Continuous data assimilation for downscaling large-footprint soil moisture retrievals

Muhammad U. Altaf\textsuperscript{a,b}, Raghavendra B. Jana\textsuperscript{a}*, Ibrahim Hoteit\textsuperscript{b}, Matthew F. McCabe\textsuperscript{a}

\textsuperscript{a}Water Desalination & Reuse Centre, Biological, Environmental Sciences & Engineering Division; 
\textsuperscript{b}Earth Sciences and Engineering Division; 
King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.

ABSTRACT

Soil moisture is a key component of the hydrologic cycle, influencing processes leading to runoff generation, infiltration and groundwater recharge, evaporation and transpiration. Generally, the measurement scale for soil moisture is found to be different from the modeling scale for these processes. Reducing this mismatch between observation and model scales is necessary for improved hydrological modeling. An innovative approach to downscaling coarse resolution soil moisture data by combining continuous data assimilation and physically based modeling is presented. In this approach, we exploit the features of Continuous Data Assimilation (CDA) which was initially designed for general dissipative dynamical systems and later tested numerically on the incompressible Navier-Stokes equation, and the Benard equation. A nudging term, estimated as the misfit between interpolants of the assimilated coarse grid measurements and the fine grid model solution, is added to the model equations to constrain the model’s large scale variability by available measurements. Soil moisture fields generated at a fine resolution by a physically-based vadose zone model (HYDRUS) are subjected to data assimilation conditioned upon coarse resolution observations. This enables nudging of the model outputs towards values that honor the coarse resolution dynamics while still being generated at the fine scale. Results show that the approach is feasible to generate fine scale soil moisture fields across large extents, based on coarse scale observations. Application of this approach is likely in generating fine and intermediate resolution soil moisture fields conditioned on the radiometer-based, coarse resolution products from remote sensing satellites.

Keywords: soil moisture, continuous data assimilation, remote sensing, modeling, vadose zone, scaling, CDA

1. INTRODUCTION

Models seeking to characterize runoff, infiltration, recharge and evapotranspiration related processes require soil moisture as an input, either directly or indirectly. It has been shown by previous studies \cite{1, 2} that better characterization of the spatial variability of soil moisture leads to better predictions from hydrologic/climate models \cite{3}. However, the scale of soil moisture measurement is generally inconsistent with the modeling scale for these processes \cite{4, 5}. Ground-based, in-situ measurements have fine resolutions \cite{6}, but become impractical in terms of coverage over large spatial extents. Satellite-based, remotely sensed data have excellent spatial coverage extents, but suffer from comparatively poorer resolution, both spatial and temporal \cite{7}. Bridging the gap between these two commonly-used measurement techniques in necessary for better hydrological modeling and thus, scaling of soil moisture data is a topic of considerable interest in the hydrologic community \cite{8-10}.

Data assimilation has been used in hydrology in general, and in unsaturated zone hydrology in particular, to improve model predictions of state variables such as the soil moisture \cite{11-14}. Most studies involve statistical techniques such as the Ensemble Kalman filter (EnKF) \cite{15, 16} or variational approaches \cite{14, 17}. Based on the assumption of a linear underlying dynamical model and Gaussian distribution of error terms, these approaches identify the state of the system that has the highest probability given the previous state and observations \cite{18, 19}. A major disadvantage of such methods is possible dynamic imbalances and sub-optimalities due to their linear nature and localization. A recently developed technique, called Continuous Data Assimilation (CDA), overcomes such drawbacks by incorporating a nudging term in the model.

*Raghavendra.Jana@kaust.edu.sa; phone +966544700173; hydrology.kaust.edu.sa
equations to constrain the large scale variability in model variability to available observations \cite{20}. This CDA approach was initially analyzed theoretically in various scenarios \cite{21,22} and recently tested numerically on the two-dimensional incompressible Navier-Stokes equations \cite{23}. However, no numerical studies have been performed yet in the field of hydrology, and as such, this study is a step towards applying this methodology to hydrological applications. In this paper, soil moisture fields generated at a fine resolution by a physically-based unsaturated zone model are subjected to data assimilation conditioned upon the coarse resolution observations using the CDA approach. This enables nudging of the model outputs towards values that honor the coarse resolution dynamics while still being generated at the fine scale. The objectives of this study were to show that the CDA approach can be used to nudge fine scale model outputs towards honoring coarse scale observations, and also to study the effect of observation scale on the CDA performance.

