AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2013WR014586

Special Section:
Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling and Data Assimilation

Soil moisture and soil properties estimation in the Community Land Model with synthetic brightness temperature observations

Xujun Han[1,2], Harrie-Jan Hendricks Franssen[2], Carsten Montzka[2], and Harry Vereecken[2]

[1]Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, People's Republic of China, [2]Forschungszentrum Jülich, Agrosphere (IBG 3), Jülich, Germany

Abstract The Community Land Model (CLM) includes a large variety of parameterizations, also for flow in the unsaturated zone and soil properties. Soil properties introduce uncertainties into land surface model predictions. In this paper, soil moisture and soil properties are updated for the coupled CLM and Community Microwave Emission Model (CMEM) by the Local Ensemble Transform Kalman Filter (LETKF) and the state augmentation method. Soil properties are estimated through the update of soil textural properties and soil organic matter density. These variables are used in CLM for predicting the soil moisture retention characteristic and the unsaturated hydraulic conductivity, and the soil texture is used in CMEM to calculate the soil dielectric constant. The following scenarios were evaluated for the joint state and parameter estimation with help of synthetic L-band brightness temperature data assimilation: (i) the impact of joint state and parameter estimation; (ii) updating of soil properties in CLM alone, CMEM alone or both CLM and CMEM; (iii) updating of soil properties without soil moisture update; (iv) the observation localization of LETKF. The results show that the characterization of soil properties through the update of textural properties and soil organic matter density can strongly improve with assimilation of brightness temperature data. The optimized soil properties also improve the characterization of soil moisture, soil temperature, actual evapotranspiration, sensible heat flux, and soil heat flux. The best results are obtained if the soil properties are updated only. The coupled CLM and CMEM model is helpful for the parameter estimation. If soil properties are biased, assimilation of soil moisture data with only state updates increases the root mean square error for evapotranspiration, sensible heat flux, and soil heat flux.

Key Points:
- Assimilation with biased soil properties can worsen surface fluxes characterization
- Spatial distribution of model states and soil properties can be updated
- Joint assimilation works better for a coupled land surface and microwave model

Correspondence to:
X. Han,
hanxj@lzb.ac.cn

Citation:
Han, X., H.-J. H. Franssen, C. Montzka, and H. Vereecken (2014), Soil moisture and soil properties estimation in the Community Land Model with synthetic brightness temperature observations, *Water Resour. Res.*, *50*, 6081–6105, doi:10.1002/2013WR014586.

Received 16 AUG 2013
Accepted 2 JUL 2014
Accepted article online 7 JUL 2014
Published online 23 JUL 2014

1. Introduction

In general, land surface models and hydrologic models use many parameterizations to characterize the water and energy balance [*Han et al*., 2012; *Liang et al*., 1994; *Niu et al*., 2011; *Oleson et al*., 2010; *Sellers et al*., 1996]. Hydrological and thermal processes like the transport of water and energy between the land surface and atmosphere are strongly controlled by the soil moisture and the temperature status of the soil. The space-time dynamics of these state variables is dependent on soil material properties such as the hydraulic and thermal conductivity. These soil material properties are difficult to obtain directly from measurements and their upscaling to scales larger than the field scale is hampered through the lack of appropriate upscaling methods and spatial information on heterogeneity. At many locations, no information is available on key soil material properties. Pedotransfer functions have been developed in the past to estimate soil hydraulic and thermal parameters from simple soil properties such as soil texture and soil organic carbon without the need for direct measurements [*Gutmann and Small*, 2007; *Vereecken et al*., 2010]. These functions have been successfully used in vadose zone and land surface models to predict water and matter fluxes [*Oleson et al*., 2010; *Rigon et al*., 2006; *Šimůnek et al*., 2008]. However, an inaccurate or incorrect specification of soil hydraulic and thermal parameters through inaccurate estimates of soil properties and/or uncertain or inaccurate pedotransfer functions (sand, clay, and/or soil organic matter density) will introduce biases into the model simulation of water and energy cycles [*Dai et al*., 2013; *Duan et al*., 2006; *Huang and Liang*, 2006; *Liu and Gupta*, 2007]. As it is difficult to obtain information about soil hydraulic and thermal parameter values and their statistics with a good spatial coverage, alternative approaches have to be used to determine these parameters. Optimal estimates of these parameters can be derived through inverse

modeling using measurements of state variables or through data assimilation approaches with spatially retrieved remote sensing measurements.

Many parameter estimation approaches have been proposed to cope with the lack of knowledge of model parameters and their statistics, such as generalized likelihood uncertainty estimation (GLUE) [*Beven and Binley*, 1992], the shuffled complex evolution method [*Duan et al*., 1992], shuffled complex evolution metropolis [*Vrugt et al*., 2003], particle swarm optimization [*Scheerlinck et al*., 2009], differential evolution adaptive metropolis [*Vrugt and Ter Braak*, 2011], multicriteria parameter estimation [*Livneh and Lettenmaier*, 2012], or Bayesian recursive estimation [*Thiemann et al*., 2001]. However, the application of these methods in combination with large-scale land surface models is often complicated, as these methods require in general a large number of model realizations, which are repeatedly applied on the complete time series under consideration. A second complication is that these methods focus on parameter estimation, whereas in land surface models other sources of uncertainty (especially uncertainty of the forcings and uncertainty/errors in the model structure) also need to be considered. Whereas it is not trivial to define an adequate statistical model of the space-time uncertainty of the forcings, it is even much more complicated to include the model structural uncertainty in the analysis [*Liu et al*., 2012]. As a complication, the incorrect definition of forcing errors and model structure in the context of parameter estimation might result in biased parameter estimates. It is to be preferred to address the influences of model structural uncertainty and forcing data uncertainty before parameter estimation [*Liu and Gupta*, 2007].

Data assimilation provides an approach to estimate soil hydraulic and thermal parameters or also vegetation parameters while considering model structural and model forcing uncertainties [*Liu and Gupta*, 2007]. In sequential data assimilation, several methods have been proposed for joint state and parameter estimation. This joint scheme expands the data assimilation framework from updating the model states only to updating both model states and model parameters. In general, two kinds of methods can be used to estimate the states and parameters jointly: (1) SODA (Simultaneous Optimization and Data Assimilation). Here the Bayesian filtering method is combined with an optimization algorithm to update the model states and parameters separately [*Vrugt et al*., 2005]; (2) State augmentation approach. Here the model state vector is augmented to include both model states and parameters in a new vector, and states and parameters are updated simultaneously in the framework of the ensemble Kalman filter [*Bateni and Entekhabi*, 2012; *Jazwinski*, 1970; *Nie et al*., 2011; *Schöniger et al*., 2012].

Due to the computational burden of the SODA method (especially for complex distributed models), it has not been applied widely. The state augmentation method has become a popular approach in recent data assimilation applications, in which the ensemble Kalman filter [*Moradkhani et al*., 2005a; *Nie et al*., 2011; *Schöniger et al*., 2012] or particle filter [*Montzka et al*., 2011; *Moradkhani et al*., 2005b] is used to estimate the model states and model parameters jointly. These studies have shown the positive impacts of joint state and parameter estimation on the characterization of model states, parameters, and fluxes. An additional advantage is that temporally variable parameters can also be characterized with this approach [*Kurtz et al*., 2012; *Montzka et al*., 2013a].

In this study, we want to evaluate the joint state and parameter estimation method for the coupled Community Land Model 4.5 (CLM) [*Oleson et al*., 2013] and Community Microwave Emission Model (CMEM) [*de Rosnay et al*., 2009] through assimilation of synthetic microwave brightness temperature data. In land data assimilation, soil moisture plays a key role. Microwave remote sensing measurements become more and more important in soil moisture data assimilation for water resources management, runoff forecasting [*Brocca et al*., 2010], and weather forecast [*Drusch*, 2007]. Multisource remote sensors can provide soil moisture products for the regional or global scale, examples are the Soil Moisture and Ocean Salinity (SMOS) mission [*Kerr et al*., 2010], Soil Moisture Active Passive (SMAP) Mission [*Entekhabi et al*., 2010], and the European Space Agency (ESA) Soil Moisture Climate Change Initiative (CCI) [*Liu et al*., 2011]. These large-scale soil moisture products provide the opportunity for regional-scale land data assimilation and parameter estimation. The European Center for Medium range Weather Forecasting (ECMWF) is using CMEM to assimilate the SMOS brightness temperature into the forecast system [*Sabater et al*., 2012]. In CMEM, the Dobson model [*Dobson et al*., 1985] is used to calculate the soil dielectric constant as a function of soil moisture, sand fraction, clay fraction, sensor frequency, and surface soil temperature. Therefore, it is not only that the CLM model results are sensitive to the sand, clay fractions, and organic matter density, but the sand and

clay fractions are also very important to the L-band microwave brightness temperature modeling in CMEM. Biased soil parameter values will result in poor model predictions of both CLM and CMEM.

