Hydrology and Earth System Sciences

Improving estimated soil moisture fields through assimilation of AMSR-E soil moisture retrievals with an ensemble Kalman filter and a mass conservation constraint

B. Li[1], **D. Toll**[2], **X. Zhan**[3], **and B. Cosgrove**[4]

[1]Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, and Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
[2]Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA
[3]NOAA NESDIS Center for Satellite Applications and Research, Camp Springs, MD, USA
[4]NOAA/NWS Office of Hydrologic Development, Silver Spring, MD, USA

Correspondence to: B. Li (bailing.li@nasa.gov)

Abstract. Model simulated soil moisture fields are often biased due to errors in input parameters and deficiencies in model physics. Satellite derived soil moisture estimates, if retrieved appropriately, represent the spatial mean of near surface soil moisture in a footprint area, and can be used to reduce bias of model estimates (at locations near the surface) through data assimilation techniques. While assimilating the retrievals can reduce bias, it can also destroy the mass balance enforced by the model governing equation because water is removed from or added to the soil by the assimilation algorithm. In addition, studies have shown that assimilation of surface observations can adversely impact soil moisture estimates in the lower soil layers due to imperfect model physics, even though the bias near the surface is decreased. In this study, an ensemble Kalman filter (EnKF) with a mass conservation updating scheme was developed to assimilate Advanced Microwave Scanning Radiometer (AMSR-E) soil moisture retrievals, as they are without any scaling or pre-processing, to improve the estimated soil moisture fields by the Noah land surface model. Assimilation results using the conventional and the mass conservation updating scheme in the Little Washita watershed of Oklahoma showed that, while both updating schemes reduced the bias in the shallow root zone, the mass conservation scheme provided better estimates in the deeper profile. The mass conservation scheme also yielded physically consistent estimates of fluxes and maintained the water budget. Impacts of model physics on the assimilation results are discussed.

1 Introduction

Soil moisture plays an important role in the energy and water exchange between the atmosphere and the land surface, as well as in agricultural applications and water resource management. Model simulated soil moisture fields are often biased due to uncertainties in model input parameters and model physics. The existence of model bias can be seen in several model inter-comparison studies which showed that model estimated soil moisture is significantly different from each other, even when identical forcing data are used (Mitchell et al., 2004; Wood et al., 1998). Recognizing the significant disparity between the models, Mitchell et al. (2004) concluded that there was a "stringent need for good absolute states of soil moisture". Reducing the bias in model estimated soil moisture fields has been shown to have a positive impact on other physical processes. Dirmeyer (2000) demonstrated that the rainfall patterns and the near surface air temperature can be improved by using a mean soil moisture data set derived from a global soil moisture data bank.

Satellite derived soil moisture retrievals represent the spatially averaged soil moisture in a footprint area (Njoku et al., 2003). If retrieved appropriately, they can be used to improve the spatial mean of modeled soil moisture fields, at finer spatial resolutions than the footprint of AMSR-E, as well as the temporal mean through continuous assimilations in time. While interest in assimilating satellite retrieved soil moisture estimates began more than a decade ago (Houser et al., 1998; Margulis et al., 2002; Walker et al., 2001), recent

studies have focused on using the anomaly information extracted from the sensor data by removing the mean of the observations priori to assimilation to improve model's anomaly detection (Bolten et al., 2008; Crow and Zhan, 2007; Draper et al., 2009; Reichle et al., 2007). While assimilation of anomalies does not directly address if models are unbiased which is required for optimal estimators (Kalnay, 2003), it preserves the water budget of forecasts.

An alternative to the offline bias-removal technique, as those used in the above studies, is to estimate the forecast bias online by adding a bias state in the filtering process (De Lannoy et al. 2007a,b; Keppenne et al., 2005). De Lannoy et al. (2007a,b) compared the performance of several online bias correction techniques with the standard EnKF using the CLM land model and profile soil moisture observations. Their results showed that the online bias correction techniques, on average, yielded slightly more reductions in root mean square error. One major obstacle for applying this approach in assimilating satellite retrieved soil moisture is that observations are only available at the surface which makes it very challenging to estimate the bias state in the deeper profile. When bias is not correctly estimated, assimilation may lead to unbalanced water budget as the assimilation may change the mean of estimated soil moisture fields. Lack of water budget closure is a weak point for many data assimilation systems as pointed out by Pan and Wood (2006), perhaps more so for land surface models whose major goal is to partition the total water budget, precipitation, into different physical processes such as evapo-transpiration (ET) and runoff.

When sensor data are less biased (relative to the truth) than model estimates, they can be used to reduce uncertainty in model estimates. Recognizing this potential, studies have been conducted to assimilate actual values of satellite data without using any bias correction techniques (Houser et al., 1998; Margulis et al., 2002; Ni-Meister et al., 2006; Walker et al., 2001). While the bias reduction at the surface was achieved in these studies, improvements in the deeper soil layers did not always occur. Houser et al. (1998) and Walker et al. (2001) showed that assimilation of surface observations actually adversely impacted the soil moisture state in the lower soil layers. The representation of the hydrological condition in the lower soil zone is often a weak point in land surface models due to lack of knowledge and observations. If model physics is flawed, it may adversely impact the outcome of data assimilation, especially for an EnKF which relies on model physics to calculate the Kalman gain matrix dynamically (Keppenne et al., 2000).

The objective of this study is to assimilate AMSR-E soil moisture retrievals as they are without any preprocessing or scaling into the Noah land surface model to improve the simulated soil moisture fields using an EnKF. To overcome the potential bias issue associated with both the model and the AMSR-E retrieval, a mass conservation updating scheme was developed to allow the upper soil layers updated using the conventional EnKF while the lower layers are updated with an equation that conserves mass of the forecast. This study differs from those recent studies on AMSR-E data assimilation (Bolten et al., 2008; Crow and Zhan, 2007; Draper et al., 2009; Reichle et al., 2007) in that AMSR-E retrievals were not pre-processed priori to assimilation while, in the other studies, the mean of retrievals were removed through matching the cumulative distribution functions (Drusch et al., 2005; Reichle and Koster, 2004). By assimilating the actual value of AMSR-E soil moisture, the objective of this study is to reduce forecast bias and estimation errors. In Sect. 2, the experiment site, data, and the Noah model are briefly described. Details of the mass conservation assimilation method along with the conventional EnKF are described in Sect. 3. Assimilation results including all land surface fluxes and water budgets are presented in Sect. 4. Impacts of model physics on model simulation and assimilation results and the limitations of the mass conservation scheme are discussed in Sect. 5.

2   Experiment site, data and model

2.1   Study area and ground validation data

The Little Washita watershed, located in southwestern Oklahoma, was chosen as the study site primarily for its abundance of in situ soil moisture measurements. With an area of 611 square kilometers, the watershed is one of the two Micronet sites maintained by the U.S.D.A. Agriculture Research Service (ARS) for hydrological and meteorological observations (http://ars.mesonet.org). Figure 1 shows the watershed boundary and the locations of the ARS stations. At each station, hourly soil moisture and temperature measurements are taken at 5, 25 and 45 cm depths below the surface, in addition to surface measurements such as precipitation. Figure 1 also shows the only Soil Climate Analysis Network (SCAN) station located within the watershed (Schaefer et al., 2007; http://www.wcc.nrcs.usda.gov/scan). The SCAN site complements the ARS stations in that it provides soil moisture measurements at the 100 cm depth which were used to verify simulated soil moisture in the deeper soil profile. Daily stream flow data recorded at the watershed by USGS (see Fig. 1 for the location of stream site 07327550) were used for validating model predicted runoff. Latent heat measurements from the Southern Great Plain (SGP) main station (http://public.ornl.gov/ameriflux) were used for validating the simulated latent heat. Although SGP, which is approximately 200 km north of Little Washita, is not located near the watershed, it is the nearest site where flux data are publicly available.

2.2   AMSR-E retrievals

The AMSR-E soil moisture product produced by the NOAA's National Environmental Satellite, Data and
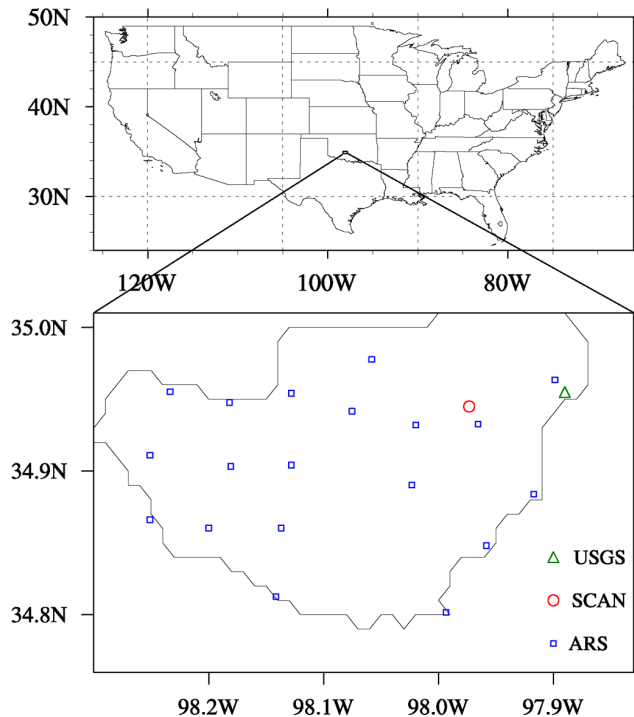
Fig. 1. The Little Washita watershed and locations of ARS, SCAN and USGS stations.