2. METHODOLOGY

2.1 Continuous Data Assimilation

In the Continuous Data Assimilation (CDA) approach, a nudging term, estimated as the misfit between interpolants of the assimilated coarse grid measurements and the fine grid model solution, is added to the model equations to constrain the model’s large scale variability by available measurements. The interpolants are applied on both model outputs and coarse data before nudging, allowing to only constrain the large scale flow of the model. Let $I_h(\phi(x))$ be an interpolation operator of a function $\phi(x)$. For instance, let:

$$I_h(\phi(x)) = \sum_{k=1}^{N_h} \phi(x_k) \chi_k(x)$$

where $Q_j$ are disjoint subsets such that $\text{diam}(Q_j) \leq h$; $\bigcup_{j=1}^{N_h} Q_j = \Omega$; $x_j \in Q_j$, $x_E$ is the characteristic function of the set $E$, and $\phi$ is a suitable interpolant. Consider the updated system of model equations:

$$\psi^u(t) = \psi(t) - \mu \left( I_h(\psi^p(t)) - I_h(\psi(t)) \right)$$

To accurately predict $\psi^u(t)$ on the interval $[t_1, t_1 + T]$, it is sufficient to have coarse observational data $I_h(\psi(t))$ accumulated over an interval of time $[0, t_1]$ linearly proportional to $T$ in the immediate past. In particular, suppose it is desired to predict $\psi^u(t)$ with accuracy $\epsilon > 0$ on the interval $[t_1, t_1 + T]$, where $t_1$ is the present time and $T > 0$ determines how far into the future one wants to predict. Let $h$ be small enough and $\mu$ be large enough, then there exist constants $C$ and $\beta$ such that:

$$\|\psi(t) - \psi^u(t)\|_{H^1(\Omega)} \leq C e^{-\beta t}, t \geq 0$$

One then uses $\psi^u(t_1)$ as the initial condition to make future predictions. Let $W$ be a solution with initial condition $W(t_1) = \psi^u(t_1)$. Known results on continuous dependence on initial conditions imply the existence of $\gamma > 0$ such that:

$$\|W(t) - \psi(t)\|_{L^2(\Omega)} \leq \|W(t_1) - \psi(t_1)\|_{L^2(\Omega)} e^{\gamma(t-t_1)}, t \geq t_1$$

Therefore,

$$\|W(t) - \psi(t)\|_{L^2(\Omega)} \leq C e^{-\beta t_1 + \gamma T}, t \in [t_1, t_1 + T]$$

provided $\beta t_1 \geq \gamma T + \ln(C/\epsilon)$. Thus $W(t)$ predicts $\psi(t)$ with accuracy $\epsilon$ on the interval $[t_1, t_1 + T]$. 
2.2 Modeling domain setup

The unsaturated zone moisture dynamics was modeled using the physically-based HYDRUS-3D hydrologic simulation software package[24]. The program solves a modified form of the Richards’ equation for water flow in saturated/unsaturated domains using numerical techniques and allows the user to study flow through saturated, partially saturated or unsaturated regions with irregular boundaries, and composed of heterogeneous soil cover. As the name suggests, HYDRUS-3D allows for three dimensional flow representations in the unsaturated zone. Since we intend to consider a realistic view of the water flow process involved in a large study domain, we decided to use the HYDRUS-3D model.

The modified form of the Richard’s equation is given by:

\[
\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x_i} \left[ K \left( K_{ij}^A \frac{\partial h}{\partial x_j} + K_{iz}^A \right) \right] - S
\]  

where \( \theta \) is the volumetric water content, \( h \) is the pressure head, \( S \) is a sink term, \( x_i(=1,2) \) are the spatial coordinates, \( t \) is time, \( K_{ij}^A \) are components of a dimensionless anisotropy tensor \( K^A \), and \( K \) is the unsaturated hydraulic conductivity, given by:

\[
K(h, x, y, z) = K_s(x, y, z) K_r(h, x, y, z)
\]

where \( K_r \) and \( K_s \) are the relative and saturated hydraulic conductivities, respectively.

