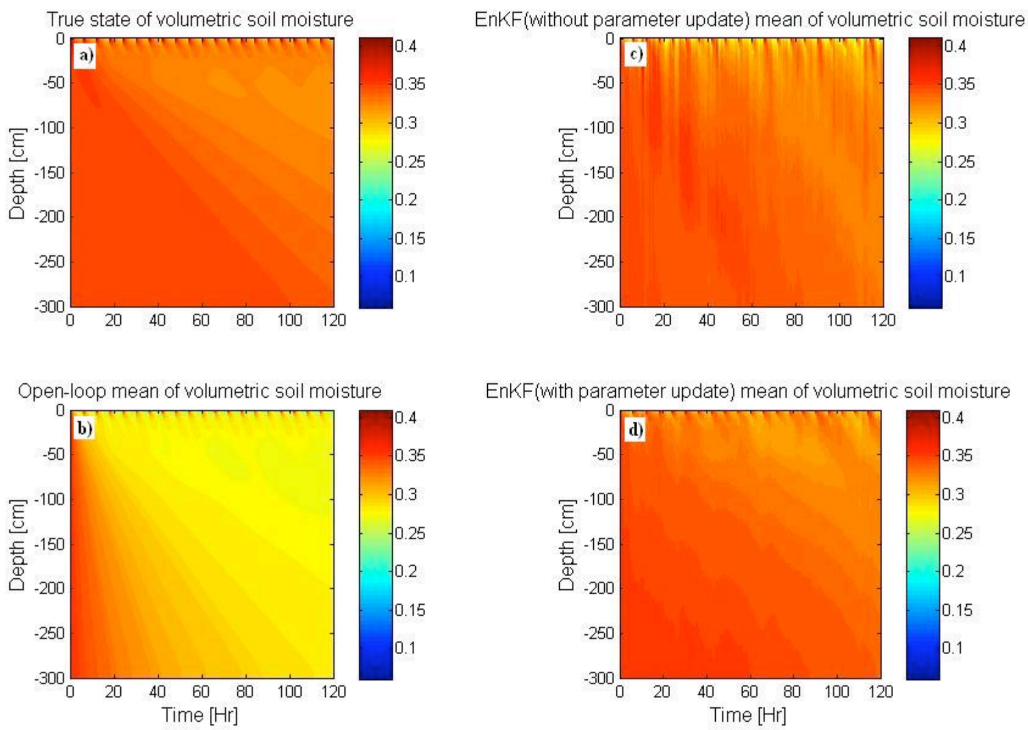

Figure 1. a) True soil moisture evolution compared to: b) open-loop, and EnKF estimates c) without parameter estimation and d) with parameter estimation. The open-loop and EnKF estimates represent the ensemble mean.

Figure 2 shows the parameter estimation (as a function of time) using the EnKF. In the beginning of the simulation the uncertainty ranges of these four parameters are very large, but after several update steps the mean of the ensemble converges to the true parameter values. This implies that under these conditions the EnKF has the ability to simultaneously estimate both states and parameters. Figure 3 shows the bottom water flux predictions in the open-loop and the EnKF without/with parameter estimation. The predicted bottom water flux in the EnKF without parameter estimation performs relatively poorly even though the soil moisture dynamics have been reasonably well captured (shown in Figure 1). However, when we update parameters and soil states simultaneously, it is possible to capture soil moisture and