Information Service (NESDIS) was used in this study. The soil moisture retrievals, based on the X-band brightness temperature measurements, were obtained through the inversion of the Single Channel Retrieval algorithm with the MODIS vegetation water content as an auxiliary variable (Jackson, 1993; Zhan et al., 2008). Zhan et al. (2008) showed that this version of AMSR-E generally has larger dynamic ranges than the official AMSR-E product (Njoku et al., 2003) even though both products show strong temporal correlations (Crow and Zhan, 2007). The spatial resolution of AMSR-E retrievals is about 25 by 25 km after re-sampling from its original sensor data (Njoku et al., 2003). The experiment site (the rectangular area shown in Fig. 1) is covered by portions of 5 to 6 AMSR-E retrievals at any observation time. On average, there are $1 \sim 2$ retrievals per day at any given location and both ascending and descending data were assimilated at the retrieval time, except in areas of dense vegetation or frozen grounds.

The sensing depth of the AMSR instrument is believed to be about 1–2 cm from the surface for the frequency range of AMSR-E (Njoku et al., 2003). This depth is shallower than the ARS surface measurement (5 cm) and the center of Noah's surface layer. However, without reliable methods to extrapolate the AMSR-E estimates, it was assumed that the AMSR-E soil moisture retrieval is representative of soil moisture in the top 5 cm soil and therefore was assimilated into Noah's top layer directly. This approximation could bring bias into the retrievals used for data assimilation.

2.3 The Noah land surface model, forcing and input parameters

The Noah land surface model (version 2.7.1) is used operationally at the NOAA's National Centers for Environmental Prediction for coupled weather and climate modeling. The soil moisture simulation in Noah is based on the one-dimensional Richards equation (Chen et al., 1996; Ek et al., 2003):

$$\frac{\partial \theta(z, t)}{\partial t} = \frac{\partial}{\partial z} \left(D(\theta) \frac{\partial \theta(z, t)}{\partial z} + K(\theta) \right) + P - R - E \quad (1)$$

where θ is the soil moisture content; K is the hydraulic conductivity; D is the water diffusivity, which is defined as $K \partial \psi / \partial \theta$, where ψ is the matric potential; P is the precipitation; R is the surface runoff; E is the ET; z is the vertical dimension with upward as the positive direction; t is the time.

Following the operational version of Noah (Ek et al., 2003), four soil layers with thicknesses of 10, 30, 60 and 100 cm were used in this experiment. The top two layers, a thin surface layer and the shallow root zone, generally show stronger and faster interactions with the atmospheric forcing. The third and the fourth layers represent the deeper root zone and water storage, respectively.

Equation (1) is solved with the following boundary condition at the 200 cm lower boundary:

$$q|_{z=200\text{cm}} = -K \quad (2)$$

where q is the subsurface runoff or base flow. Equation (2) is also referred to as the free drainage condition, meaning gravity is the only force pushing water downward (so the negative sign) and no upward diffusive movement is allowed across the lower boundary (Jury et al., 1991). The use of free drainage is very common in land surface models because it does not require any knowledge about the soil moisture state or flux in the subsurface which is impossible to obtain for large-scale modeling.

Noah uses the Campbell (1974) model to describe the nonlinear relationship between the conductivity and soil moisture:

$$K = K_{\text{s}} \left(\frac{\theta}{\theta_{\text{s}}} \right)^{2b+3} \quad (3)$$

where K_{s} is the saturated conductivity; θ_{s} is the saturated water content; b is a fitting parameter. The US general soil texture classes (STATSGO) and a look-up table, based on a unified soil hydraulic parameter set (Mitchell et al., 2004), were used to provide soil hydraulic parameters needed for solving Eq. (1). Hydraulic conductivity usually exhibits the property of a log-normal distribution and is positive skewed (Cosby et al., 1984). As a result, the subsurface runoff calculated using Eq. (2) is non-Gaussian which can lead to unrealistic ensemble mean values in an EnKF when larger ensemble spreads occurred in the lowest soil layer (De Lannoy et al., 2007a; Ryu et al., 2009).

Table 1. Annual ARS and GDAS precipitation (mm) at Little Washita in 2006 and 2007.

	2006	2007
ARS	690	1259
GDAS	959	1096

Model simulations were carried out in the NASA's Land Information System (LIS, version 5.0) which is a software interface between various land surface models and forcing/static parameter fields (Kumar et al., 2006). LIS is also equipped with a one-dimensional EnKF (Kumar et al., 2008) which will be described in the next section. The Noah model was integrated on a 0.01° grid so that spatial variability was well represented in model estimated soil moisture and flux at the watershed. To avoid the model spin up issue (Cosgrove et al., 2003a; Rodell et al., 2005;), the initial soil moisture conditions used were extracted from the output of Global Land Data Assimilation (GLDAS)/Noah model which have been continuously integrated since 1979 (Rodell et al., 2004).

Model simulations were driven by forcing data (including precipitation, radiation, wind, and temperature fields) from the NOAA/NCEP Global Data Assimilation System (GDAS, Derber et al., 1991; Rodell et al., 2004). Basin-averaged monthly GDAS and ARS precipitation for the simulation period (2006-2007) are compared in Fig. 2 and their annual precipitation amounts are listed in Table 1. Despite some underestimated and overestimated events in the GDAS forcing data, both data sets showed that 2006 is a drier year than 2007.

3   Data assimilation methods

In this section, the conventional EnKF and the mass conservation EnKF scheme are described. EnKF is a widely used technique for assimilating observations into numerical models to improve model estimates (e.g., Crow and Wood, 2003; Evensen and van Leeuwen, 1996; Keppenne et al., 2000; Pan and Wood, 2006; Reichle et al., 2007). EnKF is especially suited for a non-linear system since the error covariance, used for passing observation information from data-rich zones to data-poor zones, is calculated through an ensemble of model states (Evensen and van Leeuwen, 1996). An EnKF usually consists of two steps: the forecast step where an ensemble of model forecasts are obtained and propagated forward in time with perturbations added for forcing and state variables, and the update step where an analysis is obtained using an update equation when observations become available. The model forecast can be expressed as:

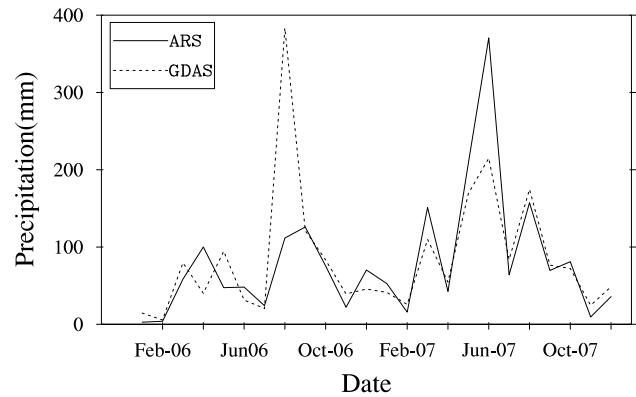$$X_t^{\mathrm{f}} = M\left(X_{t-1}^{\mathrm{a}}, F_t, U\right) \tag{4}$$

Fig. 2. Basin averaged monthly ARS and GDAS precipitation.

where X is the vector containing the four state variables of soil moisture of Noah; M represents the Noah model; F represents all the forcing fields such as precipitation and radiation; U represents static input parameters such as soil hydraulic parameters; and t indicates the time step. The superscript ($^{\mathrm{f}}$) indicates results for the forecast and ($^{\mathrm{a}}$) for the analysis. Although not explicitly noted, Eq. (4) and the following update equations are valid for each ensemble member. The conventional EnKF updating scheme for obtaining the analysis can be written as:

$$X_t^{\mathrm{a}} = X_t^{\mathrm{f}} + \underline{\mathbf{K}}\left(v_t - H\,X_t^{\mathrm{f}}\right) \tag{5}$$

where $\underline{\mathbf{K}}$ is the Kalman gain matrix computed from the ensemble statistics of the model simulated soil moisture fields (Keppenne, 2000); v is the observation (AMSR-E retrievals in this study); H is the observation operator that relates the observation to the model state and is [1, 0, 0, 0] in this study because the observation is the same type as the model state and is only available at the surface layer. The AMSR-E retrievals were used without downscaling, i.e., all model grid points within the footprint of the satellite were given the same retrieved soil moisture value, which is equivalent to a priori partition of the large scale retrieval to the finer scale with the same value assigned to each grid cell. This approach allows for direct and efficient assimilation of satellite retrievals using the current infrastructure of LIS. It is justified for the purpose of this study which is to improve the spatial mean of simulated soil moisture fields and will not be an issue for larger scale simulations where model resolutions can be made to match that of AMSR-E. Since observations are only available at the surface, the innovation, $(v - HX^{\mathrm{f}})$, is a scalar. The $\underline{\mathbf{K}}$ matrix propagates the innovation downwards to obtain the increment, $\underline{\mathbf{K}}(v - HX^{\mathrm{f}})$, for all lower layers.