A cuboidal domain measuring 144 km x 144 km x 5 m was generated with a finite element mesh spacing of 1 km. This results in a total of 21,025 finite element nodes per layer, and there were two layers used in this study. The top boundary was assigned as an “atmospheric” boundary condition that allows moisture movement both into and out of the domain in the vertical direction. The bottom boundary was assigned as a free draining boundary. The CDA scaling methodology was tested under multiple scenarios with respect to initial conditions and soil distribution to validate its applicability. Figure 1 shows two soil distribution patterns, S1 and S2, which were used in this study. Soil moisture fields corresponding to “observations” were generated on a daily time step for 30 days by initializing the simulations at the field capacity value of each soil type. These 1 km resolution outputs – considered as the “true” values – were coarsened to the study scales (9 km and 36 km) by interpolation to the nearest value at that scale. These coarsened values are used as the observations.

Initial conditions (IC; and specified as the initial water content of the soils) used in this study consisted of a uniform water content of 0.3 cm$^3$/cm$^3$ across the entire domain (IC1), a value close to the saturation water content (0.9 \( \theta_s \)) of each soil type (IC2), and a value close to the residual water content (1.1 \( \theta_r \)) of each soil type (IC3). These values were chosen to verify the robustness of the scaling methodology when initiated with widely varying conditions. A time series of rainfall data was used to inject moisture into the domain at specified time steps.

Soil moisture dynamics of the study domain were simulated for one day at a time. At the end of each day, the observed data were assimilated and the model outputs nudged towards the observations. A non-biased nudging weight of 0.5 was used to assign equal weights to the model and observed data. Following the convention of [12], any assimilation outputs that fall beyond the soil moisture range of the model (\( \theta_r \) to \( \theta_s \)) are reassigned a value such that \( \theta = \theta_r + 0.05 \) if \( \theta < \theta_r \); and \( \theta = \theta_s - 0.05 \) if \( \theta > \theta_s \).
3. RESULTS AND DISCUSSION

Figure 2 shows, for one sample soil distribution and initial condition combination, the fine scale (1 km) model outputs, the coarse scale (9 km) observations, the interpolated surfaces for the modeled and observed soil moisture (Model\textsuperscript{int} and Obs\textsuperscript{int}), the mismatch (Obs\textsuperscript{int}-Mod\textsuperscript{int}), and the updated values which become the initial conditions for the next time step of simulation in HYDRUS, at the end of the 30 day simulation period. It is seen that while the fine scale model outputs (panel a) do not appear to match well with the coarse scale observations (panel c), the interpolated values are very similar to each other (panels b and d). The mismatch (panel e) is also close to zero across much of the domain.

In order to assess how close the updated model soil moisture fields were to the true soil moisture field, the L\textsuperscript{2} norm, which provides the distance between any two datasets, was chosen as the measure of performance. The lower the L\textsuperscript{2} norm, the less is the error between the two datasets. A sample of the L\textsuperscript{2} norms for the above case, computed as the distance between the fine scale (1 km) modeled/updated soil moisture field and the “true” soil moisture field used to generate the observations is shown in Figure 3.
Figure 2. Sample soil moisture fields for soil distribution pattern S1. (a) Fine scale (1 km) model outputs; (b) Model output interpolated to coarse scale (9 km); (c) Observed data at 9 km; (d) Observed data interpolated at coarse scale; (e) Mismatch between interpolated observed and modeled data; and (f) Updated soil moisture field after CDA.
When the HYDRUS program is run on an open loop, i.e., without implementing the CDA algorithm, it is seen that the simulations, initiated with an initial condition different from the true condition, slowly converges towards the truth (solid green line with diamond markers). This behavior is expected, and is the reason why most simulation studies allow for a model spin-up period to allow the model state to reach conditions close to the truth [25]. It is also seen that implementing the CDA algorithm rapidly decreases the norm at each step, and the curve (solid black line with triangular markers) becomes asymptotic around a norm value of 1.5 by around the 6th time step. In contrast, the open loop norm continues to stay around a value of 5 even at the end of 30 time steps. This shows that the data assimilation procedure increased the efficiency of the model in generating soil moisture fields close to the truth. The dashed blue line tracks the model output at each time step before the assimilation procedure. It is seen that at each step before converging with the black line, the model outputs have a norm slightly lower than the norm of the updated soil moisture fields from the previous time step. This indicates that the model is still converging towards the truth within each time step independent of the CDA.