Synthetic passive brightness temperature observations were assimilated instead of soil moisture in this study for two reasons: (1) if directly soil moisture is assimilated (level 2 product), a microwave emission model like CMEM was already used for obtaining the soil moisture estimates, and soil texture information was used as input for these estimates. Erroneous soil texture parameters will also result in biased soil moisture estimates. Moreover, if soil hydraulic parameters are estimated together with model states this can result in soil texture/soil hydraulic parameters, which are different for CLM and CMEM and which are therefore inconsistent. The assimilation of brightness temperature can probably reduce the bias introduced by erroneous soil texture information in both CLM and CMEM; (2) the soil temperature is also a necessary input for microwave emission models like CMEM and will therefore also influence the soil moisture estimate. A land surface model allows defining this input as well in a physically consistent matter, again providing consistency between input for the microwave emission model and the land surface model.

The objectives of this study are: (1) to evaluate the feasibility of jointly estimating soil moisture and soil hydraulic and thermal parameters in CLM using sequential data assimilation of synthetic L-band microwave brightness temperature, (2) to analyze the impacts of soil properties estimation for the coupled CLM and CMEM models, and (3) to study the impact of joint state and parameter estimation on the model states (soil moisture, soil temperature) and surface fluxes. In order to be able to validate the results and in particular the quality of the updated parameters, a synthetic study was designed. Assimilation of real-world data was beyond the scope of this study. It requires a lot of additional sources of information like the temporal variation of vegetation properties (leaf area index, vegetation type). The coarse brightness temperature observations need to be downscaled, bias corrected, and validated carefully for the high-resolution study. The downscaling of brightness temperature is still an ongoing subject of study with no straightforward solution yet [*Song et al*., 2014]. Moreover, the real brightness temperature could be influenced by external noise, such as the L-band radio frequency interference (RFI) [*Oliva et al*., 2012] and Faraday rotation [*Sabater et al*., 2012]. The assimilation of multiple-incidence angle (SMOS) brightness temperature data also remains a challenge [*Montzka et al*., 2013a]. The synthetic study mimics the Rur catchment in Germany.

The structure of this paper is as follows: materials and methods (CLM, CMEM, LETKF, and updating soil states and properties at grid cells without brightness temperature measurement) are introduced in section 2. Section 3 explains the synthetic experiment setup. The results and discussion are presented in sections 4 and 5, respectively. Section 6 presents the final conclusions.

2. Materials and Methods

Several studies focused on updating model parameters like the saturated hydraulic conductivity or the Mualem-van Genuchten parameters of the governing Richards equation [*Montzka et al*., 2011; *Nie et al*., 2011]. However, some land surface or hydrologic models (e.g., Community Land Model) use soil properties (i.e., soil texture and soil organic matter density) in combination with model internal pedotransfer functions to derive soil hydraulic and thermal parameter values. In this study, we focus on the update of these texture classes and therefore on the associated soil hydraulic and thermal properties.

2.1. Community Land Model

We used the well-developed and versatile CLM to model the water and energy balance for the Rur catchment in Germany (see section 3 for a description of the catchment). CLM uses a modified Richards equation to predict the one-dimensional multilayer vertical soil water flow and energy transport. The Monin-Obukhov similarity theory is used to derive the land surface fluxes. The spatial land surface heterogeneity is considered in CLM and soil moisture, soil temperature, sensible heat flux, latent heat flux, and soil heat flux are model output [*Oleson et al*., 2013]. CLM uses the pedotransfer functions of sand and clay [*Clapp and Hornberger*, 1978; *Cosby et al*., 1984] and organic properties of the soil [*Lawrence and Slater*, 2007] to derive soil hydraulic and thermal parameters.

In CLM, the soil profile is divided into 15 layers and the soil moisture is only simulated for the first 10 layers. If i is the layer number, the hydraulic conductivity K (mm/s) is defined at the depth of the interface of two adjacent layers i and $i+1$. It is calculated as function of the saturated hydraulic conductivity K_{sat} (mm/s) and soil moisture content of the two layers i and $i+1$ [θ_i (mm^3/mm^3) and θ_{i+1} (mm^3/mm^3)], saturated soil

moisture content θ_{sat} (mm^3/mm^3) of the same two layers i and $i+1$, the ice impedance factor ϑ_{ice}, and the exponent B [*Oleson et al.*, 2013].

$$K=\left\{\begin{array}{ll}\vartheta_{ice}K_{sat}\left[\frac{0.5(\theta_i+\theta_{i+1})}{0.5(\theta_{sat,i}+\theta_{sat,i+1})}\right]^{2B+3} & 1\le i\le 9\\ \vartheta_{ice}K_{sat}\left[\frac{\theta_i}{\theta_{sat,i}}\right]^{2B+3} & i=10\end{array}\right\} \quad (1)$$

$$\vartheta_{ice}=10^{-6F_{ice}} \quad (2)$$

$$F_{ice}=\frac{\theta_{ice}}{\theta_{sat}} \quad (3)$$

where F_{ice} is the ice-filled fraction of the pore space, the ice impedance factor ϑ_{ice} is used to quantify the increased tortuosity of the water flow when part of the pore space is filled with ice [*Swenson et al.*, 2012].

The soil water matric potential φ (mm) is defined as:

$$\varphi=\varphi_{sat}S^{-B} \quad (4)$$

where φ_{sat} is the saturated matric potential (mm) and S (-) is the wetness of the soil layer with respect to saturation (values between 0.01 and 1.0):

$$\varphi_{sat}=(1-f_{om})\varphi_{sat,min}+(-10.3)*f_{om} \quad (5)$$

$$\varphi_{sat,min}=-10.0*10^{1.88-0.0131(\%sand)} \quad (6)$$

where $\varphi_{sat,min}$ is the saturated mineral soil matric potential, $\%sand$ is the sand fraction (in %), f_{om} is the organic matter fraction (in %) [*Oleson et al.*, 2013]. The direct input of CLM is organic matter density ρ_{om} (kg/m^3), the organic matter fraction is defined as a function of organic matter, and 130 (kg/m^3) is the maximum organic matter density where soil is assumed to act like peat:

$$f_{om}=\rho_{om}/130.0 \quad (7)$$

The saturated soil water content θ_{sat} is:

$$\theta_{sat}=(1-f_{om})\theta_{sat,min}+0.9*f_{om} \quad (8)$$

where the porosity of the mineral soil $\theta_{sat,min}$ is:

$$\theta_{sat,min}=0.489-0.00126(\%sand) \quad (9)$$

The exponent B (an empirical parameter in Clapp-Hornberger parameterizations) [*Clapp and Hornberger*, 1978] is defined as:

$$B=(1-f_{om})(2.91+0.159(\%clay)+2.7*f_{om} \quad (10)$$

where %clay is the clay fraction (in %).

The saturated hydraulic conductivity for mineral soil $K_{sat,min}$ (mm/s) is related to the sand fraction [*Cosby et al.*, 1984]:

$$K_{sat,min}=0.0070556*10^{-0.884+0.0153(\%sand)} \quad (11)$$

The saturated hydraulic conductivity K_{sat} (mm/s) is related to $K_{sat,min}$:

$$f_{perc}=\left\{\begin{array}{ll}0.5*(f_{om}-0.5)^{-0.139}*f_{om} & f_{om}\ge 0.5\\ 0 & f_{om}<0.5\end{array}\right\} \quad (12)$$

$$K_{sat} = (1 - f_{perc}) * \left[\frac{1 - f_{om}}{K_{sat,min}} + \frac{f_{om} - f_{perc}}{K_{sat,om}} \right]^{-1} + f_{perc} * K_{sat,om} \quad (13)$$

where f_{perc} is the fraction of connected soil organic matter fraction for each grid cell and $K_{sat,om}$ (mm/s) is the saturated hydraulic conductivity for organic soils.

The thermal conductivity of soil solids, λ_s (W/mK) is related to the organic matter fraction, sand fraction, and clay fraction:

$$\lambda_s = (1 - f_{om})\lambda_{s,min} + 0.25 * f_{om} \quad (14)$$

where the mineral (*min*) soil solid thermal conductivity $\lambda_{s,min}$ (W/mK) is:

$$\lambda_{s,min} = \frac{8.8(\%sand) + 2.92(\%clay)}{(\%sand) + (\%clay)} \quad (15)$$

The sand fraction, clay fraction, and organic matter density were chosen as the parameters to be estimated instead of the soil hydraulic parameters of the equations given above. The direct update of hydraulic or thermal parameters in CLM is not trivial as these equations are hard coded in CLM and only sand fraction, clay fraction, and organic matter density are direct input data.

2.2. Synthetic L-Band Brightness Temperature Observation

Passive microwave remote sensing has been extensively applied in soil moisture retrieval. The microwave emission radiometer at L-band (1.4 GHz) has been shown to be the best frequency for soil moisture detection and it is currently used in the SMOS satellite mission and will be used in the upcoming SMAP mission. The Community Microwave Emission Model (CMEM) is developed for passive microwave soil moisture retrieval and data assimilation purposes [*de Rosnay et al*., 2009; *Holmes et al*., 2008].