When both the model and the observation are unbiased, the mean of the innovation (and increments) is zero. When either or both of them are biased, the analysis (X^{a}) obtained through Eq. (5) may not possess the same mean as the forecast (X^{f}) which is enforced by the mass balance Richards

Table 2. Additive perturbation errors (same unit as volumetric water content) given to the four Noah soil moisture variables (θ_1, θ_2, θ_3, and θ_4) and their cross-correlations (the last four columns).

Variable	Standard deviation	Correlation in time	Cross-correlation with θ_1	Cross-correlation with θ_2	Cross-correlation with θ_3	Cross-correlation with θ_4
θ_1	0.002	12 h	1.0	0.6	0.4	0.1
θ_2	5.0e-4	12 h	0.6	1.0	0.5	0.2
θ_3	6.0e-5	12 h	0.4	0.5	1.0	0.4
θ_4	6.0e-6	12 h	0.1	0.2	0.4	1.0

equation. The CDF matching technique used by previous studies (e.g., Reichle et al., 2007) renders the mean of the retrievals equal to that of the model and therefore preserves the mean of the forecast. The tradeoff of this scaling approach is that it discards the mean value of retrievals which may be useful in improving the mean of model estimates.

In order to assimilate the actual value of retrievals which may not have the same mean as model estimates, the loss of water mass (relative to the forecast) needs to be handled in the updating scheme. Pan and Wood (2006) used a two-step constrained Kalman filter to redistribute the mass imbalance caused by assimilating multiple types of observations (ET, stream flow and soil moisture). When only the surface soil moisture observation is available for assimilation, the redistribution of mass imbalance can be carried out within the four soil layers. Specifically, while the top two layers are updated using Eq. (5), a different updating scheme can be used for the lower two layers:

$$y_t^{\mathrm{a}} = y_t^{\mathrm{f}} - \sum_{k=1}^{2} (\Delta C_k \, d_k)/(d_3 + d_4) \qquad (6)$$

where y represents the soil moisture at layer 3 or 4; d represents the thickness of each soil layer; subscripts $(_k)$, $(_3)$, and $(_4)$ indicate the soil layer; ΔC_k represents the increment, i.e., water (in soil moisture content) lost or gained, for the top two layers when they are updated using Eq. (5). Equation (6) redistributes the mass imbalance (amount of water) incurred in updating the top two layers to the lower layers and therefore, guarantees that the total water storage remains the same for each ensemble member after the ensemble update. The division of layer thicknesses in Eq. (6) is to convert the amount of water to volumetric soil moisture content to match the unit of the state variable. Equation (6) is performed each time when the upper two layers are updated so that the column water of the analysis remains the same as the forecast (but with a different soil moisture profile). By maintaining the water storage within a soil column, the mass conservation scheme also preserves the long-term water budget of the control run (without any data assimilation) since ET and runoff are calculated based on the column water storage and perturbations added to the forcing and state variables are unbiased.

Because of the enforcement of mass conservation of the control run, this scheme (Eq. 5 for top two layers and Eq. 6 for the two lower layers) is referred to as the mass conservation updating scheme. Note that no assumption was made about the observation and the model, both of which can be biased, in deriving Eq. (6).

In addition to preserving mass, Eq. (6) avoids updating the lower layers with the conventional EnKF which has been shown to yield undesired increments due to inappropriate model physics (Houser et al., 1998; Walker et al., 2001). Preserving water mass does not necessarily lead to improved soil moisture estimates in the lower layers, but Eq. (6) keeps the increments small due to the larger thickness of lower two layers relative to the upper two layers, and thus minimizes any potential adverse impacts.

The ensemble of model states was generated by adding zero-mean perturbations (errors) to the forcing fields and state variables to represent random errors in them. Following Reichle et al. (2007), precipitation, long and short wave radiation fields which have the largest impact on soil moisture were perturbed using the same parameters given by Reichle et al. (2007) as the same forcing data were used in both studies. Perturbations for precipitation and shortwave radiation were assumed to be multiplicative and additive for longwave radiation. The perturbation frequency for these forcing fields was 5.5 h.

Perturbations were also added to soil moisture variables to account for errors in the input parameters such as soil hydraulic conductivity and model physics using parameters listed in Table 2. Smaller perturbations (in volumetric soil moisture content) were given to the lower two layers because of their larger thickness and the fact that perturbations added in the top two layers can travel downward through the dynamics of the Richards equation. In addition, the issue with the calculation of ensemble mean base flow due to the skewness of the hydraulic conductivity function (De Lannoy et al., 2007a; Ryu et al., 2009) also requires smaller perturbations in the lower layers to ensure physically consistent ensemble runoff. All soil moisture variables were assumed to have additive zero-mean Gaussian errors with vertical cross-correlations among four layers given in Table 2. The perturbation frequency for soil moisture was 24 h. Noah Soil
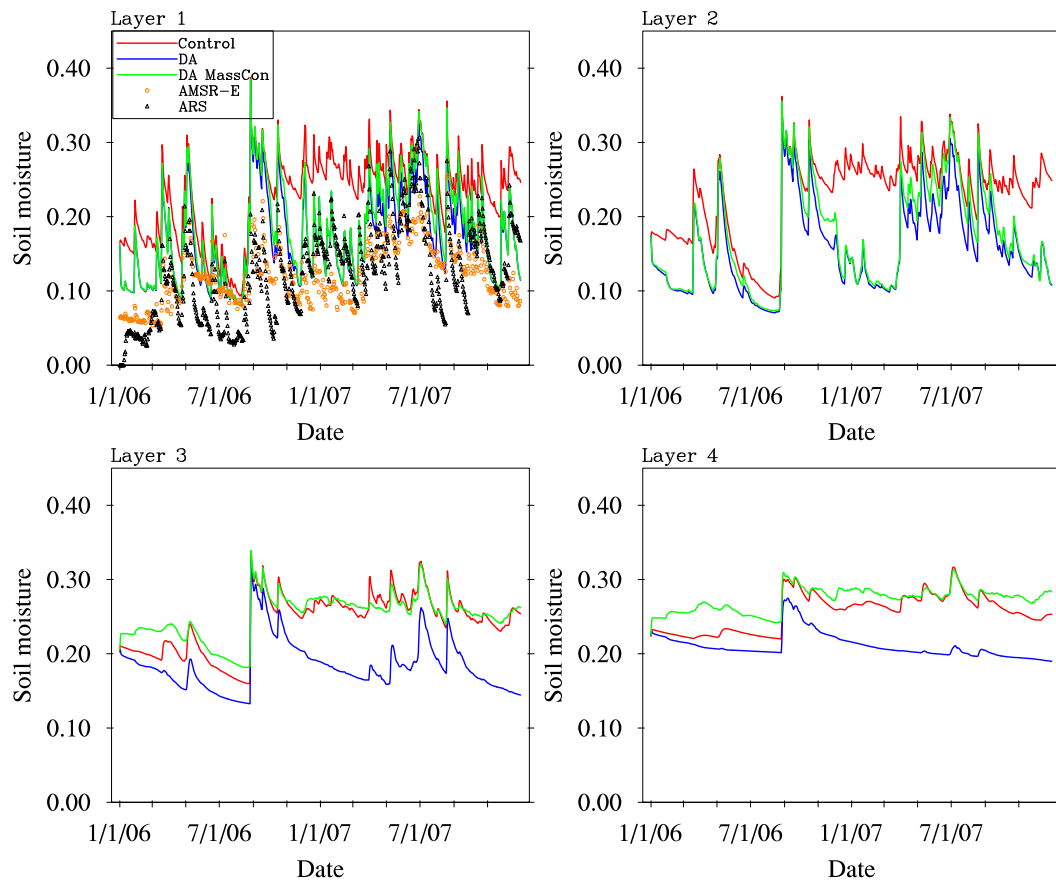
Fig. 3. Time series of basin averaged daily soil moisture ($cm^3 cm^{-3}$) from Control, DA, and DA MassCon for Noah soil layers 1 to 4. Simulated soil moisture at layer 1 is compared to basin averaged ARS measurements at the 5 cm depth and the AMSR-E retrievals.

moisture moves very slowly in drier conditions, which is why the longer perturbation frequency was used to avoid ensemble bias. Despite all zero-mean perturbations, ensemble bias could still exist in the ensemble soil moisture field due to the nonlinear relationship among various processes and the strong influence of model physics. Parameters in Table 2 were chosen because they yielded unbiased ensemble (without data assimilation) soil moisture fields relative to a single member control run. The 0.03AMSR-E error (Njoku et al., 2003) was used in the filter to account for errors in the observation. The same filter parameters were used for both updating schemes.

