RECENT ADVANCES ON SOIL MOISTURE DATA ASSIMILATION

Wenge Ni-Meister

Department of Geography
Hunter College and Program in Earth and Environmental Sciences
The City University of New York
New York, New York 10021

Abstract: This study reviews recent progress on soil moisture data assimilation. Data assimilation is a process of merging observations with a system dynamic model to provide an improved estimate of the states of the environment. The application of data assimilation in hydrology is relatively new, however, rapid progress has been made in the last decade or so with the available remotely sensed soil moisture data. After briefing the history of soil moisture data assimilation, the review focuses on the most common data assimilation methods and recent progress made in soil moisture data assimilation through a case study of the soil moisture initialization activities for NASA's seasonal and interannual climate prediction. The example demonstrates that soil moisture data assimilation has made great progress in the last decade, however is still in its infancy. Good quality remotely sensed soil moisture data with accurate uncertainty information at continental and global scale are needed to ensure the success of the operational use of soil moisture data assimilation technique. Further advancement on the current soil moisture data assimilation methods is necessary to be able to assimilate multisource hydrological remote sensing data into land surface models for the best use of various remote sensing data sources at continental and global scales. [Key words: data assimilation, soil moisture, remote sensing, Kalman filter.]

INTRODUCTION

Soil moisture is a key hydrological state variable that integrates much of the land surface hydrological and biophysical processes. Soil moisture interacts with the overlying atmospheric boundary layer through exchanges of sensible and latent heat fluxes, thus influences the atmospheric circulation. Soil has the ability to store precipitated water from periods of excess for later evaporation and to “remember” the wet or dry weather conditions longer than atmospheric processes (Koster and Suarez, 2000; Dong, Ni-Meister, and Houser, 2007). Accurate knowledge of the spatial distribution and temporal variation of soil moisture would provide insight into larger-scale hydrological processes and would serve as good land surface moisture initialization states in fully coupled climate system models for improved seasonal-to-interannual climatological and hydrological prediction (Koster et al., 2004).

Obtaining accurate soil moisture datasets at large spatial scales over a long period is not an easy task. In situ soil moisture measurements are generally expensive and no large-area soil moisture networks exist to measure soil moisture at high temporal frequency in multiple soil depths. Alternatively, many land surface process models were developed and are available for calculating soil moisture,
however, modeled soil moisture data are often prone to error due to insufficient model physics, inaccurate parameterization, and initialization states.

Remote sensing, on the other hand, is ideal for obtaining data at large scales. Airborne microwave remote sensing data are available to retrieve large-scale soil moisture (Jackson et al., 1995, 1999; Jackson, 1997). Thanks to satellite remote sensing, global quarter-degree resolution near surface soil moisture content has been derived using C-band passive microwave observations from the Nimbus-7 satellite Scanning Multi-frequency Microwave Radiometer (SMMR) for the 1979 to 1987 period of operation (Owe et al., 2001). Moreover, the Advanced Microwave Scanning Radiometer for the Earth observing system (AMSR-E) on the Aqua satellite launched in 2001 provides global C-band passive microwave soil moisture data at 25 km spatial resolution. While no C-band measurements are available between SMMR and AMSR-E, low-latitude soil moisture has been estimated using Tropical Rainfall Measuring Mission (TRMM) X-band microwave observations since 1998 (Bindlish et al., 2003; Gao et al., 2004, 2006). Further, L-band passive microwave soil moisture data is expected to be available from 2008 with the launch of the Soil Moisture and Ocean Salinity (SMOS) mission. The advantage of satellite data is that they provide a uniform spatial and temporal coverage of the land surface soil moisture data. But one limitation of these microwave-based soil moisture estimates is their confinement to the top few centimeters of soil and spatial and temporal gaps in areas with significant vegetation, soils, and roughness error sources. However, remote sensing data, when integrated with a land surface process model, will be able to provide estimates of soil moisture both at surface and deep in the rootzone with higher spatial and temporal resolution with less error than either remotely sensed data or model simulations alone. Data assimilation offers a means to combine the advantages of modeling with those of remote sensing data.