Figure 4 displays the norms for open loop and assimilated soil moisture fields for the soil distribution pattern S1 for the three initial conditions and at the two spatial resolutions (9 km and 36 km). The dashed lines represent open loop simulations, while the solid lines represent the CDA updated values. It can be seen that for both data resolutions, the norms for the CDA updated values are far lower than the open loop values. Further, it is seen that the model, when left to itself, does not appear to converge towards the truth when initial conditions close to the residual water content of the soils is used. However, the CDA algorithm overcomes this, and performs almost similar to the case where a uniform soil moisture is used as the initial condition. It can also be deduced from the figure that the case where the simulations were run with initial soil moisture conditions close to the saturation value of the soils performs the best, and displays the lowest norms when the CDA algorithm is implemented.
Figure 4. L2 norms for open loop and post-CDA data for soil distribution pattern S1 at two spatial resolutions, for each initial soil moisture condition. Dashed lines represent open loop data, solid lines represent CDA updated data.

Similar behavior is displayed in Figure 5, where the norms for soil distribution pattern S2 are plotted. This consistency across soil distribution patterns indicates that the CDA algorithm is robust, and can be applied across different domains. A check for the RMSE values showed that they were consistently around $0.02 \text{ cm}^3/\text{cm}^3$ for the updated fields once the norm curve became asymptotic, while the RMSE for the open loop soil moisture was consistently above the $0.05 \text{ cm}^3/\text{cm}^3$ value. Most satellite remote sensing platforms have estimation errors around $0.04 \text{ cm}^3/\text{cm}^3$ \cite{6}. Thus, while the norms not reaching an ideal value of 0 indicates the updated soil moisture fields are not exactly similar to the truth, the errors between the two is less than the accuracy of the remote sensors. This means that the CDA updated soil moisture fields could be considered statistically similar to the true soil moisture.

The above results indicate that the CDA methodology can be successfully applied to unsaturated zone hydrology to nudge fine scale model outputs towards honoring coarse resolution observations. As a result, given the coarse resolution observations (e.g., from remote sensing platforms), fine resolution soil moisture maps can be generated across large spatial extents. Further, it is also possible to generate soil moisture maps at intermediate resolutions (between the model and observation resolutions) that can be useful for other hydrological modeling efforts, including improvements in the estimation of spatially distributed evaporation \cite{26, 27}, informing other downscaling approaches \cite{28, 29}, or in verifying land surface model output \cite{30, 31}.

It was seen that the spatial resolution of the observations influence the performance of the scaling. In most cases, data at 36 km resulted in higher norms for the updated soil moisture fields in comparison to those at 9 km. However, the difference was small. Further detailed experiments are necessary to assess the sensitivity of the CDA algorithm to the observation scale.
During the study, it was observed that in domains with completely random soil distributions, the CDA algorithm failed to perform well. While that is a completely hypothetical condition not expected to be encountered in the real world, it does bring out a limitation of the method. In the real world, soils are distributed in patches, and have some geostatistical correlation lengths. This structured distribution enables the interpolation to be more realistic. In a completely random distribution, the fine scale variations between the interpolation points (coarse grid) are entirely ignored, thus leading to updated soil moisture fields that may not resemble the true states. The norms in that case would be high for the updated data.

4. CONCLUSIONS

The feasibility of applying the Continuous Data Assimilation approach to update fine scale modeled soil moisture fields towards honoring coarse resolution observations was studied. The study results show that the approach is applicable to such efforts and improves the speed of convergence of the model towards the true states. It was shown that fine resolution soil moisture maps can be produced across large spatial domains, provided the coarse resolution observations are available. Application of this approach is likely in generating fine and intermediate resolution soil moisture fields conditioned on the radiometer-based, coarse resolution product from NASA’s SMAP satellite. Further studies to assess the sensitivity of the CDA algorithm to observation scales, topography, and rainfall patterns are envisaged.
REFERENCES