4 Results

Three simulation runs were performed at the Little Washita watershed for the 2006 to 2007 period. The control run (Control), which represents the baseline performance of the Noah model, was driven by the GDAS forcing and all the parameter fields in their unperturbed states. The other two simulations featured assimilations of AMSR-E soil moisture

retrievals using the conventional (DA) and mass conservation (DA MassCon) updating schemes.

Given the objective of this study which is to improve the mean of analyzed soil moisture fields, basin averaged daily bias and root mean square errors (RMSE) were used to evaluate the assimilation results. All statistics were calculated with respect to the ground validation data described in Sect. 2.

4.1 Soil moisture

Figure 3 shows the comparison of soil moisture in the four Noah soil layers from the three simulations. The upper left panel also includes basin averaged AMSR-E soil moisture retrievals and ARS measurements at the 5 cm depth. Overall, the AMSR-E soil moisture compares well with ARS by capturing the seasonal change and the mean value of in situ measurements. The daily variation of AMSR-E is small due to the twice per day (maximum) retrieval interval. Control also captured the wetting and drying cycles of the surface soil moisture, exhibiting strong correlation with the ARS measurements. However, it consistently overestimated the surface soil moisture throughout the simulation period,

Table 3. Basin averaged bias and root mean square error (rmse) of daily simulated soil moisture at the 5, 25, 45 and 100 cm depths, latent heat (W m^{-2}) and AMSR-E soil moisture retrievals for the two-year period. Statistics were calculated with respect to daily values of ground measurements at ARS, SCAN and SGP. Numbers in bold represent the minimal bias or rmse among Control, DA and DA MassCon.

	Control		DA		DA MassCon		AMSR-E	
	bias	rmse	bias	rmse	bias	rmse	bias	rmse
soil moisture (5 cm)	0.11	0.11	**0.05**	**0.07**	0.06	0.08	−0.002	0.05
soil moisture (25 cm)	0.10	0.10	**0.03**	**0.06**	0.04	0.07	–	–
soil moisture (45 cm)	0.09	0.09	**0.02**	**0.05**	0.03	0.06	–	–
soil moisture (100 cm)	−0.10	0.11	−0.16	0.16	**−0.09**	**0.09**	–	–
latent heat	8.27	28.87	−5.76	28.80	**−0.33**	**28.38**	–	–

even in the period from December 2006 to June 2007 when GDAS underestimated the precipitation (see Fig. 2). The same overestimation was also observed (not shown) when the model was driven by the North America Land Data Assimilation System (NLDAS, Cosgrove et al., 2003b) forcing data which yielded nearly unbiased monthly precipitation estimates against ARS measurements (not shown). These results indicate that the bias at the surface was not initiated by errors in the precipitation forcing data. Figure 3 also shows that the overestimation by Noah was more severe in winter periods when ET and precipitation were low, which limits the likelihood that incorrect runoff and ET algorithms may have left excessive water at the surface. Flux results that will be discussed in Sect. 4.2 also do not show any negative bias. In a separated study (to be submitted), NLDAS/Noah was compared with SCAN soil moisture for the continental US and the similar overestimation was found in the western US. The likely cause for this persistent overestimation in such a large area may be the static parameters such as soil hydraulic conductivity. The vertical drainage of soil moisture in Noah is controlled by the nonlinear function of soil hydraulic conductivity as shown in Eqs. (1) and (3). The parameters in Eq. (3) were obtained through linear regressions (Cosby et al., 1984) which may not capture all the nonlinear behaviors of hydraulic conductivity. If the hydraulic conductivity value is lower than expected in the drier range of soil moisture, it would explain why Noah failed to drain soil moisture quickly in Little Washita and the western US.

DA and DA MassCon both reduced the overestimation of Control in the surface layer (reduction in bias and RMSE are listed in Table 3). The degree of correction is not uniform, especially in very wet conditions where the assimilation failed to nudge the soil moisture towards the AMSR-E retrievals. This is because the perturbation parameters for the filter had to be tuned to work with the driest condition in order to avoid ensemble bias. If overly perturbed to fit the need of wetter conditions, ensemble bias would appear in drier periods because some of the ensemble members would hit the lower bound of soil moisture (Reichle and Koster, 2002).

Figure 3 also shows that both updating schemes decreased soil moisture in layer 2. However, for layers 3 and 4, the two schemes acted differently. DA lowered the soil moisture in layers 3 and 4 as it did with the top two layers. DA MassCon increased the soil moisture in the lower layers because it captured the amount of water removed from the top two layers in the lower layers. As shown in Fig. 4 which compares the simulations with in-situ soil moisture measurements at various measuring depths, DA and DA MassCon both improved over Control at 25 and 45 cm by reducing bias and rmse (see Table 3). But only DA MassCon improved over Control at 100 cm with reduced bias and RMSE while DA further worsened the bias and RMSE of Control. Table 3 also lists the statistics for the AMSR-E retrievals which are nearly unbiased relative to in situ measurements, although the mass conservation scheme does not require retrievals to be unbiased.

The improvement made by DA MassCon at 100 cm may be debatable due to the co-existence of the overestimation in the upper layers and the underestimation of soil moisture in the lower layers, which is found to be true for the Noah model in the western US. Houser et al, (1998) also showed the similar model behavior with a different model. When the overestimation and underestimation do not occur concurrently, the mass conservation algorithm may not lead to improved soil moisture estimates in the lower soil profile, but it does not cause significant changes (relative to Control) to the lower soil moisture states, as seen in Fig. 3, because of the smaller increments given by Eq. (6).

The conventional updating scheme generated increments with the same sign for all layers that significantly decreased soil moisture in the lower profile, a result not supported by in situ measurements at 100 cm. For Noah, the fact all increments have the same sign is due to the free drainage condition which has to adjust the soil moisture in the lower layers according to changes in the upper layers in order to maintain the downward flow direction prescribed by Eq. (2). Negatively cross-correlated soil moisture perturbations between the upper layers and lower layers were also tested for the conventional EnKF (not shown). They did not change the sign
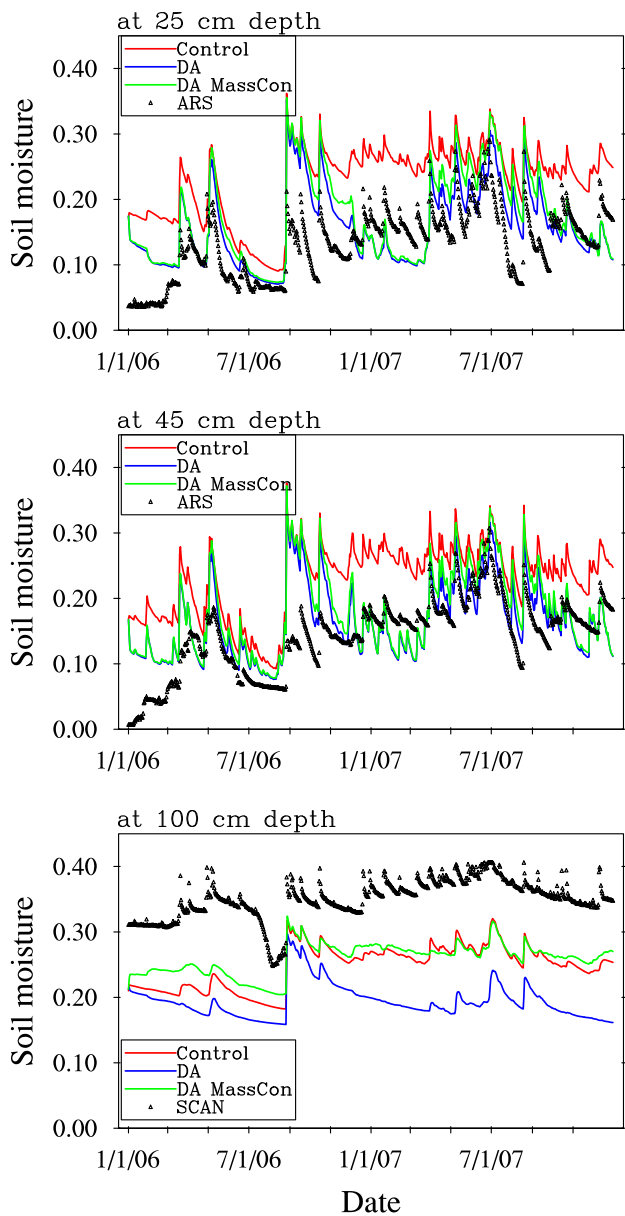
Fig. 4. Comparison of basin averaged daily soil moisture ($cm^3 cm^{-3}$) from Control, DA, and DA MassCon, interpolated at 25, 45 and 100 cm depths, with measurements from ARS stations and the SCAN site.

of the increments but slightly lowered their magnitudes, with the soil moisture estimates in the lower two layers slightly wetter than those shown in Figs. 3 and 4 but still much worse (drier) than Control when compared to in situ measurements. Cross-correlations of perturbations only partially influence the outcome of the increments which also strongly depend on model physics. As model physics largely determines the mean behavior of soil moisture, it is difficult for the zero-mean perturbations alone to overcome the large difference between DA and Control seen in Fig. 4 (lower panel).