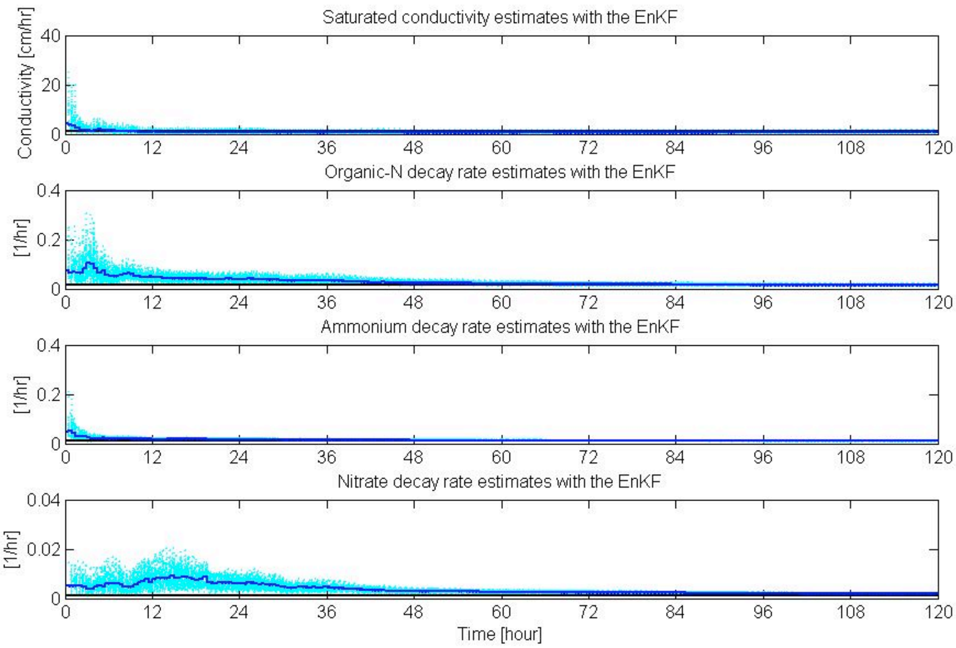

Figure 2. Parameter estimates of the saturated hydraulic conductivity and three first-order decay rates with the EnKF scheme. The true parameter value is shown with the black line, the ensemble replicates are shown as cyan lines, and the ensemble mean is shown with the blue line.

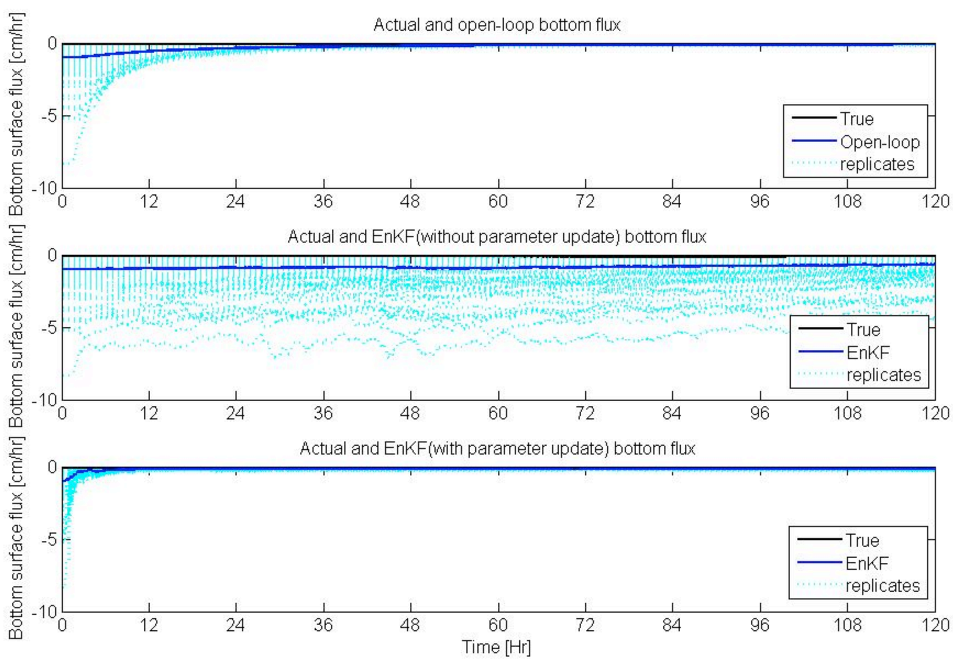

Figure 3. Comparisons of bottom water flux evolutions in the open-loop and the EnKF (with/without parameter estimates) simulations. The true bottom water flux is shown with the black line, the ensemble replicates are shown as cyan lines, and the ensemble mean is shown with the blue line.

State	Index		Bias	RMSE
Soil moisture	Open-loop		-0.0467	0.0483
	EnKF without parameter update		-0.0091	0.0167
	EnKF with parameter update		-0.0028	0.0066
Organic-N (ppm)	Open-loop		-1.9945	2.2683
	EnKF without parameter update		-1.9989	2.3826
	EnKF with parameter update		-0.7202	0.9935
Ammonium (ppm)	Open-loop		-5.4908	5.8278
	EnKF without parameter update		-0.3709	0.8775
	EnKF with parameter update		-0.1782	0.5299
Nitrate (ppm)	Open-loop		3.5167	4.4962
	EnKF without parameter update		0.6446	1.0860
	EnKF with parameter update		0.0747	0.4177
Bottom water flux (cm/hr)	Open-loop		0.1551	0.2647
	EnKF without parameter update		0.7431	0.7475
	EnKF with parameter update		0.0329	0.1021

Table 1. Error metrics for the ensemble mean during a five-day simulation period under biased and uncertain model inputs.

bottom water flux evolution with the EnKF scheme. This experiment implies that to obtain reliable flux estimations, parameter estimates are necessary. Table 1 shows the error metrics to evaluate the performance of different simulations during a five-day simulation. The results clearly indicate that the EnKF with parameter estimations improves both soil state and flux estimates.