The use of data assimilation in hydrology is relatively new, however, it has shown promise for improving hydrologic predictions by incorporating remotely sensed soil moisture (e.g., Houser et al., 1998; Walker et al., 2001, 2002; Reichle, McLaughlin, et al., 2002; Reichle, Walker, et al., 2002; Reichle and Koster, 2005; Zhang et al., 2005; Ni-Meister et al., 2006) and snow (e.g., Rodell and Houser, 2004; Sun et al., 2004; Dong, Walker, et al., 2007) into land surface models. The focus of this paper is to review the recent progress made in soil moisture data assimilation. Section 2 summarizes the history of soil moisture/hydrological data assimilation; Section 3 introduces the basic concepts and common methods used in soil moisture data assimilation; Section 4 presents a case study of soil moisture data assimilation. The final section discusses the future direction of soil moisture data assimilation.

BRIEF HISTORY OF SOIL MOISTURE DATA ASSIMILATION

Data assimilation techniques have widely been used in meteorology and oceanography for decades (see Daley, 1991, for detailed descriptions). With substantial enhancement of the remote sensing instruments measuring various atmospheric quantities and the improvements in data assimilation techniques over the last few years, satellite data have been assimilated in Numerical Weather Prediction (NWP) models operationally. This resulted in the improvement of middle range (3 days to
two weeks) weather predictions to be as good as short range (1 day) predictions (Kalnay, 2003). The use of data assimilation in oceanography also resulted in improved prediction of ocean dynamics (Bennett, 1992). Data assimilation methods have been well demonstrated as an effective technique to improve predictions of atmospheric and oceanic dynamics processes.

With the well-developed data assimilation methods from meteorology and oceanography, hydrological data assimilation, particularly soil moisture data assimilation has made great progress in the last decade or so. Jackson et al. (1981) was the first to use the direct insertion approach to update soil moisture with remotely sensed surface soil moisture when observations are available. Bernard et al. (1981) uses synthetic observations to predict surface fluxes based on the direct insertion method. Prevot et al. (1984) and Brucker and Witono (1989) use a similar approach, but with real remote sensing data. More recent progress made on soil moisture data assimilation is to test the feasibility of optimal soil moisture data assimilation methods. The research focuses on: (1) the use of Kalman filtering based sequential data assimilation strategy and (2) the Kalman smoothing or variational data assimilation strategy.

Kalman filter (Gelb, 1974) is an optimal sequential data assimilation method, which provides the optimal estimate for the current and future states for linear systems when observations are available. To apply it to nonlinear hydrological processes, the extended Kalman filter has been successfully applied in soil moisture data assimilation by linearizing the hydrological processes (Entekhabi et al., 1994; Walker and Houser, 2001; Walker et al., 2001, 2002). It has demonstrated how data assimilation can extend remotely sensed measurements of surface conditions into the underlying soil column. Ensemble Kalman filter was implemented as an alternative to avoid the calculation of the derivative of a tangent linear model to approximate the nonlinear hydrological processes (Margulis et al., 2002; Reichle, McLaughlin, et al., 2002; Reichle, Walker, et al., 2002; Crow and Wood, 2003).

The feasibility of synthesizing distributed fields of soil moisture using the variational (four dimensional) data assimilation was first demonstrated by Houser et al. (1998). A more complicated variational approach has been applied in hydrology (Reichle and McLaughlin, 2001; Reichle et al., 2001; Margulis and Entekhabi, 2003). This approach uses observations distributed in space and time with the knowledge of temporal evolution of the model state through the calculation of the adjoint, which is obtained by linearizing the forward model along a trajectory producing the tangent-linear model. However, for the nonlinear hydrologic processes, calculating the adjoint can be difficult. To overcome this problem, Dunne and Entekhabi (2005, 2006) developed an ensemble Kalman filter based smoothing approach, which uses ensemble Kalman filter to update not only the current and future states, but also the states in the past while avoiding calculating the adjoint. They demonstrated that it can be used operationally for soil moisture data assimilation (Dunne and Entekhabi, 2006).