Significantly increasing perturbations for soil moisture is not permitted because it will lead to ensemble bias in soil moisture and base flow.

Similar to surface overestimation, the underestimation by Control in the lower profile cannot be explained by any precipitation forcing errors. In fact, the underestimation is caused by the free drainage condition which drains excessive water away and prevents moisture moving up from below the land surface. The deficiency of the free drainage condition is why the surface overestimation did not occur at 100 cm. The underestimation has also been observed with other models which employ the same boundary condition (Zeng and Decker, 2009; Houser et al., 1998). This deficiency in model physics is why the conventional EnKF cannot obtain increments favorable for improvements in the lower profile. Presumably, the underestimation of soil moisture in the lower soil moisture profile could also be balanced out with base flow which would require either deeper soil moisture measurements or observations of base flow to create the innovation. With only the surface soil moisture observation available, the mass conservation scheme focuses on improving the soil moisture fields first and let the mass conservation to constrain flux estimates.

Figure 5 features the contour plot of the annual mean surface soil moisture at the watershed. Control revealed that 2006 is drier than 2007, confirming the earlier analysis regarding GDAS precipitation (Fig. 2 and Table 1). NOAA AMSR-E retrievals also captured the difference in annual precipitation, as DA and DA MassCon all show the wetter soil moisture condition in 2007. DA MassCon yielded slightly higher soil moisture estimates than DA because the former has a wetter lower soil profile (see Fig. 3) which pushed the surface soil moisture slightly higher via the capillary force. This is why DA achieved slightly better statistics than DA MassCon for the upper three observation levels shown in Table 3. However, for the root zone soil moisture (consisting of the upper three Noah soil layers), Fig. 6 shows that while DA MassCon and Control yielded wetter soil moisture conditions in 2007, DA soil moisture barely reflects this variation in annual precipitation, further confirming the failure of DA in updating the lower layers.

Figures 5 and 6 also show that the spatial variability of Control was generally preserved by the assimilation schemes even though the AMSR-E retrievals were assimilated directly without any spatial downscaling. Note that spatial variability of soil moisture may be lost slightly at the assimilation time, but it recovered quickly afterwards because of the high resolution soil and vegetation parameters.

4.2 Flux

One of the important roles of any land surface model is to simulate water and energy fluxes based on soil moisture fields. Improvements on soil moisture do not necessarily lead to improvements in the calculation of flux because of
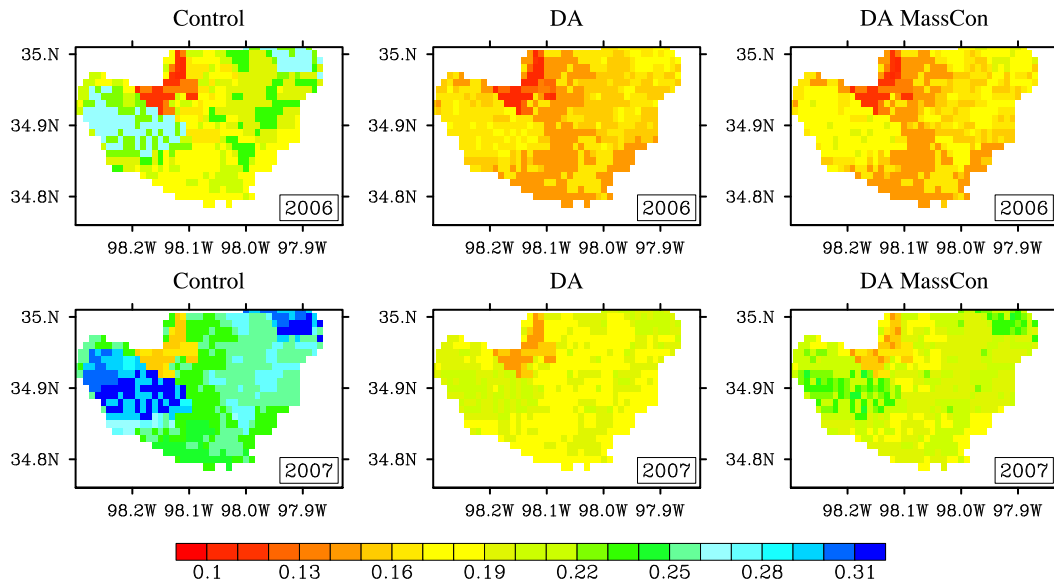
Fig. 5. Mean annual surface soil moisture ($cm^3 cm^{-3}$) from Control, DA and DA MassCon in 2006 and 2007.

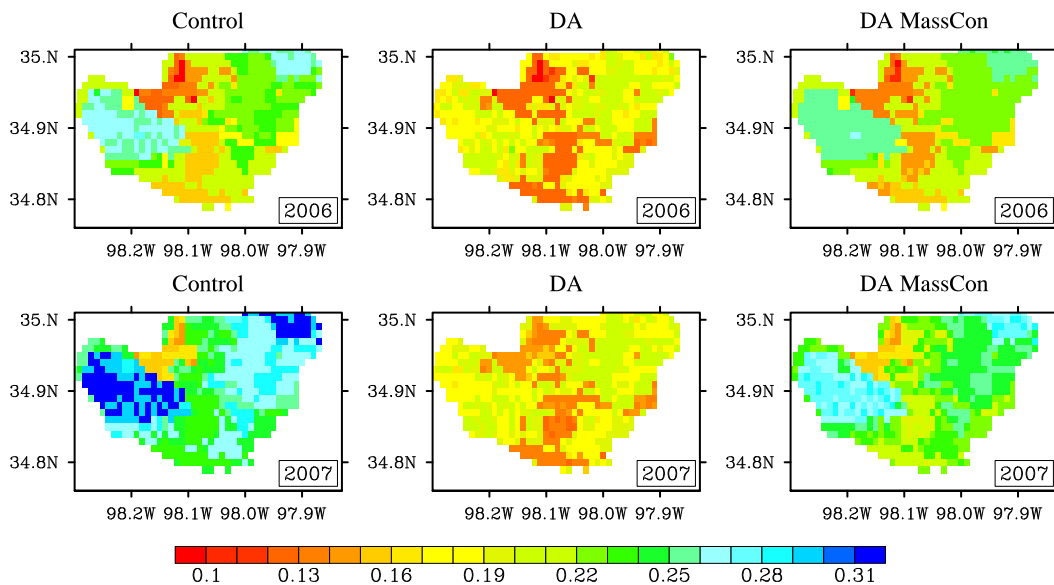**Fig. 6.** Mean root zone (upper 100 cm) soil moisture ($cm^3 cm^{-3}$) from Control, DA and DA MassCon in 2006 and 2007.

imperfect model physics and complex relationship among various processes. Therefore, it is important to examine all components of model estimates to prevent unexpected flux estimates.