In CMEM, the soil brightness temperature TB_{toa} (K), which is measured by the passive microwave remote sensor, is expressed as:

$$TB_{toa} = TB_{au} + \exp(-\tau_{atm}) * TB_{tov} \quad (16)$$

where TB_{au} (K) is the upwelling atmospheric brightness temperature, and τ_{atm} is the atmospheric optical depth. The top of vegetation brightness temperature is obtained for the case that the vegetation is represented as a single scattering layer above a rough surface:

$$TB_{tov} = TB_{soil} * \exp(-\tau_{veg}) + TB_{veg} * \left[1 + r_r * \exp(-\tau_{veg})\right] + TB_{ad} * r_r * \exp(-2 * \tau_{veg}) \quad (17)$$

where the soil contribution TB_{soil} (K) is defined as:

$$TB_{soil} = T_{eff} * e \quad (18)$$

and the vegetation contribution TB_{veg} (K) as:

$$TB_{veg} = T_c * (1 - \omega) * (1 - \exp(-\tau_{veg})) \quad (19)$$

TB_{ad} (K) is the downwelling atmospheric brightness temperature. r_r is the reflectivity of the rough surface (equal to one minus the emissivity *e*), τ_{veg} is the vegetation optical depth along the viewing path, and τ_{veg} is defined as the function of vegetation water content, which is calculated as $0.5 * Leaf_Area_Index$. ω is the single scattering albedo. T_{eff} (K) is the effective temperature of the surface medium and T_c (K) is the canopy temperature.

The emissivity *e* is computed as a function of the smooth surface emissivity [*Wigneron et al*., 2007]. The calculation of smooth surface emissivity in CMEM is based on the Dobson model [*Dobson et al*., 1985] and Wilheit model [*Parrens et al*., 2014; *Wilheit*, 1978], which represent the soil as a stratified medium where the soil dielectric constant and soil temperature vertical profiles are used to compute the resulting air-soil interface emission and the contribution from each soil layer.

For the horizontal polarization, the reflectivity r_r is defined as:

$$r_r = \left| \frac{\mu_s \cos(\theta) - \sqrt{\mu_s \varepsilon_b - \sin^2 \sin(\theta)}}{\mu_s \cos(\theta) + \sqrt{\mu_s \varepsilon_b - \sin^2 \sin(\theta)}} \right|^2 \quad (20)$$

where μ_s is the soil magnetic permeability, θ (°) is the sensor incidence angle, and ε_b is the complex, smooth, bare soil dielectric constant.

With the Dobson model, the dielectric constant ε_b of wet soil can be calculated as:

$$\varepsilon_b = \left[1 + \frac{\rho_b}{\rho_s} \left(\varepsilon_{pa}^{\alpha} - 1 \right) + SM^{\beta'} \varepsilon_{sfw}^{'\alpha} - SM \right]^{\frac{1}{\alpha}} - j(SM^{\beta''} \varepsilon_{sfw}^{''\alpha})^{\frac{1}{\alpha}} \quad (21)$$

$$\beta' = (127.480 - 0.519 * \%sand - 0.152 * \%clay / 100.0) \quad (22)$$

$$\beta'' = (133.797 - 0.603 * \%sand - 0.166 * \%clay / 100.0) \quad (23)$$

where ρ_b (g/cm^3) is the soil bulk density; ρ_s (g/cm^3) soil particle density; $\varepsilon_{pa}^{\alpha}$ is the dielectric constant of solid particles; α is 0.65; SM (cm^3/cm^3) volumetric water content of soil; $\varepsilon_{sfw}^{'\alpha}$ and $\varepsilon_{sfw}^{''\alpha}$ are dielectric constants of free water included in the soil.

CMEM needs leaf area index, vegetation fractions, vertical profiles of soil moisture content and soil temperature (for seven layers), and land surface temperature, in our simulation experiments calculated by CLM, as input for calculating L-band brightness temperature.

2.3. Local Ensemble Transform Kalman Filter

The Local Ensemble Transform Kalman Filter (LETKF) is one of the popular ensemble Kalman filter variants and used frequently in atmospheric data assimilation [*Hunt et al*., 2007; *Miyoshi and Yamane*, 2007]. The LETKF updates all model grid cells separately in an assimilation cycle and can be easily parallelized [*Han et al*., 2012; *Hunt et al*., 2007]. This scheme is suitable for simulations with a large number of grid cells, such as high-resolution regional data assimilation or global-scale data assimilation. This was an important reason to select LETKF as the assimilation algorithm for this study. The reader is referred to *Hunt et al*. [2007] for further details about LETKF. In order to estimate the soil moisture and soil properties simultaneously within the data assimilation framework, the state augmentation method was used [*Franssen and Kinzelbach*, 2008; *Li and Ren*, 2011; *Moradkhani et al*., 2005a; *Yang and Delsole*, 2009], in which the soil moisture and soil properties for each grid cell were augmented into one state vector of LETKF and updated jointly using the same observations. In data assimilation, we first used Wilheit model [*Wilheit*, 1978] to calculate the weights (the contribution of each soil layer to the air-soil interface emission) of different soil layers and then treated the weighted soil moisture as the model prediction state to be updated in assimilation. In theory, the soil moisture and soil properties can be updated simultaneously. Because of the large computer memory requirements in joint state and parameter estimation, a dual step approach was used, which estimates the parameter vector and the state vector separately and saves memory. In the parameter estimation step, the augmented state vector contains the weighted soil moisture, sand fraction, clay fraction, and organic matter density; in the state estimation step, it includes only the weighted soil moisture followed by 10 layers of soil moisture values for each grid cell. The dual step approach gives the same results as the one step approach but reduces memory requirements; the memory requirement of each separate step is reduced, and this is especially helpful for large-scale assimilation on supercomputers with limited memory.

The analysis of LETKF is divided into two steps, the global operations and the local analysis:

1. Global operations: two global matrices are constructed after the forecast step:

$$\mathrm{X}^b = [x_1^b - \overline{x}^b, \ldots, x_N^b - \overline{x}^b] \quad (24)$$

$$\mathrm{y}_i^b = H(\mathrm{x}_i^b) \quad (25)$$

$$Y^b = [y_1^b - \overline{y}^b, \ldots, y_N^b - \overline{y}^b] \quad (26)$$

where $x_1^b, \ldots x_N^b$ are the model forecast ensemble members, N is the ensemble size, $\overline{x}^b$ is the ensemble mean calculated over all ensemble members $x_1^b, \ldots x_N^b$, H is the observation operator (CMEM model in this

study). In the parameter updating step, x_i^b and $\overline{x}^b$ are composed of weighted soil moisture $\theta_{weighted}$, sand fraction, clay fraction, and organic matter density; for state updating it consists of the weighted soil moisture and soil moisture values for 10 soil layers and for each grid cell $(\theta_1, \ldots, \theta_{10})$. Therefore, the dimension of the augmented state vector for the parameter update is 4 (for each model grid cell) whereas it is 11 for state update. The same weighted soil moisture content was used in both parameter update and soil moisture update.

For the state update, the state vector $\overline{x}^b$ is:

$$\overline{x}^b = \begin{bmatrix} \theta_{weighted} \\ \theta_1 \\ \vdots \\ \theta_{10} \end{bmatrix} \tag{27}$$

For the parameter update, the state vector $\overline{x}^b$ is:

$$\overline{x}^b = \begin{bmatrix} \theta_{weighted} \\ Sand \\ Clay \\ Organic \end{bmatrix} \tag{28}$$

Global operations are a preparation step of LETKF. LETKF updates the model grid cells separately, but the state/observation vectors of all model grid cells only need to be prepared once. Moreover, because all the data are recorded in the file, it is better to open the file once and read all the ensemble members of all model grid cells instead of opening the file many times.

2. Local analysis: select the local observations for each model grid cell and calculate the local analysis error covariance and perturbations in the ensemble space.

Calculate analysis error covariance matrix:

$$\mathrm{P}^a = [(N-1)I + Y^{bT} R^{-1} Y^b] \tag{29}$$

Next, the perturbations in ensemble space are calculated as:

$$\mathrm{W}^a = [(N-1)\mathrm{P}^a]^{1/2} \tag{30}$$

The perturbations in ensemble space are computed and added to each column of W^a to get the analysis ensemble in ensemble space:

$$\overline{w}^a = \mathrm{P}^a Y^{bT} R^{-1} (y^o - \overline{y}^b) \tag{31}$$

The new analysis is calculated by:

$$X^a = X^b \mathrm{W}^a + \overline{x}^b \tag{32}$$

where R is the observation error covariance matrix, y^o contains the observations and X^a contains the model ensemble members after the update with LETKF (analysis).

An important issue for the ensemble-based method is that the assimilation performance depends strongly on the ensemble spread. The ensemble spread may narrow down in the course of the parameter estimation such that most of the ensemble members would become very close to the ensemble mean value. This will result in filter divergence [*Whitaker and Hamill*, 2012] and ensemble inflation methods have been proven to be an effective way to avoid filter divergence. In this study, the multiplicative inflation algorithm proposed by *Whitaker and Hamill* [2012] was applied to the soil moisture and soil properties ensemble, in which the

inflation of the posterior ensemble is proportional to the amount that observations reduce the ensemble spread, resulting in more inflation in regions of dense observations.

$$X_i^a = X_i^a \left(\frac{\sigma^b - \sigma^a}{\sigma^a} + 1 \right) \tag{33}$$

where X_i^a is the updated model state of ensemble member i. σ^b is the prior ensemble standard deviation and σ^a is the posterior ensemble standard deviation.