To facilitate the operational application of data assimilation in hydrology for land surface initialization, the Land Data Assimilation Systems (LDAS) conceptual framework has been developed to provide the best estimation of the current state of land surfaces through integration of satellite observations and modeling at a wide range
of scales. Several LDAS systems have been implemented in near real time and at high spatial resolutions for North American (Mitchell et al., 2004; Cosgrove et al., 2003), European (Van den Hurk, 2002), and global domain (Rodell et al., 2004; Kumar et al., 2006; Peters-Lidard et al., 2007). These LDASs are forced with real time numerical weather prediction models outputs with satellite observed surface parameters. Data assimilation methods are being implemented in these systems to incorporate satellite-observed state conditions as a constraint to the model dynamics. For example, direct insertion, rule-based approaches, and ensemble Kalman Filter (EnKF) were implemented in the current Land Information System (http://lis.gsfc.nasa.gov) to assimilate multiple observations (soil moisture and snow water equivalent [SWE]) into multimodels (Kumar et al., 2008). This has improved land surface energy balance and water cycle estimations.

DATA ASSIMILATION METHODS

Soil moisture data assimilation methods range from the simplest to the most complicated. The simplest data assimilation method is direction insertion, in which the model soil moisture values are replaced with observations whenever they are available with the assumption that the observations are perfect. The relatively more complicated method is the static updating technique, in which the background/model error is approximated and fixed over time and the inversion is simplified and/or solved iteratively. Common static assimilation methods include statistical correction, successive correction, optimal interpolation, analysis correction, Newtonian nudging, and 3-D variational data assimilation. The above two methods are easy to implement and often used for operational applications. The most complex method, but commonly used in soil moisture data assimilation, is optimal data assimilation algorithms that adjust uncertain parameters to obtain a best fit to observations at each time step. The selection of a data assimilation procedure is dependent on a given application together with a balance between making the best use of available information (optimality), computational efficiency, flexibility, and robustness. The following provides more details on optimal data assimilation methods.

Optimal data assimilation refers to the process varying the model and observational error at each time in a dynamical system to minimize the error between the model predictions and observations. It can be divided into two categories: sequential data assimilation or filtering and 4-D variational data assimilation or smoothing. In the sequential approach, observations are used to update the state at the current and the future states. In the variational assimilation, observations are used to update not only the current and the future state, but also the past state and it uses observations distributed in space and time to update the knowledge of the temporal revolution of the state. The difference of these two schemes is illustrated in Figure 1.

For a nonlinear land surface hydrological processes model, the evolution of the states from time $t_k$ to time $t_{k+1}$ is given by

$$x_{k+1} = f_k(x_k, μ_k) + w_k$$  \(1\)
where $x_k$, a model state vector at time $t_k$; $f_k$ is a nonlinear function describing the evolution of the states from time $t_k$ to time $t_{k+1}$; $w_k$ is the model error vector, which represents errors in the model forcing data, initial conditions and parameters and model physics with error covariance $Q_k = (ww^T)$.

The observations are assumed to be related to the system states by the equations

$$y_k = h_k(x_k) + \delta_k,$$

(2)

where $h_k$ is a nonlinear transformation function. The observational errors, $\delta_k$ are assumed to be unbiased, serially uncorrelated, Gaussian random vectors with covariance matrices $R_k = (\delta \delta^T)$.

The data assimilation scheme consists of a series of forecasting (background) and update (analysis) steps. During the forecast step, the state error covariance is estimated, during the update step, the estimate of the state is updated based on the observation and the state error covariance. Updates to these forecast state and covariance values are made periodically when observations become available, with the correction being the weighted difference of the observation and model predicted observation. The forecasting and update steps are described as:

$$x_k^b = f_{k-1}(x_{k-1}, u_{k-1})$$

$$x_k^a = x_k^b + K_k [h_k(x_k^b) - y_k],$$

(3)

where $x_k^b$ is the predicted state of the system at the time $t_k$, known as the background states. $x_k^a$, is known as the analysis state. The innovation vector, referring to the predicted and measured observation vector difference, $[h_k(x_k^b) - y_k]$, is used to make a correction to the model predicted system state vector to improve state
estimates. The matrix $K$, known as the gain matrix, must be chosen to ensure that the analysis states converge to the true states of the system over time.