Experiment 2: Point scale experiment with heterogeneous (stratified) soil media

For the Palmdale site, borehole results indicate that the soil is a stratified medium. In previous experiments, the EnKF has shown to be capable of estimating system states using sparse near-surface observations for an idealized homogeneous soil case. However, the layered soil media may reduce correlations between states at observed and unobserved locations, which should reduce the effectiveness of the DA system. In this point scale experiment, we tested the ability of the EnKF to capture the dynamics in a stratified (layered) soil system with the same sampling frequency of the simplified homogeneous soil system. In this experiment, we consider two sets of observation depths. One is the same as the setup in the simplified homogeneous system (i.e. at the depth of 20, 40, 60, and 80cm). For another observation setup, we consider observations are available at the center of different soil type layers (i.e. the sensor locations are at the depth of 40, 100, 150, and 200 cm). In this observing system simulation experiments (OSSEs), we set up a three-day spin up period with the open-loop simulation to obtain initial conditions for each replicates. Figure 4 shows the results from the two different sampling cases. The effectiveness of DA system to estimate soil states at deeper locations is reduced since the correlations between observed and unobserved locations are decreasing in the layered soil system. However, based on the figure 4, we can see if the information at deeper locations is available it is possible to have better soil state estimates at deeper locations. The results of this experiment indicate that different sampling locations lead to different results of soil state estimation. For next step, it is worthy to find the optimization of sampling locations through the OSSEs while sensors are limited.

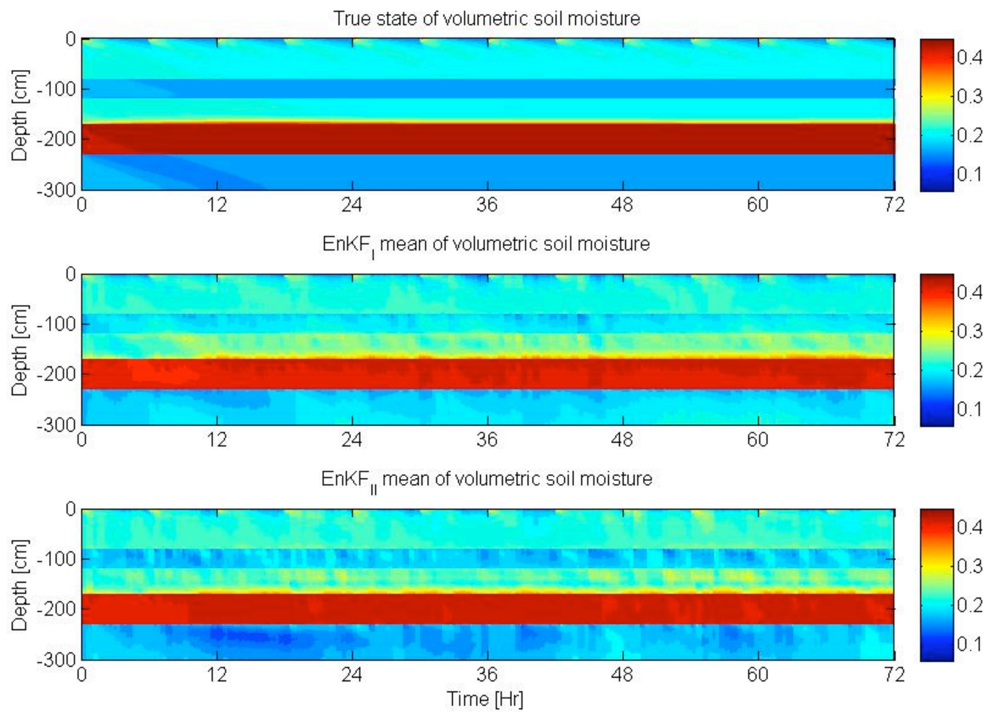

Figure 4. True soil moisture evolution and the EnKF simulations (EnKFI has measurements available at the depths of 20, 40, 60, and 80 cm; EnKFII has measurements available at the depths of 40, 100, 150, and 200 cm).

Experiment 3: Point scale heterogeneous experiment with dynamic vegetation growth conditions

In order to apply the developed EnKF scheme in the Palmdale site for long-term simulations, the coupled unsaturated water flow and solute transport model was modified to account for growing vegetation and intervening fallow periods. This implementation will allow us to consider the crop effects over the Palmdale site dynamically. Figure 5 shows the soil moisture and temperature evolution in the open-loop simulation (with an ensemble of 160 replicates) for one-year simulation. The outcomes will be treated as initial conditions for following-on experiments.