This methodology is designed to keep ensemble spread. For small ensemble sizes, which are often used in combination with high-resolution physically based models, this is necessary to avoid a strong underestimation of ensemble variance and associated filter divergence. The consequence of the multiplicative inflation applied here is that the impact of data assimilation on reduction of prediction uncertainty cannot be assessed properly, as the spread is kept artificially large. This is a drawback of the applied inflation.

2.4. Local Analysis of Brightness Temperatures

For dense and high vegetation (e.g., forests), the soil contribution to the microwave emission is low and masked by the vegetation. Soil moisture cannot be retrieved effectively for those areas and the uncertainty is too high for adequate consideration during assimilation. Therefore, parts of the catchment will lack information on the brightness temperature and will remain noncovered. We used a local analysis method proposed in LETKF that exploits the presence of observations near noncovered regions to update the brightness temperature at these locations [*Greybush et al*., 2011; *Han et al*., 2012; *Hunt et al*., 2007]. *Han et al*. [2012] showed that a limited number of locations can be used to improve the assimilation results of noncovered regions. This scheme was extended in this study to not only update soil moisture but also soil properties for both covered and noncovered areas. Surrounding local brightness temperature observations were selected, based on the spatial correlation between the observation location and the grid cell for which an update is required, and used for updating both states and soil properties at the given grid cell. The best fitted semivariogram model was chosen from spherical model, exponential model, and Gaussian model at each assimilation step for the brightness temperature data, and was used to weight the local selected observations [*Han et al*., 2012]. A correlogram was calculated using the normalized semivariogram value. The correlogram is equal to 1 at the observation location, and it gradually reduces toward 0.0 as the distance from the model grid cell increases [*Han et al*., 2012]. The selected observations were assimilated directly without interpolation. The observation variance R was divided by the correlogram value corresponding to the observation location and model grid cell location. This increases the variance for observations situated far away from the grid cell, which is updated, and decreases the contributions of those observations.

LETKF assimilates each model grid cell separately. In the LETKF analysis step, the observations to be assimilated for each model grid cell need to be chosen based on the spatial correlation characteristics of observations. For those model grid cells for which a brightness temperature observation was available (i.e., covered grid cells), only one local observation was assimilated for the joint state parameter estimation procedure. For the nonobserved model grid cells, nine surrounding observations were used in the data assimilation procedure. This is in correspondence with findings by *Han et al*. [2012]. However, for the local analysis of nonobserved model grid cells that include parameter estimation, the optimum number of local observations was first evaluated in the scenarios of joint estimation.

3. Experiment Setup

3.1. Study Area

This synthetic assimilation study mimics the Rur catchment, which is located in the west of Germany bordering to Belgium and Netherlands. The area of the Rur catchment is 2354 km^2 (Figure1). The southern region corresponds to the hills of the Eifel and is covered by needleleaf and broadleaf forests and grassland, and the annual precipitation and potential evapotranspiration for this region are 850–1300 and 450–550 mm/yr, respectively. In the northern region, fertile agricultural land predominates, and the annual precipitation and potential evapotranspiration are 650–850 and 580–600 mm/yr, respectively [*Montzka et al*.,

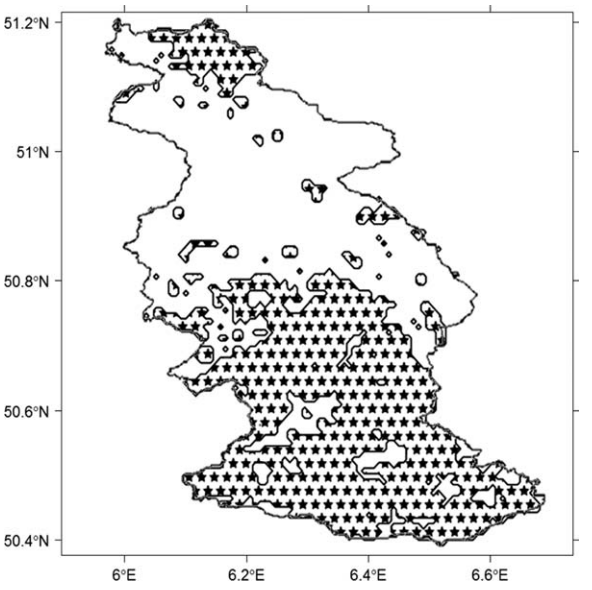

Figure 1. Forest area (marked as star) map of the Rur catchment.

2008]. The Terrestrial Environmental Observatories (TERENO) initiative has established a dense network of measurement instruments at the Rur catchment, such as doppler radar for rainfall intensity determination and automatic weather stations, soil moisture networks, eddy-covariance systems, cosmic ray sensors, and radiometer systems [*Bogena et al.*, 2010; *Zacharias et al.*, 2011]. The Rur Catchment is also a validation site for the ESA SMOS mission [*Montzka et al.*, 2013b] and the planned NASA SMAP mission.

3.2. Reference Run

A reference run with, as input, the true sand, silt, and clay fractions and true organic matter density as well as the true model forcings was made for the period from 1 April 2010 to 30 September 2010 after 1 year spin-up. This run serves as the truth or reference for comparison in this Observing System Simulation Experiment (OSSE). The 3 months OSSE period was characterized by a very dry first month, but in May 2010, rainfall was increased with a series of rainfall events.

The CLM spatial resolution of the Rur catchment was 0.00833° (approximately 750 m) with in total 4340 active grid cells. We used the MODIS 500 m Plant Functional Type (PFT) product MCD12Q1, which was projected and resampled to 0.008333° with the nearest neighbor method [*Sun et al.*, 2008] for fixing the vegetation type. Soil types were fixed based on the Harmonized World Soil Database v1.1 (HWSD) [*FAO et al.*, 2010]. The two layer soil data of HWSD were linearly interpolated to generate soil properties for 10 layers, and the soil layer thicknesses of the first 10 layers are shown in Figure 2. The atmospheric forcing provided by the Global Land Data Assimilation System (GLDAS) project [*Rodell et al.*, 2004] was interpolated using the bilinear IPOLATES grid interpolation library (http://www.nco.ncep.noaa.gov/pmb/docs/libs/iplib/ipolates.html) of National Centers for Environmental Prediction (NCEP). The sand and clay content were rechecked to make sure that their sum was less than 98% (we assumed the silt fraction is always $>2\%$). If the sum of sand and clay was greater than 98% after interpolation, the sand and clay content are adjusted by subtracting the quantity $((\text{sand}+\text{clay})-98)/2.0$. The organic matter density was constrained by a maximum possible value of 0,130 kg/m^3 [*Oleson et al.*, 2010]. The details of the input data preparation can be found in *Han et al.* [2012].

CLM Layer	Depth (cm)
1	1.75
2	4.51
3	9.06
4	16.56
5	28.91
6	49.29
7	82.89
8	138.28
9	229.61
10	380.19

Figure 2. CLM soil layer depth distribution.

The synthetic L-band brightness temperature observations were calculated using the CMEM model and the soil moisture and soil temperature data from the reference run of CLM. The synthetic brightness temperatures were perturbed using spatially correlated noise derived from a spatially correlated Gaussian random field with mean 0.0 K and an exponential semivariogram model with nugget 0.0 K^2, variance 4.0 K^2 and a range of 10 km. The variance value was based on literature value [*Entekhabi et al.*, 2010]. After the generation of the

Table 1. Summary of Perturbation Parameters for Atmospheric Forcing Data and the Cross-Correlation Coefficients Used to Model the Correlations Among the Perturbations for the Different Variables

Variables	Noise	Standard Deviation	Spatial Correlation Scale	Forcing Cross Correlation
Precipitation	Multiplicative	0.5	10 km	1.0, −0.8, 0.5, 0.0,
Shortwave radiation	Multiplicative	0.3	10 km	−0.8, 1.0, −0.5, 0.4,
Longwave radiation	Additive	20 W/m^2	10 km	0.5, −0.5, 1.0, 0.4,
Air temperature	Additive	1 K	10 km	0.0, 0.4, 0.4, 1.0

correlated Gaussian field, only the observations located in the nonforested areas were selected to be used in assimilation.

The spatial resolution we used for the measured synthetic L-band brightness temperature is higher than that of the common passive microwave sensors such as SMOS and SMAP, which would be too coarse to be applied at the catchment scale. We would only obtain few grid cells of observation data from the coarse passive microwave sensors and would not catch the spatial heterogeneity of the soil moisture distribution. Recently, various downscaling approaches [*Mascaro et al*., 2010; *Merlin et al*., 2013; *Piles et al*., 2011] have been proposed to downscale the coarse microwave data to the scale of several kilometers using high spatial resolution remote sensing data (e.g., MODIS). Similarly, several multiscale data assimilation approaches have been proposed in the literature, where observations measured at a certain scale are assimilated into a model at another scale [*Montzka et al*., 2012]. This was the motivation to assimilate higher-resolution downscaled L-band brightness temperature data in this study.