Sequential and variational methods use different approaches to choose the gain matrix, $K$ to minimize the mismatch between the model and the observations. The optimal analysis is given by the maximum likelihood a priori Bayesian estimate of the system states. This is equivalent to minimizing the square error between the predicted observations by the model and the measured observations, weighted by the inverse of the covariance matrices, over the assimilation window with the cost function (Nichols, 2002),

$$ J = \frac{1}{2} (x_0 - x_0^b)^T B_0^{-1} (x_0 - x_0^b) + \frac{1}{2} \sum_{k=0}^{N-1} [h_k(x_k) - y_k]^T R_k^{-1} [h_k(x_k) - y_k] $$  (4)

where subscripts $0$ and $k$ indicate the system initial state and the system state at time $t_k$, respectively. $B_0 = E[(x - x_0^b)(x - x_0^b)^T]$ is the error covariance of the initial state, $x_0^b$ is the background/model initialization state. This problem can be solved directly to give a sequential assimilation (dynamic observer) scheme, or it can be solved indirectly to give a four-dimensional variational assimilation (direct observer) scheme.

The sequential data assimilation method derives the optimal sequential solution for the nonlinear hydrologic processes system by minimizing the cost function with the background states as the prior/initial state. The optimal solution is given by

$$ x_k^a = x_k^b + K_k [h_k(x_k^b) - y_k], $$  (5)

where $K_k = B_k H_k^T (H_k B_k H_k^T + R_k)^{-1}$ is the Kalman gain with $H_k = \frac{\partial h_k}{\partial x_k}^b$ (the matrix $H_k$ is the Jacobian of the function $h_k$ with respect to $x$ evaluated at $x_k^b$), $B_k = E[(x - x_k^b)(x - x_k^b)^T]$ is the error covariance matrix for the background states $x_k^b$ at time step $t_k$. Note that $K_k$ is based on the relative magnitudes of the model and observation error covariances (Gelb, 1974).

A linear system has an exact optimal solution; this scheme is called Kalman filter (Kalman, 1960). But for nonlinear systems, the calculation of derivatives of linearized equations (Jacobian matrix) is required to propagate the error covariance to approximate the nonlinearities of the hydrologic model. This method is often called extended Kalman filter.

However, the derivatives of the nonlinear land surface model used in the extended Kalman filter sometimes are impossible, which hinders its use with conventional land surface model. As an alternative to the extended Kalman filter for nonlinear problems, the ensemble Kalman filter was developed (Evensen, 1994; Houtekamer and Mitchell, 1998), the model covariances are estimated using a Monte Carlo approach to produce an ensemble of model trajectories. The advantages of ensemble Kalman filter are that it can be used in any forward model and the model does not need to be differentiable. It also offers great flexibility in the specification of model error. The disadvantage is that estimates are conditioned on past measurements only.

The variational data assimilation method finds the model initial conditions that minimize a misfit between model results and observations for the whole analysis
period. This approach calculates the derivatives of the cost function $J$ in equation 4 with respect to each of the initial vector values, known as the adjoint using the Lagrange multiplier. The analysis period is usually much longer than the model time step (Reichle and McLaughlin, 2001; Reichle et al., 2001; Margulis and Entekhabi, 2003). The advantage of variational techniques is that all observations are used in a match mode to obtain the best states. However, variational techniques require that the system is differentiable.

Data assimilation offers several advantages over traditional methods for retrieving soil moisture from microwave satellite observations and model estimates. First, it provides a framework to best use both model predictions and observations. If the observations are unreliable, the model correction will be very small, and if the observations are reliable the model correction will be such that the model estimate becomes very close to the observed value. Good sets of remote sensing data are critical to ensure improved estimates of soil moisture. Second, the assimilation procedure estimates not only soil moisture but related energy and mass fluxes. Third, estimates of soil moisture are provided not only near the surface but throughout the soil column. Fourth, it provides continued soil moisture estimates in time and space rather than updates soil moisture at the specific time and location when/where observations are available. Fifth, the data assimilation procedure can generate estimates at finer resolution than the microwave observations. Sixth, the data assimilation procedure provides information on both the accuracy and uncertainty of soil moisture estimates (Reichle and McLaughlin, 2001; Reichle et al., 2001; Margulis et al., 2002; Reichle, McLaughlin, et al., 2002). It offers the advantage to integrate various amounts of satellite observations and model predictions for improved estimates of soil moisture states at fine spatial and temporal scales. The following section uses an example to demonstrate some of the advantages listed here.

