Development of Optimization and posterior inference tools in NASA Land Information System (LIS)

Sujay Kumar, Joseph Santanello, Christa Peters-Lidard, Rolf Reichle, Ken Harrison, Soni Yatheendradas

NASA Goddard Space Flight Center, Hydrological Science Branch, Greenbelt, MD USA

1. Introduction

The NASA Land Information System (LIS; http://lis.gsfc.nasa.gov; Kumar et al. 2006, Peters-Lidard et al. 2007) is a high-resolution, high-performance, land surface modeling and data assimilation system to support a wide range of land surface research and applications. LIS integrates various community land surface models, ground and satellite-based observations, and ensemble-based data assimilation tools to enable assessment and prediction of hydrologic conditions at various spatial and temporal scales of interest.

LIS facilitates the integration of observations from Earth-observing systems and predictions and forecasts from Earth System and Earth science models into the decision-making processes of partnering agency and national organizations. Due to its flexible software design, LIS can serve both as a Problem Solving Environment (PSE) for hydrologic research to enable accurate global water and energy cycle predictions, and as a Decision Support System (DSS) to generate useful information for application areas including disaster management, water resources management, agricultural management, numerical weather prediction, air quality and military mobility assessment. More recently, the LIS infrastructure has been enhanced with the development of a comprehensive optimization infrastructure that provides a suite of deterministic and stochastic algorithms, for applications such as the optimal estimation of model parameters. The ongoing work extends this infrastructure to include a suite of uncertainty modeling approaches that enables posterior probabilistic model predictions.

2. Approach

To demonstrate the new optimization infrastructure in LIS, three optimization algorithms were implemented: (1) Levenberg-Marquardt (LM), (2) Genetic Algorithm (GA) and (3) Shuffled Complex Evolution from University of Arizona (SCE-UA). The three algorithms represent a range of search strategies suitable to different problem types. LM exploits local information (numerically evaluated derivatives) to efficiently identify the optimum in problems in which a single local optimum exists. In contrast, GAs use a random search strategy that helps avoid convergence to local, inferior optima. SCE-UA combines local and random search strategies.

Two test cases involving the estimation of soil parameters were formulated to demonstrate the implementation of the algorithms and the optimization infrastructure. Test case I was based on the prior work (Santanello et al. (2007)) that involved determining soil properties given estimates of near surface soil moisture derived from passive (L-band) microwave remote sensing over the Walnut Gulch watershed in Southeastern Arizona. Errors in the simulated versus observed soil moisture were minimized by adjusting the soil texture, which in turn controls the hydraulic properties through the use of pedotransfer furnctions (PTF). To further test the algorithms, test case II was formulated by modifying test case I to directly estimate the soil properties. Test case II, therefore presents a more challenging optimization problem, since it involves the estimation of more parameters compared to test case I.

3. Results

The soil moisture simulations are conducted over two sites in the Walnut Gulch watershed: (1) Lucky Hills and (2) Kendall. The Noah land surface model in LIS is employed at 40m spatial resolution, forced with observed meteorology at these locations. Figure 1 shows the results from test case II, where the soil moisture simulations from the using the default model parameters (control) is compared against the estimates generated by the model using the optimized parameters, using the three optimization algorithms (LM, GA, SCE-UA). It can noticed that at both the sites, the optimized parameters improves the ability of the model to match the simulated soil moisture much more closely than when compared to the control that relied on mappings of hydraulic parameters from soil texture classifications.