3.3. Data Assimilation Experiments

The OSSE was designed to assimilate synthetic L-Band brightness temperature data into CLM using a combination of the CMEM model and the LETKF algorithm. The assimilation cycle mimics the ascending mode passing time of the SMOS-satellite and is therefore carried out every 3 days at 06:00Z from 1 April 2010 to 30 June 2010. At these overpassing time steps, brightness temperature is assimilated into the model and the radiative transfer model CMEM links measured brightness temperature and soil moisture content. From 1 July to 30 September, the CLM was driven by the updated soil properties without data assimilation.

Representing the different sources of uncertainty that affect model prediction is very important. There are three main sources of uncertainties: (1) uncertainty in model parameters, (2) uncertainty in model forcing data, and (3) model structure uncertainty including the formulation of the PTFs. The model structure uncertainty is hard to quantify, and therefore in this study, the model prediction uncertainty is represented only by uncertain model parameters and model forcing data. It is acknowledged that results will be affected by neglecting model structural uncertainty and that results will be overoptimistic. However, this is one of the first studies on the joint calibration of distributed states and parameters of a land surface model, using sequential data assimilation, and it therefore provides a first indication of its feasibility. Sand fraction and clay fraction were perturbed by adding a spatially uniform distributed noise in the range of [−10%, +10%], and the range for perturbing organic matter density was [−10 (kg/m^3), +10 (kg/m^3)]. The soil properties (sum of sand and clay content, organic matter density) were also in this case checked and properly adjusted. The multiplicative inflation [*Whitaker and Hamill*, 2012] was applied to the soil properties after each assimilation step to keep the standard deviation of the parameter ensemble members equal to the standard deviation of the uniform distribution with the range of [−10%, +10%] for sand and clay fractions, and [−10 (kg/m^3), +10 (kg/m^3)] for organic matter density.

In order to perturb precipitation, shortwave radiation, longwave radiation, and air temperature, spatially correlated noise was added to these model forcings. The spatial correlated noise was generated using a Fast Fourier Transform approach [*Park and Xu*, 2009], in which the perturbations were kept physically consistent (e.g., a positive perturbation of incoming shortwave radiation is related to a negative perturbation of incoming longwave radiation and a positive perturbation of air temperature) to conserve the atmospheric balance among radiation, clouds, and air temperature [*Reichle et al*., 2010]. The perturbation parameters according to *Reichle et al*. [2007] are summarized in Table 1. The additive and multiplicative perturbations were assumed to be normal distributed and lognormal distributed, respectively. The mean value for the perturbative factors was equal to zero for the additive case and one for the multiplicative case.

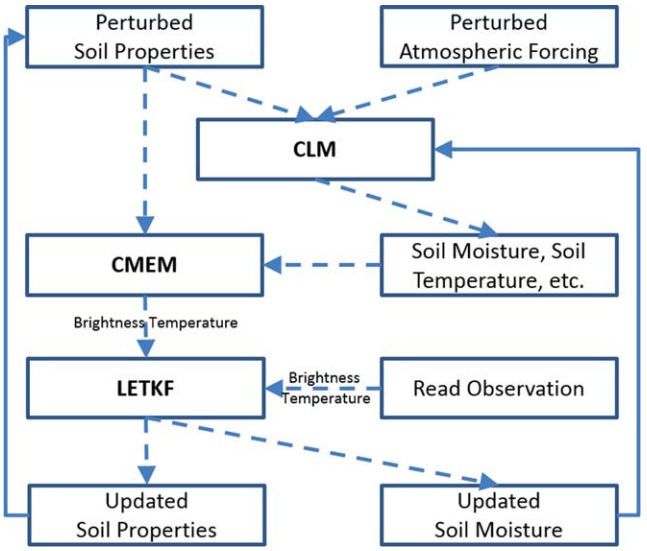

Figure 3. Data assimilation flowchart.

The biased soil properties used in the open loop run and data assimilation runs were generated as follows: the true sand fraction was multiplied by 0.5 and the true clay fraction and organic matter density by 2.0, and these perturbed soil inputs were used as the background value in the open loop run and data assimilation experiments and represent biased a priori information on the soil textural composition. Spatial correlated noise was added to the reference sand fraction, clay fraction and organic matter density. The spatially correlated perturbation was generated by the sequential Gaussian simulation method [*Goovaerts*, 1997] with values in the interval of [−5%, +5%], an exponential semivariogram model with nugget 0.0, variance 100.0, and range 10 km. The range of reference sand fraction is [5%∼80%], for clay fraction it is [3%∼40%], and for organic matter density [1∼130]. The spatial correlation of soil properties can vary considerably for different areas, but a range value that is about 1/10 of the simulation domain is a very typical value often used in simulation studies. The range of the perturbations between −5% and +5% is chosen to introduce a more realistic spatial heterogeneity. In the open loop run and assimilation runs, no spatial heterogeneity of soil properties was considered and the soil properties for each ensemble member were obtained by adding random noise to the homogeneous prior soil properties.

CLM divides the soil properties into 10 layers, but input soil texture and organic matter density cannot be obtained for the 10 layers at the catchment scale. For simplicity, only the first layer of sand fraction, clay fraction, and organic matter density was involved in the parameter estimation. The soil properties for the other (lower) layers were generated based on the a priori ratio of these properties.

The following state and parameter estimation strategies were evaluated and compared:

1. No data assimilation (Open Loop).

2. Data assimilation without soil hydraulic parameter estimation (Only_Assim).

Table 2. Configurations of Different Simulation/Assimilation Scenarios[a]

Scenario Name	Description	Parameter Estimation	Number of Local Observations (State)	Number of Local Observations (for Parameter Estimation)
Open_Loop	Run CLM ensembles only	No	N.A.	N.A.
Only_Assim	Update soil moisture only	No	9	N.A.
Joint_1_Obs	Update soil moisture and soil properties with perturbation of the forcings	Yes	9	1
Joint_5_Obs		Yes	9	5
Joint_9_Obs		Yes	9	9
Joint_16_Obs		Yes	9	16
Only_CLM	Updated sand and clay fractions are only used in CLM	Yes	9	1
Only_CMEM	Updated sand and clay fraction are only used in CMEM	Yes	9	1
Only_Parameter	Update the sand fraction, clay fraction and organic matter density only, no soil moisture update	Yes	9	1

[a]The number of local observations (to be assimilated for each model grid cell) refers to updating soil moisture and soil properties at forest grid cells.

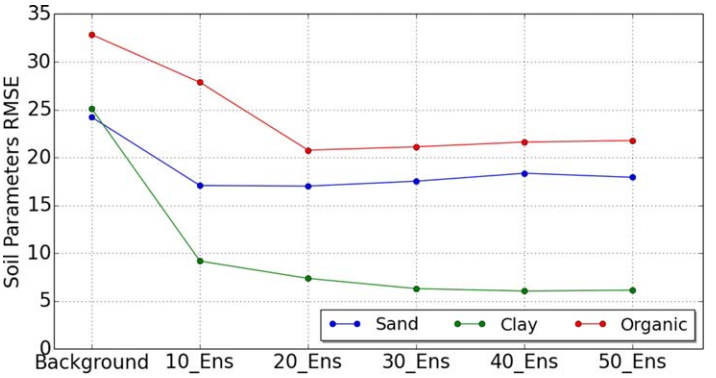

Figure 4. RMSE for sand (%), clay (%), and organic matter density (kg/m^3) for different numbers of ensemble members (scenarios Joint_1_Obs_10, Joint_1_Obs_20, Joint_1_Obs_30, Joint_1_Obs_40, Joint_1_Obs_50), averaged over all grid cells.

3. Joint update of soil moisture and soil properties with perturbation of the forcings. For this scenario, three variants were evaluated:

3a. Optimized sand fraction, clay fraction, and organic matter density were used in CLM and CMEM, in which different number of local observations were assimilated for each assimilation step (Joint_1_Obs, Joint_5_Obs, Joint_9_Obs and Joint_16_Obs).

3b. Optimized sand fraction, clay fraction, and organic matter density were only used in CLM, but not used in CMEM (Only_CLM).

3c. Optimized sand fraction, clay fraction, and organic matter density were only used in CMEM, but not used in CLM (Only_CMEM).

4. Updating soil properties (both in CLM and CMEM) only without soil moisture update (Only_Parameter).

The data assimilation flow layout is illustrated in Figure 3.

The detailed configuration of each scenario is summarized in Table 2.

In order to evaluate the impact of the number of ensemble members on the parameter estimation results, five additional simulation experiments for the scenario of Joint_1_Obs with 10, 20, 30, 40, and 50 ensemble members were evaluated. Figure 4 shows the RMSE for the soil properties for the scenarios of joint estimation with one local observation and different numbers of ensemble members. With more ensemble members, the estimation of the clay fraction is improved further, but the characterization of the sand fraction and organic matter density worsened. The RMSE values of latent heat flux and sensible heat flux for the assimilation period and the verification period (Figure 5) illustrate that simulation results with more ensemble members improved compared to simulations with less ensemble members. Therefore, 50 ensemble members were used in this study, balancing simulation accuracy and computing efficiency.