Figure 7 shows the simulated latent heat fluxes in comparison with SGP observations. The differences among the three simulations are relatively small, with total ET for the two-year period estimated at 1362, 1013 and 1147 mm, for Control, DA and DA MassCon, respectively. Part of the reason is that the ET algorithm in Noah is more sensitive to the vegetation greenness fraction than soil moisture (Chen et al.,

1996). In addition, the watershed is mostly covered by vegetations with shallow root zone depths such as shrubs and grasses which do not strongly depend on the soil moisture state in the lower profile where DA differs from DA MassCon the most. Nevertheless, Table 3 shows that Control yielded the largest bias (positive) in latent heat estimation. DA reduced the bias but DA MassCon produced the smallest bias. Although the improvement by DA MassCon and DA should not be overstated given that the SGP is not located near the watershed, the impact of the different soil moisture fields on the latent heat estimation is demonstrated.
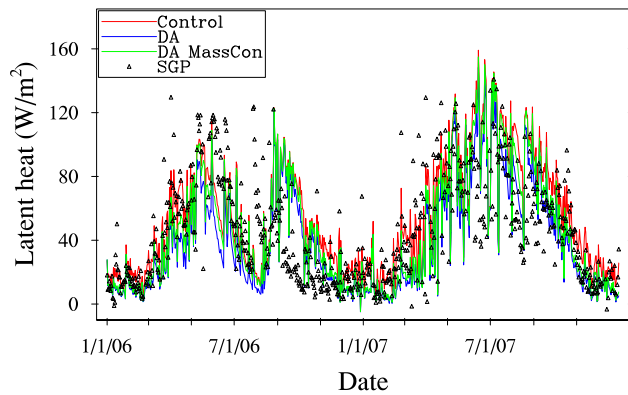
Fig. 7. Comparison of daily latent heat from Control, DA, and DA MassCon versus the SGP flux data. The daily latent heat estimates are averaged values from 06:00 a.m. to 06:00 p.m. LT (local time) for both the SGP measurements and Noah estimates.

Noah employs the Simple Water Balance (SWB) model by Schaake et al. (1996) to partition the precipitation into surface runoff and infiltration. Soil moisture deficits in the entire profile and precipitation intensity are accounted for in the implementation of SWB in Noah (Ek et al., 2003). Figure 8 (upper panel) shows that the three simulations yielded very similar surface runoff. DA, which produced the driest soil moisture profile, as expected, yielded the lowest surface runoff. The insensitivity of surface runoff to soil moisture is probably due to the fact that there were no prolonged precipitation periods and that the soil in the basin remained relatively dry which left enough room for infiltration.

On the other hand, the assimilation of AMSR-E has a much larger impact on base flow as shown in Fig. 8 (lower panel). As mentioned early, Noah uses Eq. (2) to calculate base flow which has a monotonic relationship with the soil moisture in layer 4. As a result, DA yielded the lowest base flow while DA MassCon generated the largest base flow. Compared to Control, DA MassCon significantly increased base flow in winter months when more corrections were made to the soil moisture fields. Notice that DA generated significantly smaller amounts of base flow in 2007 than in 2006 which is relatively drier. Frequent rainfall in 2007, which restored bias in the surface, means more water was removed and not captured by DA. Overall, the assimilation results support the findings by Li et al. (2009) who concluded that the initial soil moisture condition has a larger impact on base flow while precipitation uncertainty has a larger impact on surface runoff.

Additional information beyond surface runoff and base flow is required to compute the stream flow for the watershed. In the western US where significant groundwater recharges may occur, simple summation of base flow with surface runoff will lead to overestimation of stream flow. Figure 9 (upper panel) shows the comparison of simulated total runoff (surface runoff plus base flow) with the USGS stream
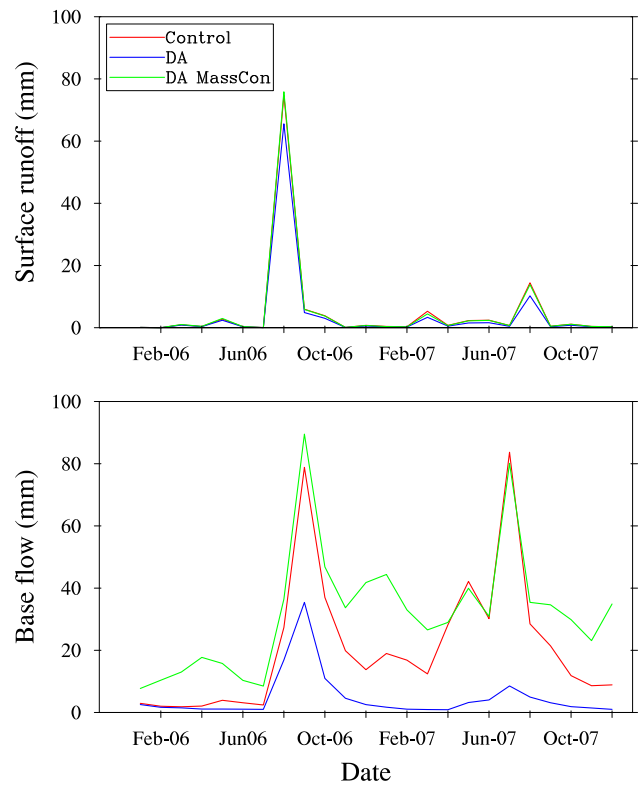

Fig. 8. Monthly surface (top panel) and base flow (lower panel) from Control, DA and DA MassCon.

flow data. The predicted total runoffs are much higher than gauged values, except the result by DA in 2007. Based on the study by Schaller and Fan (2009), about 30 % of total runoff in the Little Washita area contributes to the stream flow. Using this information, the simulated stream flow, which was taken as 30 % of the total runoff, was plotted in the lower panel of Fig. 9. The stream flow estimation by Control and DA MassCon now compare well with the gauge data except for the overestimations in August 2006 and underestimations in June 2007 which were caused by GDAS forcing errors (see Fig. 2). Bias in forcing data cannot be corrected through soil moisture data assimilation since the forcing was assumed to be unbiased. While these comparisons do not constitute accurate validations (which is why no statistics were calculated for stream flow), they illustrate the potential impact of AMSR-E retrievals on runoff by different algorithms. Note that the estimated monthly total runoff and stream flow in Fig. 9 are simple aggregations of the estimated surface and subsurface runoff at all the grid points within the basin. No routing algorithm or time delay was used in producing them, which can be justified given the relatively smaller basin size and the large time scale.
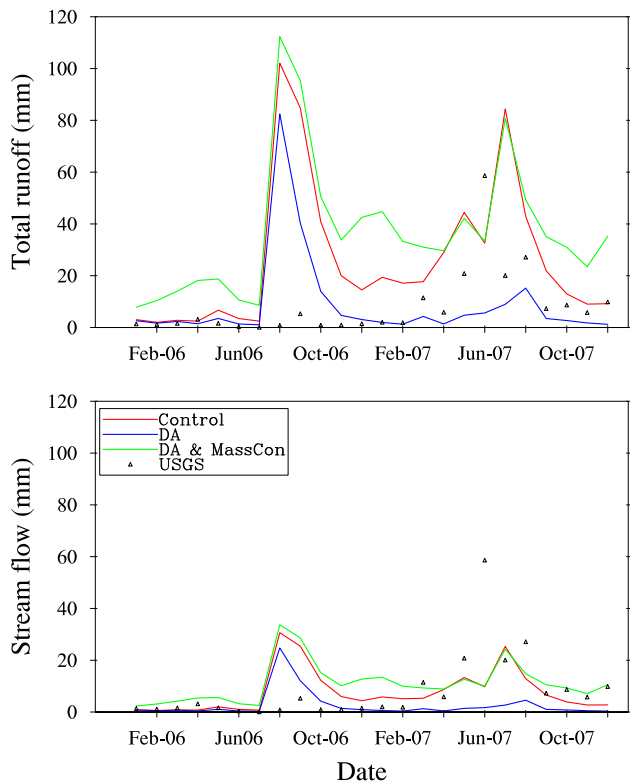
Fig. 9. Comparison of monthly total runoff (top panel) and stream flow (lower panel) from Control, DA and DA MassCon versus USGS gauge data.

4.3 Water budget

As mentioned early, a unique challenge in assimilating remotely sensed data is that the observation is only available for a thin surface layer. An assimilation method, which may look successful based on the verification of soil moisture near the surface, may fail in the lower soil zone. For instance, DA could have been declared a success based on the verification of soil moisture in the shallow root zone and the latent heat. Yet, it degraded soil moisture estimates in the deeper soil profile that led to the deterioration of base flow and failure to show annual precipitation changes in the root zone. Lack of both complete observations and a full set of soil moisture constraints is the root cause for this inconsistent conclusion. To avoid this problem a quality check, independent of any soil moisture measurement, is needed.

Water budget checks represent one way to ensure the assimilation results are physically consistent across all the processes. The water budget here is defined as the sum of ET, surface runoff, base flow and the net change in column water. Since the forcing perturbations were assumed unbiased, the assimilation runs should produce the same water budget as Control in time scales much longer than the perturbation frequency. To assess the overall performance of the two assimilation methods, monthly GDAS precipitation and water

budgets from the three simulations, which, in theory, should be equal to precipitation, are displayed in Fig. 10. While the difference between GDAS precipitation and the water budget of Control is due to numerical errors associated with the discretization of the Richards equation, the difference between Control and the two data assimilation runs can only be attributed to the Kalman filters. The failure of DA is clearly evident because it does not have water budget closure in every month. Failure to capture mass loss from the top two layers and the inappropriate update in the lower layers contribute to the loss of water in the DA run. On the other hand, DA MassCon, in general, achieved monthly water balance throughout the two-year period. Some ensemble bias still existed in DA MassCon in January and February of 2006 when the soil was so dry that the perturbations used in the filter were likely slightly larger than needed.