CONTAM 03.5 Accomplishments

Based on the outcomes of the OSSEs described above, significant improvements in state estimate were made when the EnKF is applied under uncertain initial conditions, forcing, and model parameters. Accurate flux estimates can be made with the proposed DA approach while parameter estimates are considered simultaneously. Currently, we are applying the DA framework to a stratified soil system expected in the Palmdale study site, and results show that soil state estimates still can be improved. The coupled unsaturated water flow and solute transport model incorporated with a SVAT model has been developed and more realistic tests involving dynamic vegetation growth are ongoing.

CONTAM 03.6 Future Directions

The specific research tasks to be performed in the future are described briefly below.

Task 1: Online forward model calibration

Based on the outcomes of the OSSEs in homogenous soil media, a significant improvement of system state and flux estimates is made when the EnKF is applied. However, flux estimation for a heterogeneous (layered) soil system has not been fully tested and proven. Experiments will be undertaken to assess whether the DA approach can correct the time-invariant soil hydraulic parameters/first-order decay rates and improve flux estimates in a stratified soil system.

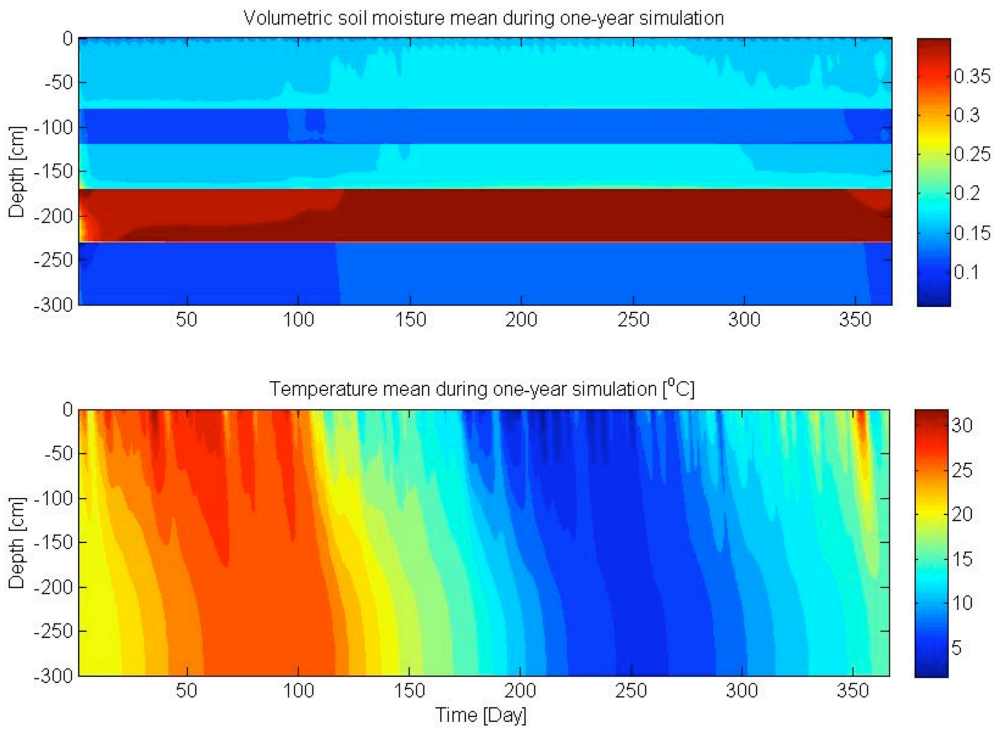

Figure 5. Soil moisture and temperature evolution in the open-loop simulation (with an ensemble of 160 replicates) for one-year simulation.

Task 2: Spatially-distributed experiments

In general, it is desirable to have state estimates over the entire three-dimensional domain of interest, especially when horizontal heterogeneities in the system or system forcing (in this case irrigation) exist. For the Palmdale site, the primary variability in soil properties is in the vertical. Therefore, horizontal spatial heterogeneity exists largely due to the pivot irrigation system, which introduces a radially-symmetric heterogeneity in water and contaminant input. To avoid the computational expense of a three-dimensional model for the Palmdale site we will take advantage of the symmetry of the problem and use a two-dimensional (2D) strip domain (i.e. along a particular radius). This setup will allow for estimates over the entire domain along the irrigation pivot trajectory. The assimilation system can then use measurements at select locations in the domain to estimate states everywhere in between (and at depth) and ultimately provide feedback on the irrigation rate in future pivot rotations. OSSEs similar to those described above will be used to test whether the states and fluxes at unobserved areas in the distributed domain can be estimated accurately.