SOIL MOISTURE ASSIMILATION—A CASE STUDY

Seasonal climate model prediction accuracy is currently limited due to poor soil moisture state initialization. However, initial soil moisture state prediction accuracy can potentially be enhanced by the assimilation of remotely sensed near-surface soil moisture data in off-line simulation. The following uses the current NASA–Global Modeling and Assimilation Office (GMAO)’s land initialization activities as an example to demonstrate the use of Kalman filter based soil moisture data assimilation for improved soil moisture initialization for seasonal to interannual climate prediction.

Data Assimilation Algorithm Development

NASA land initialization activities started with developing innovative data assimilation algorithms to merge satellite data and model predictions. To this end, Walker and Houser (2001) included an extended Kalman filter surface soil moisture data assimilation strategy in the GMAO’s catchment-based land surface model (Koster et al., 2000) with the simplified assumption that errors in different catchments are
uncorrelated. A one-dimensional Kalman filter was used to update the soil moisture prognostic variables of the CLSM because of its computational efficiency.

This assimilation system was refined using a synthetic twin experiment. Walker and Houser (2001) found that the unique catchment model physics are well suited for surface soil moisture assimilation, as the dominant prognostic moisture state variable (catchment deficit) has a significant correlation with surface soil moisture content except in very deep soils. Traditional land surface models generally have a vertical layering structure whose correlation is comparatively modest. Their study has shown that by assimilating near surface soil moisture observations, errors in forecast soil moisture profiles as a result of poor initialization may be removed and the resulting predictions of runoff and evapotranspiration improved.

The ensemble Kalman filter has also been implemented in the catchment model (Reichle, Walker, et al., 2002), being simply an alternative methodology for propagating the state covariance matrix that does not require model linearization. Reichle, Walker, et al. (2002) applied the ensemble Kalman filter to the soil moisture estimation problem and found it performs well. Reichle, Walker, et al. (2002) evaluated the relative benefits of the ensemble and extended Kalman filters using synthetic soil moisture observations and they found that ensemble approach is appealing because it does not require the calculation of derivatives of linearized equations (Jacobian matrix) to propagate the error covariance to approximate the nonlineairities of the hydrologic model and the measurement process, that are required in the extended Kalman filter approaches. The error covariance is propagated in the ensemble scheme through a finite ensemble of model trajectories. Furthermore, the ensemble scheme allows nonstationary spatially and temporally correlated representations of the model input and measurement errors and it allows for a wider range of model errors. It is extremely computationally expensive to consider the horizontal correlation in the model using the extended scheme. Small ensemble members are required for computational efficiency in the ensemble scheme to match the performance of the extended Kalman filter.

However, the performance of both filters is dependent on the assumption that model predictions and observations are unbiased, and that reliable estimates of model and observation error are available, which are difficult to estimate without an estimate of the truth for comparison. Therefore, characterizing the model and satellite observation error becomes an important part of assimilation studies that use real satellite data (e.g., Dong et al., 2005; Ni-Meister et al., 2005).

Assimilation of SMMR Data into CLSM

Kalman filter based soil moisture data assimilation strategy in the GMAO's catchment model was also used to assimilate SMMR-derived real surface soil moisture data (Owe et al., 2001; Fig. 2) into the CLSM for the period of 1979–1987 as one of the first studies to demonstrate the assimilation of real satellite data at a large continental scale over a long period. Two types of studies were conducted to assimilate SMMR data into the GMAO's catchment model. One type is conducted by Reichle and Koster (2005) and the other is by Ni-Meister et al. (2005, 2006). In the study by Reichle and Koster (2005), (1) the satellite-based soil moisture contents in their
study were scaled to match the model’s climatology before assimilation, (2) their model error was generated based on calibrated model error parameters (Reichle, Walker, et al., 2002) and not on realistic model error values, and (3) their evaluation was conducted by comparing the correlations between the assimilated and observed data and between the modeled and observed data. The logic behind this type of study is that satellite observed and modeled soil moistures have different physical meanings and the model will not benefit much by direct assimilating satellite observed soil moisture for initializing seasonal climate prediction models.