5 Discussions

The difficulty of using the surface observation to improve root zone soil moisture has been reported by Walker et al. (2001) and Houser et al. (1998) who showed that soil moisture estimates in the lower soil zone deteriorated with the assimilation of surface observations. As pointed out by Walker et al. (2001) that data assimilation could only achieve what model physics is capable of delivering, the failure of the conventional EnKF in updating the lower layers, as shown in this study, is a result of inappropriate model physics. As shown in Figs. 3 and 4, Noah (the Control run) failed to capture the trend of increasing wetness with depth as observed by in situ measurements. As a result, the conventional EnKF was not able to yield increments favoring the improvements in the lower profile. Even for assimilation methods that do not depend on model physics, EnKF could also lead to undesired impacts on the lower soil layers (Houser et al., 1998). Lack of observations in the entire profile to constrain the increments is the root cause for these difficulties. The mass conservation scheme avoids the interference of imperfect model physics for the lower layers by using a model-independent updating equation that also preserves the mass of the forecast.

Moving mass imbalances to the lower layers as presented in the mass conservation scheme was largely based on analyses of model simulation results and considerations of model physics. As reasoned in the Results section, the overestimation at the surface was likely caused by the lower than expected hydraulic conductivity values, given the persistent occurrence of overestimation, especially in winter periods when precipitation and ET were very low. For this reason, moving the surface overestimation to the lower layers acts to mitigate the inaccuracy of model parameters. Redistributing mass imbalances to the lower layers is also computationally efficient and simple to implement as it requires no additional information on fluxes. In comparison, the approach
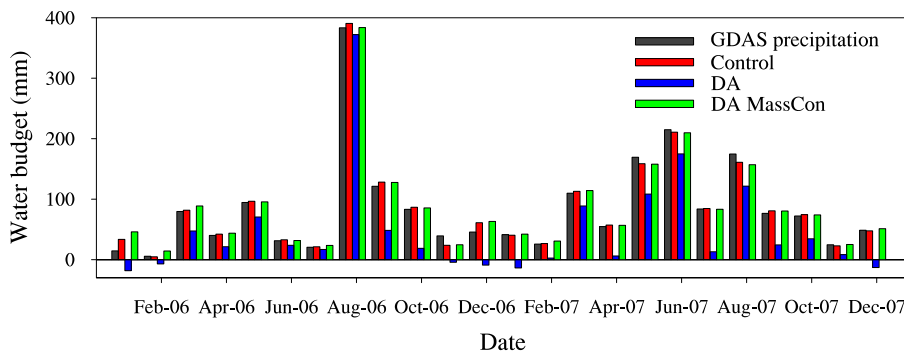
Fig. 10. Monthly GDAS precipitation and water budget for Control, DA and DA MassCon.

by Pan and Wood (2006) redistributes mass imbalances to fluxes, in addition to states, based on an error covariance matrix, and requires gridded ET and runoff observations which are not often available for large scale modeling. The smaller mass imbalances caused by updating just the two upper layers and the fact that it avoids the significant adverse impact of conventional EnKS on the lower soil layers makes the mass conservation scheme more practical for large-scale land data assimilation.

As shown in Fig. 10, the mass conservation scheme does not allow precipitation to fluctuate with assimilation of soil moisture. This restriction prevents potential detrimental impacts on the water budget while assimilating state observations when the error in the state is not linked to error in precipitation. Using Fig. 3 as an example, given the consistent positive bias in Control at the surface, if the water budget were allowed to change with soil moisture data assimilation, the assimilation would always lead to reduced water budgets, which would contradict the precipitation validation in Fig. 2 where GDAS actually overestimated precipitation or was nearly unbiased in some months.

The reduction of bias in both upper and lower layers by DA MassCon changed the soil moisture profile which became more aligned with in situ observations. As mentioned in the Introduction, simulated soil moisture fields from different models exhibit significant disparities which have greatly affected their applications in drought monitoring. Mo (2008) showed that the correlations of model-based drought indices are so low in the western US that they are not reliable for drought monitoring. Assimilating the actual value of AMSR-E retrievals into these models can reduce the uncertainty associated with model physics and should lead to more consistent soil moisture fields and thus, more reliable model-based drought indices. Although we emphasized the improvement in mean soil moisture due to the mismatch of spatial resolutions between AMSR-E retrievals and model estimates, the assimilation changed the full magnitude of soil moisture and thus its impact goes beyond the mean soil moisture fields. The anomaly correlation was not evaluated in this study due to the short simulation period from which meaningful climatology cannot be obtained.

With the current framework of LIS, parameter uncertainties are implicitly represented in errors added to soil moisture variables. Alternatively, parameters uncertainty can be represented through directly perturbing parameters (Margulis et al., 2002; Ng et al., 2009; Qin et al., 2009). The assimilation results should remain similar as they are determined by the relative error of the observation versus that of the model and constrained by the observation and the control run. Perturbing parameters can also be used to simultaneously retrieve model parameters as shown by Qin et al. (2009) who retrieved surface soil moisture and soil texture parameters using a particle filtering technique. Their study showed that changes in initial conditions can lead to completely different retrieved parameter values. Lack of constraints in parameters, particularly the knowledge about their mean values, may be responsible for this behavior. For the case studied here, the hydraulic conductivity is likely biased relative to the truth and its uncertainty can hardly be represented by a zero-mean Gaussian process. Bias, in either parameters or state variables, is an important issue that needs to be considered when assimilating real observations.

6 Conclusions

This study demonstrates that modeled soil moisture fields are significantly biased due to errors in static parameters and inappropriate model physics. Less biased satellite derived soil moisture data can be used to reduce the bias. However, the difference between the mean of model estimates and that of sensor data can also lead to mass imbalances when the bias is corrected near the surface. In addition, since satellite retrievals only represent information in the top few centimeters of the soil, effectively passing the surface information to the deeper soil layers without causing adverse impacts poses additional challenges.

The mass conservation updating scheme developed in this study preserves the water budget of the model forecast by transferring the mass imbalance incurred in updating the top two layers to the lower layers. By updating the lower layers with small increments which also preserves mass balances,

it avoids the negative impact of a conventional EnKF on the lower profile. Although it was found that updating the top two layers is more appropriate at Little Washita, studies in different climate conditions are needed to examine how far the surface measurements can influence the deeper soil layers through the conventional EnKF without causing adverse impacts. A general form of Eq. (6) for a model with N soil layers and the upper L layers are to be assimilated using the conventional EnKF is:

$$y_t^a = y_t^f - \sum_{k=1}^{L} (\Delta C_k \, d_k) / \sum_{j=L+1}^{N} d_j \qquad (7)$$

where y now represents any of the soil moisture state in layers $L+1$ to N. In general, the smaller number of upper layers that are assimilated using the conventional EnKF, the less impact the assimilation has on the rest of soil layers and flux estimates.

With the biased model shown in this study, the estimates by DA and DA MassCon were not optimized, i.e., the estimation error was not minimized. The same is true if the retrievals are scaled, priori to assimilation, using model climatology (Drusch et al., 2005; Reichle and Koster, 2004), because model estimates (Control) were still biased. For the example presented here, more reductions in the estimation error for the surface layer can be obtained by directly inserting the AMSR-E retrievals into the model, as Table 3 shows that AMSR-E retrievals have the smallest bias against the ARS measurements. However, retrievals may not always be better than modeled estimates, in which case data assimilation will yield better estimates than direct insertion. In addition, direct insertion is not as effective as an EnKF in reducing errors in root zone soil moisture because data assimilation techniques can force the surface observation to impact the adjacent soil layer while direction insertion, relying on model physics, may not be effective in passing the information downward (Crow and Wood, 2003).

References

Bolten, J., Crow, W. T., Zhan, X., Reynolds, C., and Jackson, T.J.: Assimilation of a satellite-based soil moisture product in a two-layer water balance model for a global crop production decision support system, Data Assimilation for Atmospheric, Oceanic, and Hydrologic Applications, edited by: Park, S. K., Springer-Verlag, doi:10.1007/978-3-540-71056-1, 449–465, 2008.

Campbell, G. S.: A simple method for determining unsaturated conductivity from moisture retention data, Soil Sci., 117, 311–314, 1974.

Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H., Koren, V., Duan, Q., Ek, M., and Betts, A.: Modeling of land surface evaporation by four schemes and comparison with 1GFIFE observations, J. Geophys. Res., 101, 7251–7268, 1996.

Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R.: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils, Water Resour. Res., 20, 682–690, 1984.

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E., Schaake, J., Robock, A., Marshall, C., Sheffield, ., Duan, Q., Luo, L., Higgins, R., Pinker, R., Tarpley, D., and Meng, J.: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, J. Geophys. Res., 108, 8842, doi:10.1029/2002JD003118, 2003a.