Other land data assimilation studies used often a smaller number of ensemble members. Twelve ensemble members were used by *Kumar et al*. [2009] and *Reichle et al*. [2010] and 20 members by *De Lannoy et al*. [2012] and *Pan and Wood* [2010]. If the number of model grid cells is large, computational efficiency and hard disk space requirements can become prohibitive as time series of all grid cells and all ensemble members need to be stored for the post analysis.

4. Results

The results for the different simulation experiments were evaluated by the Root Mean Square Error (RMSE), calculated for soil moisture and soil properties. The RMSE value for each grid cell over the complete time series (and also separately for the verification and assimilation period) was calculated (this includes the 3 months assimilation period and 3 months verification period), and in addition, the average RMSE value over all grid cells was determined according to:

$$\mathrm{RMSE}=\sqrt{(1/M)\sum_{m=1}^{M}(1/N)\sum_{i=1}^{N}(\overline{\theta}_{i,m}-\theta_{i,m})^2} \tag{34}$$

where $\overline{\theta}_{i,m}$ is the ensemble mean (either for open loop or data assimilation experiments), $\theta_{i,m}$ corresponds to the reference run, N is the number of time steps (4392 in this study). For the soil properties, we only

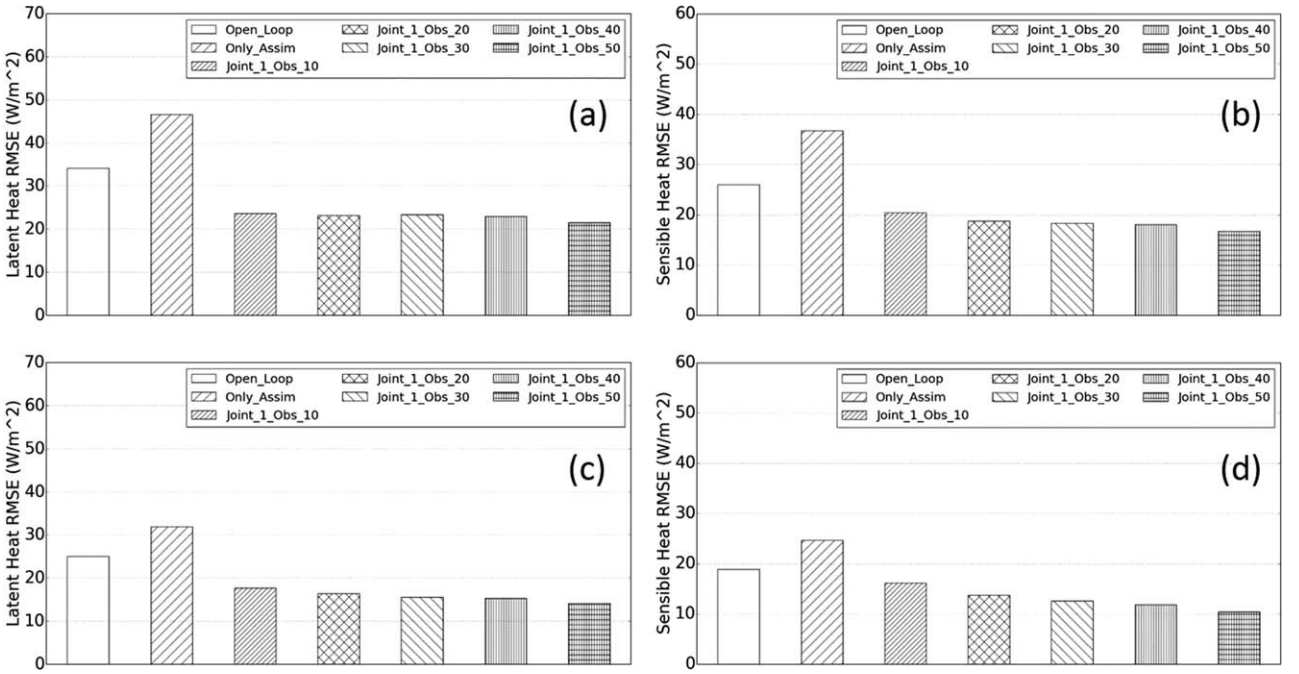

Figure 5. RMSE values of (left) latent heat flux and sensible heat flux (right) for open loop simulation (Open_Loop), single assimilation scenario (Only_Assim), and different numbers of ensemble members (scenarios Joint_1_Obs_10, Joint_1_Obs_20, Joint_1_Obs_30, Joint_1_Obs_40, Joint_1_Obs_50) averaged over (top) assimilation period and (bottom) verification period for all grid cells.

analyzed parameter values at the final assimilation step, so N is 1. M is the number of active grid cells (4340). The smaller the RMSE value is, the better the assimilation results are.

Figure 6 shows the RMSE for the soil properties for the scenarios of joint estimation with different number of local observations. It is obvious that the best estimation is for one local observation (Joint_1_Obs). The use of more local observations (Joint_5_Obs, Joint_9_Obs and Joint_16_Obs) did not improve the parameter estimation. In the rest of the paper, the results are discussed for the joint estimation scenario with one local observation (Joint_1_Obs) if not otherwise indicated.

In Figure 7, the RMSE values for the characterization of soil properties are given for the scenarios where soil properties are only updated in CLM (Only_CLM), only updated in CMEM (Only_CMEM), or updated in both CLM and CMEM but without soil moisture update (Only_Parameter). The scenario Only_Parameter gives the best characterization of soil properties. The performance of the scenarios Only_CLM and Only_CMEM is similar.

The estimated spatial distributions of the sand and clay fractions and organic matter density for the first soil layer, estimated by data assimilation combined with parameter estimation, are compared to the true values (reference) and background (open loop) values for the scenario Joint_1_Obs in Figure 8. The sand fraction increased during data assimilation, whereas clay fraction and organic matter density values decreased during data assimilation for most of the area. Assimilation of brightness temperature improved the estimates of all soil properties with values much closer to the reference values. The estimates of the soil properties

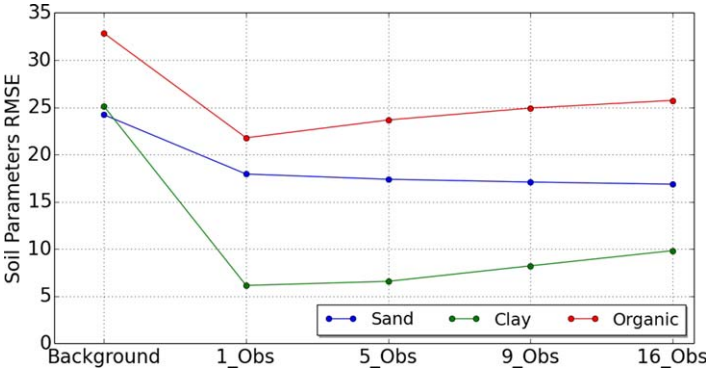

Figure 6. RMSE for sand (%), clay (%), and organic matter density (kg/m^3) for different amounts of assimilated observations for the joint estimation scenario.

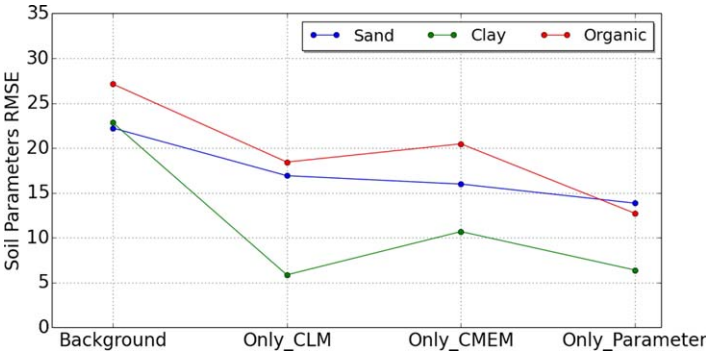

Figure 7. RMSE for sand (%), clay (%), and organic matter density (kg/m^3) for different scenarios where soil properties are only updated for CLM (Only_CLM), only updated for CMEM (Only_CMEM), or soil properties are estimated without soil moisture update (Only_Parameter).

(sand fraction and clay fraction) for the forest area also improved, although no direct observations for these grid cells were available, by taking profit of the assimilation of surrounding local observations by LETKF.