Ni-Meister et al. (2005, 2006) uses reliable model and satellite observation error information from a ground observation comparison. Ni-Meister et al. (2005) characterizes the model and SMMR errors using the Eurasian in-situ soil moisture observation network, which is the most extensive soil moisture data set available during the SMMR time period. Moreover, the ensemble Kalman filter soil moisture improvement performance is evaluated using the Eurasian station soil moisture observation network. The following describes more details of the work by Ni-Meister et al. (2006).

Historical Eurasian soil moisture observations archived in the Soil Moisture Data Bank (SMDB; Robock et al., 2000) were used to evaluate the data assimilation results presented in this study. The SMDB covers large areas including 43 Chinese, 36 Mongolian, and 130 Russian soil moisture monitoring stations over long time periods. Figure 3 shows a map of the Eurasian station measurement network and the corresponding catchments used in the evaluation.
The soil moisture estimate improvement through data assimilation was assessed as the difference between the mean absolute error in model predictions with and without assimilation, for both the surface and root zone soil moisture for each of the four seasons. The Kalman filter updates the system states based on the relative magnitudes of the model and satellite observation errors. In this study, both the predicted model and satellite observation errors are allowed to remain smaller than the actual errors, but their ratio is constrained to the actual ratio. Two types of errors were used. One is the matched error ratio, which agrees more closely with the actual error ratio. This ratio used suggests that the model errors are, on average, slightly smaller than the SMMR error. The other is the unmatched error ratio, which suggests a much smaller predicted mean model error than SMMR error, meaning that the data assimilation scheme would trust the model more than the SMMR observation. Figure 4 shows the summertime assimilated soil moisture improvement for 1979–1987 for all stations. Blue areas indicate a reduction in soil moisture estimate error through data assimilation while orange areas indicate an increase in soil moisture estimate error due to the data assimilation. Near zero values (within a range of -1 and 1% v/v) in grey indicate a negligible impact from data assimilation on the soil moisture estimate. Figure 4 shows more blue than orange areas, particularly for root zone soil moisture for the assimilation with matched error, but is
Fig. 4. Assimilated summertime soil moisture improvement (absolute model error minus absolute assimilation error) in Eurasia for 1979–1987. (A) Surface SM: DA with unmatched error; (B) Surface SM: DA with matched error; (C) Rootzone_SM: DA with unmatched error; Rootzone_SM: DA with matched error.
Fig. 5. As for Figure 4, but only for cases where the SMMR error is less than the model error. (A) Surface_SM: DA with unmatched error; (B) Surface_SM: DA with matched error; (C) Rootzone_SM: DA with unmatched error; (D) Rootzone_SM: DA with matched error.
mixed with a lot of grey areas, indicating data assimilation with the matched error produces better results than those with the unmatched error.

Figure 5 is similar to Figure 4, except it only shows results for the stations that have a smaller climatological mean SMMR error compared to the model error. In contrast to the results for all stations, Figure 5 shows soil moisture improvement through data assimilation for most stations in Mongolia and Russia and a few stations in China, indicating when the observation errors are smaller than the model error, the data assimilation improves soil moisture estimation. Many stations in China have wet biased SMMR observations compared to in-situ soil moisture measurements. Moreover, many Chinese stations show a data assimilation improvement in root zone soil moisture with corresponding degradation in surface zone soil moisture. When the SMMR mean error is greater or equal to the model error, data assimilation typically leads to an increase in soil moisture bias, particularly in the surface soil moisture estimates. In this case, bias correction is required before the observations can be rigorously used in the ensemble Kalman filter, since one assumption of the Kalman filter is that both the model and satellite observation errors are unbiased.

To investigate the temporal response of soil moisture data assimilation, modeled soil moisture with and without assimilation for both matched and unmatched error cases are compared with satellite-derived and station-measured surface zone and root zone soil moisture for one China station (Zhumadian; Fig. 6). Figure 6 shows SMMR-derived surface soil moisture overestimates ground observations, leading to the overestimation of surface soil moisture after assimilation. However, the model underestimates the root zone soil moisture, and assimilation helps bring the underestimated soil moisture up to the observed values, resulting in a somewhat better agreement with the station observations.