Cosgrove, B. A., Lohman, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Sheffield, J., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T., and Tarpley, J. D.: Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS), J. Geophys. Res., 108, 8845, doi:10.1029/2002JD003316, 2003b.

Crow, W. and Wood, E.: The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman Filtering: a case study based on ESTAR measurements during GSP97, Adv. Water Res., 26, 137–149, 2003.

Crow, W. T. and Zhan, X.: Continental-scale evaluation of remotely sensed soil moisture products, IEEE Geosci. Remote Sens. Lett., 4, 451–455, 2007.

De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Pauwels, V. R. N., and Verhoest, N. E. C.: Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter, Water Resour. Res., 43, W09410, doi:10.1029/2006WR005449, 2007a.

De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Pauwels, V. R. N., and Verhoest, N. E. C.: Correction to "Correcting for forecast bias in soil moisture assimilation with the Ensemble Kalman Filter", Water Resour. Res., 43, W10799, doi:10.1029/2007WR006542, 2007b.

Derber, J. C., Parrish, D. F., and Lord, S. J.: The new global operational analysis system at the National Meteorological Center, Weather Forecast., 6, 538–547, 1991.

Dirmeyer, P. A.: Using a Global Soil Wetness Dataset to Improve Seasonal Climate Simulation, J. Climate, 13, 2900–2922, 2000.

Draper, C. S., Mahfouf, J.-F., and Walker, J. P.: An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme, J. Geophys. Res., 114, D20104, doi:10.1029/2008JD011650, 2009.

Drusch, M., Wood, E. F., and Gao, H.: Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture, Geophys. Res. Lett., 32, L15403, doi:10.1029/2005GL023623, 2005.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J. D.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res., 108, GCP 12-1–12-16, doi:10.1029/2002JD003296, 2003.

Evensen, G. and van Leeuwen, P. J.: Assimilation of Geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasigeostrophic model, Mon. Weather Rev., 124, 85–96, 1996.

Houser, P. R., Shuttleworth, W. J., Famiglietti, J. S., Gupta, H. V., Syed, K. H., and Goodrich, D. C.: Integration of soil moisture remote sensing and hyrologic modeling using data assimilation, Water Resour. Res., 34, 3405–3420, 1998.

Jackson, T. J.: Measuring surface soil moisture using passive microwave remote sensing, Hydrol. Process., 7, 139–152, 1993.

Jury, W. A., Gardner, W. R., and Gardner, W. H.: Soil Physics, John Wiley & Sons, 328 pp., 1991.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 341 pp., 2003.

Keppenne, C. L.: Data Assimilation into a Primitive-Equation Model with a Parallel Ensemble Kalman Filter, Mon. Weather Rev., 128, 1971–1981, 2000.

Keppenne, C. L., Rienecker, M. M., Kurkowski, N. P., and Adamec, D. A.: Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction, Nonlin. Processes Geophys., 12, 491–503, doi:10.5194/npg-12-491-2005, 2005.

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., Houser, P. R., Geiger, J., Olden, S., Lighty, L., Eastman, J. L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E. F., and Sheffield, .: Land Information System – An Interoperable Framework for High Resolution Land Surface Modeling. Environ. Modell. Softw., 21, 1402–1415, 2006.

Kumar, S. V., Reichle, R. H., Peters-Lidard, C. D., Koster, R. D., Zhan, X., Crow, W. T., Eylander, J. B., and Houser, P. R.: A Land Surface Data Assimilation Framework using the Land Information System: Description and Applications, Adv. Water Resour., 31, 1419–1432, 2008.

Li, H., Luo, L., Wood, E. F., and Schaake, J.: The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting, J. Geophys. Res., 114, D04114, doi:10.1029/2008JD010969, 2009.

Margulis, S. A., McLaughlin, D., Entekhabi, D., and Dunne, S.: Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment, Water Resour. Res., 38, 1299, doi:10.1029/2001WR001114, 2002.

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Cosgrove, B. A., Sheffield, J., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T., Tarpley, J. D., Lettenmaier, D. P., Marshall, C. H., Entin, J. K., Pan, M., Shi, W., Koren, V., Meng, J., Ramsay, B. H., and Bailey, A. A.: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, J. Geophys. Res., 109, D07S90, doi:10.1029/2003JD003823, 2004.

Mo, K. C.: Model-based drought indices over the United States, J. Hydrometeorol., 9, 1212–1230, 2008.

Ng, G.-H. C., McLaughlin, D., Entekhabi, D., and Scanlon, B.:

Using data assimilation to identify diffuse recharge mechanisms from chemical and physical data in the unsaturated zone, Water Resour. Res., 45, W09409, doi:10.1029/2009WR007831, 2009.

Ni-Meister, W., Houser, P. R., and Walker, J. P.: Soil moisture initialization for climate prediction: Assimilation of scanning multifrequency microwave radiometer soil moisture data into a land surface model, J. Geophys. Res., 111, D20102, doi:10.1029/2006JD007190, 2006.

Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K., and Nghiem, S. V.: Soil moisture retrieval from AMSR-E, IEEE T. Geosci. Remote, 41, 215–229, 2003.

Pan, M. and Wood, E. F.: Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter, J. Hydrometeorol., 7, 534–547, 2006.

Qin, J., Liang, S., Yang, K., Kaihotsu, I., Liu, R., and Koike, T.: Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal, J. Geophys. Res., 114, D15103, doi:10.1029/2008JD011358, 2009.

Reichle, R. H. and Koster, R.: Land data assimilation with the ensemble Kalman filter: assessing model error parameters using innovations, Proceedings of XIV International Conference on Computational methods in water resources, Delft, Neitherlands, 2002.

Reichle, R. H., and Koster, R.: Bias reduction in short records of satellite soil moisture, Geophys. Res. Lett., 31, L19501, doi:10.1029/2004GL020938, 2004.

Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P. P., Njoku, E. G., and Owe, M.: Comparison and assimilation of global soil moisture retrievals from the Advanced Microware Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), J. Geophy. Res., 112, D09108, doi:10.1029/2006JD008033, 2007.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., J. P. Walker, D. Lohmann, K., and Toll, D.: The Global Land Data Assimilation System, B. Am. Meteorol. Soc., 85, 381–394, 2004.

Rodell, M., Houser, P. R., Berg, A. A., and Famiglietti, J.S.: Evaluation of 10 Methods for Initializing a Land Surface Model, J. Hydrometeorol., 6, 146–155, 2005.

Ryu, D., Crow, W. T., Zhan, X., and Jackson, T. J.: Correcting unintended perturbation biases in hydrologic data assimilation using Ensemble Kalman filter, J. Hydrometeorol., 10, 734–750, doi:10.1175/2008JHM1038.1, 2009.

Schaake, J. C., Koren, V., Duan, Q., Mitchell, K., and Chen, F.: Simple water balance model for estimating runoff at different spatial and temporal scales, J. Geophys. Res., 101, 7461-7475, 1996.

Schaefer, G. L, Cosh, M. H., and Jackson, T. J.: The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN), J. Atmos. Ocean. Tech., 24, 2073–2077, 2007.

Schaller, M. F. and Fan, Y.: River basins as groundwater exporters and importers: Implications for water cycle and climate modeling, J. Geophy. Res., 114, D04103, doi:10.1029/2008JD010636, 2009.

Walker, J. P., Willgoose, G. R., and Kalma, J. D.: One-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: a simplified soil moisture model and field application, J. Hydrometeorol., 2, 356–373, 2001.

Wood, E. F., Lettenmaier, D. P., Liang, X., Lohmann, D., Boone, A., Chang, S., Chen, F., Dai, Y., Dickinson, R., Duan, Q., Ek, M., Gusev, Y., Habets, F., Irannejad, P., Koster, R., Mitchel, K., Nasonova, O., Noilhan, J., Schaake, J., Schlosser, A., Shao, Y., Shmakin, A., Verseghy, D., Warrach, K.,Wetzel, P., Xue, Y., Yang, Z., and Zeng, Q.: The Project for Intercomparison of Land-surface Parameterization Schemes_PILPS/Phase 2_c/Red–Arkansas River basin experiment: 1. Experiment description and summary intercomparisons, Global Planet. Change, 19, 115–135, 1998.

Zeng, X. and Decker, M.: Improving the numerical solution of soil moisture-based Richards equation for land models with a deep or shallow water table, J. Hydrometeorol., 10, 308–319, 2009.

Zhan, X., Liu, J., Jackson, T. J., Meng, J., Weng, F., and Mitchell, K.: Merging satellite soil moisture retrievals and model simulations for a blended consistent operational global soil moisture data product, SPIE Annual Conference on Atmospheric and Environmental Remote Sensing Data Processing and Utilization IV: Readiness for GEOSS II, San Diego, CA, 10–14 August 2008.