As an example, for one selected grid cell in the study area, the ensemble of temporal evolutions of estimated sand and clay fractions, organic matter density, soil hydraulic parameters *B*, and saturated hydraulic conductivity (for the first soil layer) are plotted. The reference and background values and rainfall are also indicated in Figure 9. The mean of the ensemble members approaches the true sand and clay content as well as the true soil organic matter density after 1 month of assimilation (i.e., after the first 10

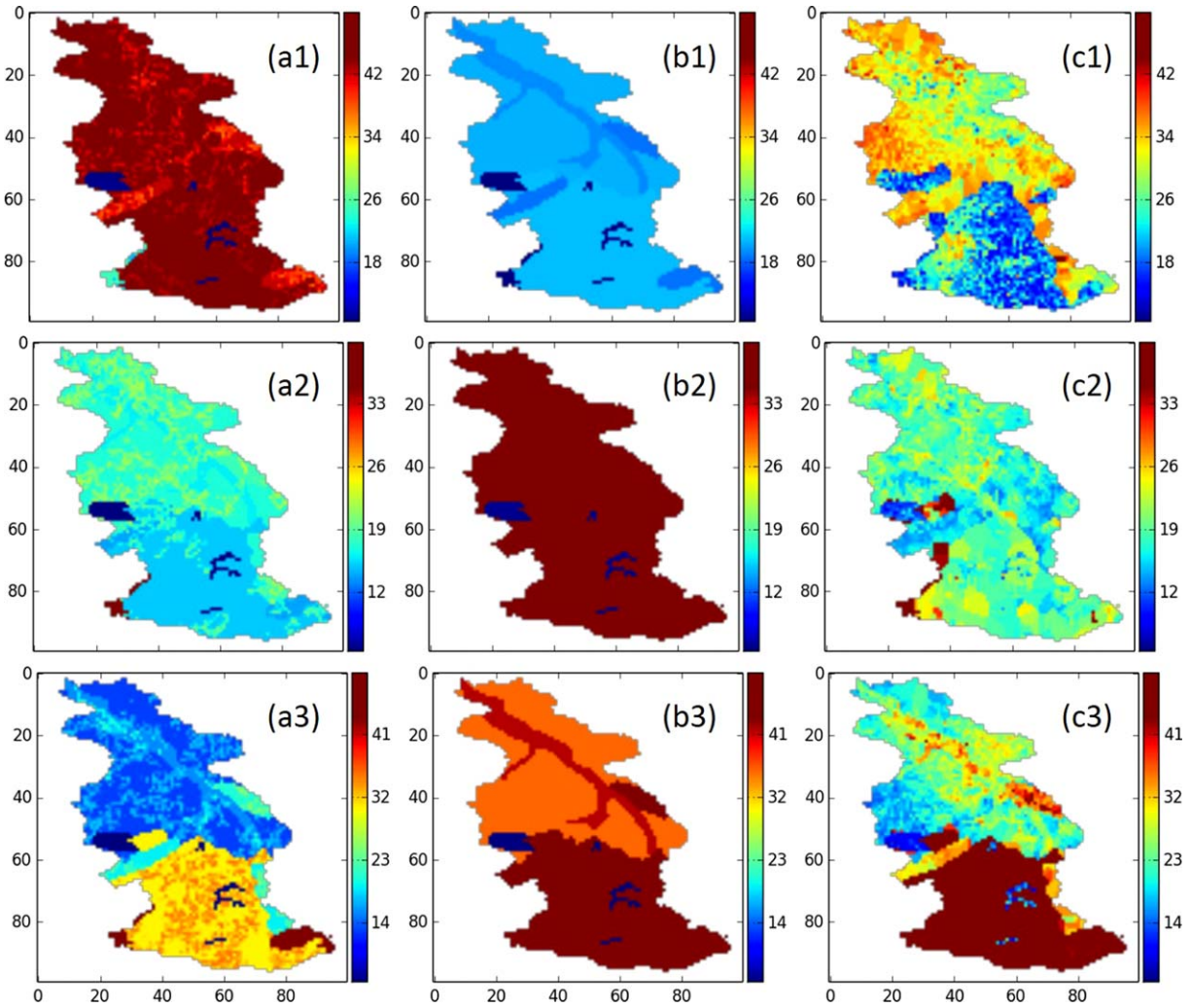

Figure 8. (top) Sand fraction (%), (middle) clay fraction (%), and (bottom) organic matter density (kg/m^3) for the first soil layer. The true values are given in a1–a3, the background values in b1–b3, and the results for the scenario Joint_1_Obs in c1–c3.

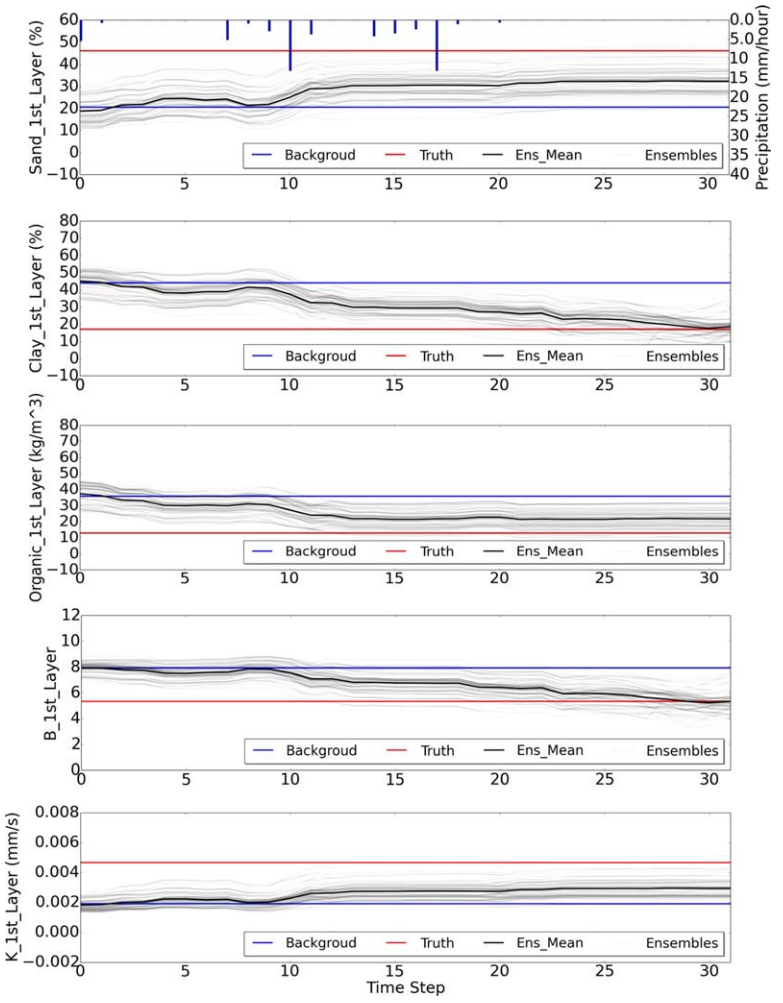

Figure 9. Sand fraction (%), clay fraction (%), organic matter density (kg/m^3), soil hydraulic parameter *B*, and saturated hydraulic conductivity (mm/s) for the upper soil layer for the reference, background, ensemble mean, and 50 ensemble members. Results are given for the scenario Joint_1_Obs at a crop land location (50.95°N, 6.2°E).

assimilation steps). The optimal parameter values are those for which the ensemble mean values are close to the true value while keeping the spread of the ensemble large enough. A too narrow spread of the soil parameter ensemble may result in filter divergence. The figures illustrate that covariance inflation helped to keep the adequate ensemble spread. The evolution of the parameter values indicates that especially the first rainfall events were helpful to improve parameter estimates. The CLM parameters *B* and saturated hydraulic conductivity for mineral soil calculated by the equations (10) and (11), also show a clear improvement.

The RMSE values of soil moisture and soil temperature for all scenarios at 5, 10, 30, and 50 cm depth over assimilation period and verification period are plotted in Figures 10 and 11, respectively. Figure 10 indicates that the assimilation of L-band brightness temperature data improves the soil moisture estimation. The results with parameter estimation are better than those obtained without parameter estimation, with best results for the scenario Only_Parameter. For the verification period, soil moisture bias was again increased for the scenario of Only_Assim, but if parameters were calibrated, soil moisture bias was reduced for the verification period (Figure 11), except for the scenario Only_CMEM.

The results for soil temperature indicate as well an improvement although less pronounced as for soil moisture. The soil temperature is related to the soil moisture through the heat fluxes in CLM. The scenario of Only_Assim resulted in worse latent heat flux, sensible heat flux, and soil heat flux, as compared to open loop simulations and these biased heat fluxes affected the soil temperature simulation [*Oleson et al.*, 2013]. An improvement of the simulated land surface fluxes is also expected to improve the temperature simulation. If the optimized soil properties were only used in CLM (Only_CLM) or only in CMEM (Only_CMEM), the updated soil moisture contents were better than for Only_Assim, but worse than for joint estimation scenarios where soil properties were both updated in CLM and CMEM. For the characterization of soil temperature, optimized soil properties for the scenarios Only_CLM and Only_CMEM did not result in as good results as for joint estimation scenarios.