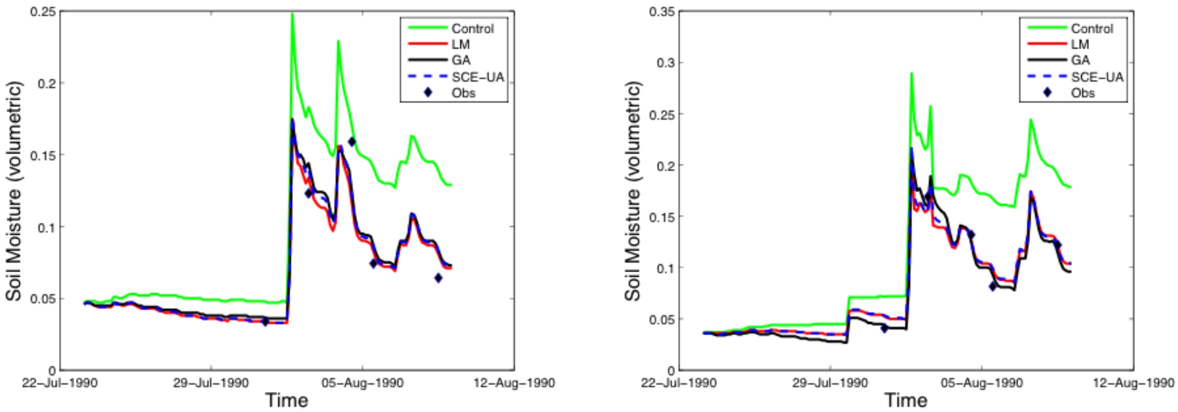
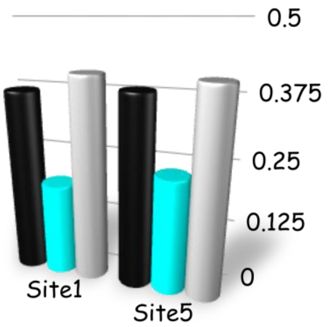
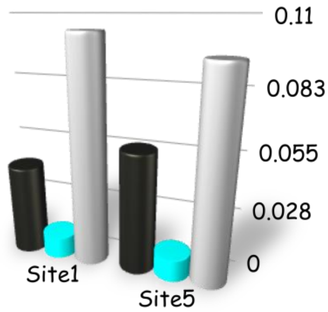


Figure 1: Comparison of soil moisture simulations at Lucky Hills (left) and Kendall (right). The control represents the model simulation using the default parameters, and LM, GA, and SCE-UA represents the model simulations using the optimized parameters which were estimated by LM, GA, and SCE-UA, respectively. The PBMR soil moisture observations are also shown.

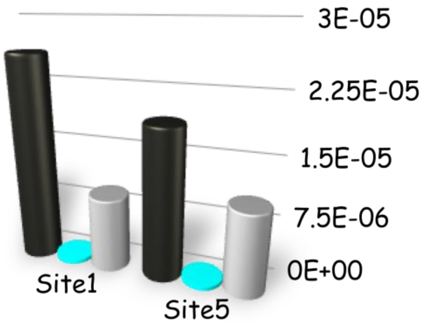
The optimized parameters from test case I and II are compared against independent estimates of the soil properties available at these locations and is shown in Figure 2. It can be observed that the PTF based (test case I) parameters show more agreement with the in-situ measurements, relative to the parameters estimates from test case II. This result suggests that though the use of PTF allows for calibrated parameters to be constrained by physical plausibility and forces physical consistency. In contrast, in estimating the parameters directly in test case II, the physical realism of the parameters may be lost.



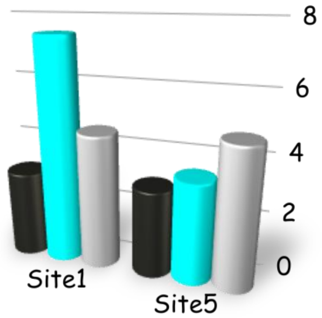
(a)



(b)



(c)



(d)

Figure 2: Shows a comparison of parameter values from test cases I and II and Lucky Hills (site 1) and Kendall (site 5): (a) soil porosity, (b) saturated matric potential, (c) saturated hydraulic conductivity, (d) pore size distribution index. The black bars represent the PTF-based parameters from test case I, the cyan bars represent parameters estimated directly from test case II and the gray bars represent the in-situ estimates.