This study is one of the first to demonstrate the potential using actual remote sensing data together with field observations to provide a better estimate of land surface soil moisture states to improve seasonal climate prediction. This study demonstrates that the Kalman filter corrects model-generated soil moisture toward the satellite observation, with the size of the correction dependent on the relative magnitudes of the satellite observation and model errors. It has been shown that with the correct model and observation errors, data assimilation provided more accurate root zone soil moisture estimates. This improvement in root zone soil moisture estimates has substantial implications for seasonal climate prediction improvement, since better knowledge in root zone soil moisture is more important than surface zone soil in terms of land surface-atmosphere interactions. Moreover, these results highlight the need for accurate model and observation error estimates. Analysis of the impact of SMMR error on the assimilation results showed that when the climatologic mean SMMR error was less than the model error, the data assimilation improved both surface zone and root zone soil moisture estimates, with the greatest improvement being in the root zone, which indicates the need of more accurate remotely sensed soil moisture observations to ensure data assimilation works. Additionally, most stations in China resulted in biased surface soil moisture estimates as a result of wet biased SMMR observations in that region. This indicates that any bias
must be removed before assimilation to ensure improvement in soil moisture estimates.

CONCLUSION AND DISCUSSION

New remote sensing technologies and spatially distributed models offer great potential for improving our understanding of hydrologic processes, especially at large scales in the coming decades. The example presented in this study has demonstrated that a Kalman filter based assimilation scheme is able to integrate surface soil moisture observations into land surface models and to provide a more accurate soil moisture estimate particularly at the root zone than modeling alone.

There is, however, much work to be done before soil moisture data assimilation algorithms can be used operationally. Currently available global remotely sensed soil moisture datasets themselves exist larger differences even in their multiyear means and temporal variability and are biased compared to in-situ soil moisture measurements (Reichle and Koster, 2007). The study presented above shows that larger errors exist in SMMR data. Even with the newest X-band (10.7 GHz) passive...
microwave brightness temperature observations by the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSRE), the retrieved soil moisture data show different multiyear means and temporal variability from SMMR-derived soil moisture data and also different from in-situ soil moisture observations (Reichle and Koster, 2007). The differences are likely caused by different retrieval algorithms and a different physical meaning of remotely sensed soil moisture as compared to in-situ measurements, as discussed in Ni-Meister et al. (2005). Only when good remote sensing soil moisture data are available, it will be possible to obtain dramatic improvement of soil moisture for large areas and consistent soil moisture estimation through data assimilation (Wagner et al., 2007). Walker and Houser (2004) found that near-surface soil moisture observations must have an accuracy better than 0.05 m$^3$/m$^3$ to positively impact soil moisture forecasts. More investigation is needed to better characterize the model and observation errors and develop better remotely sensed soil moisture datasets at continental and global scale. The future mission of European Space Agency’s Soil Moisture and Ocean Salinity (SMOS) will provide global L band observations from which a better global soil moisture data set can be obtained.

Further research is necessary to advance on current data assimilation methods to make the best use of current remote sensing data. One approach is through bias correction during data assimilation. More work is needed to develop a persistent bias model to ensure the efficiency of bias correction (De Lannoy et al., 2007). Another modification would be directly assimilating microwave brightness temperature into a coupled radiative transfer and land surface process model (Dunne and Entekhabi, 2005, 2006). The radiative transfer model would allow to calculate the brightness temperature. This approach would provide a consistent way to estimate soil moisture by assimilating brightness temperature. This approach would also allow to assimilate multisource remote sensing data to constrain soil moisture estimates. Operational application of a soil moisture data assimilation scheme will also require a sufficiently sophisticated data assimilation algorithm to be able to assimilate multiple data sources that observe different processes acting at different scales.

Acknowledgments: This work was funded by NASA contract number NNG04GL08G. Special thanks go to anonymous reviewers for their time and effort to reviewing this manuscript and the special issue editors for their patient waiting for this manuscript.

REFERENCES


Houser, P. R., Shuttleworth, W. J., Gupta, H. V., Famiglietti, J. S., Syed, K. H., and Goodrich, D. C. (1998) Integration of soil moisture remote sensing and