Figure 12 compares the latent, sensible, and soil heat flux with the equivalents from the reference fields for the assimilation period and verification period separately. The results obtained for data assimilation without

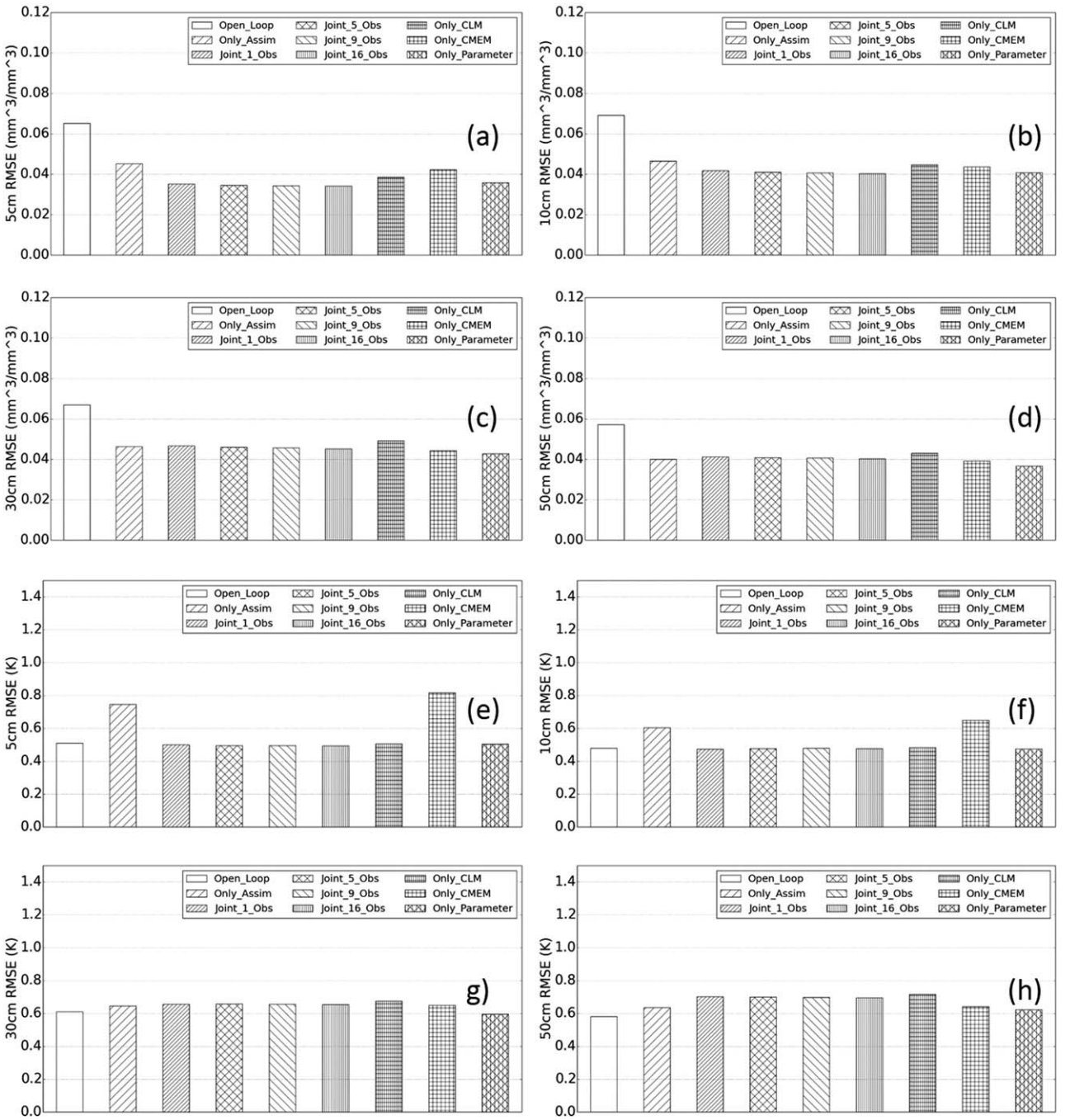

Figure 10. RMSE values of (rows 1 and 2) soil moisture and (rows 3 and 4) soil temperature at (a and e) 5 cm, (b and f) 10 cm, (c and g) 30 cm, and (d and h) 50 cm depth for all scenarios and averaged over the assimilation period and all grid cells.

parameter estimation (Only_Assim) are worse than for open loop simulations; all surface fluxes clearly deteriorated after assimilation without parameter estimation. This can be explained by the fact that soil properties are biased, and now only states are updated, which in combination with the erroneous soil properties can result in a worse estimate of the evaporative flux and other fluxes. Updating soil moisture (without updating parameters) in case of biased or incorrect soil properties (the common situation in land surface models) will in many cases not improve at all the characterization of the evaporative flux. Latent and sensible heat flux estimation did not deteriorate in case soil properties were updated together with the states, and were in this case also slightly better than the open loop simulations. We think that the relatively low

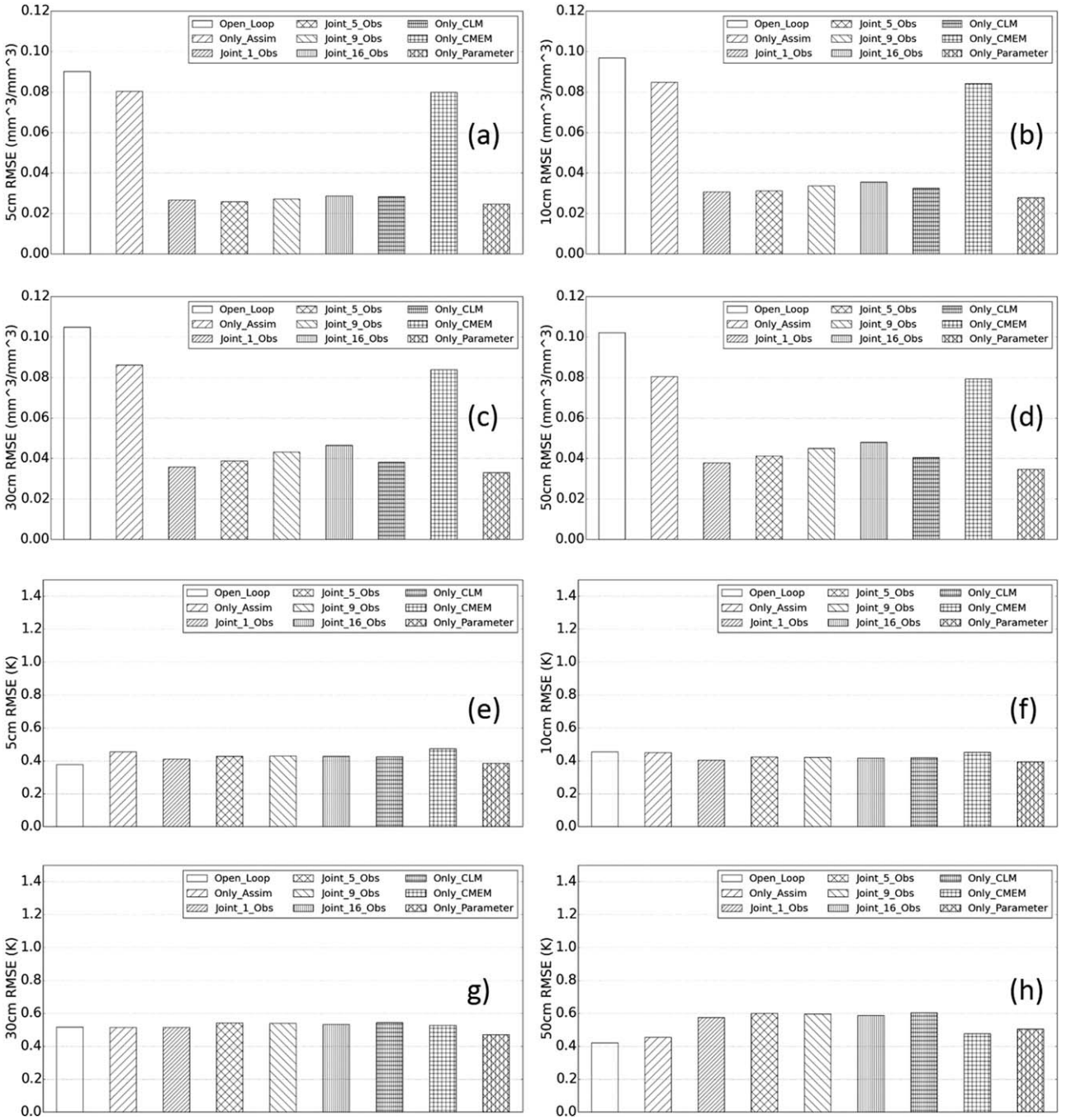

Figure 11. RMSE values of (rows 1 and 2) soil moisture and (rows 3 and 4) soil temperature at (a and e) 5 cm, (b and f) 10 cm, (c and g) 30 cm, and (d and h) 50 cm depth for all scenarios and averaged over the verification period and all grid cells.

RMSE values for the open loop run can also explain the limited improvement achieved for this scenario. In reality, the RMSE of latent heat flux and sensible heat flux could be larger than the values found here in this synthetic study. In this synthetic study, only the soil properties and the model forcings were uncertain, and simulation errors of latent heat flux and sensible heat flux were relatively small. Uncertainty concerning precipitation could be larger and also uncertainty with respect to leaf area index is important for soil moisture and land fluxes simulation. Additional bias can be included in these two inputs to increase the open loop deviation from the truth. Assimilation of soil moisture data might result then, in reality, in larger improvements of the latent and sensible heat flux estimates. A high clay fraction was used in the open loop runs,

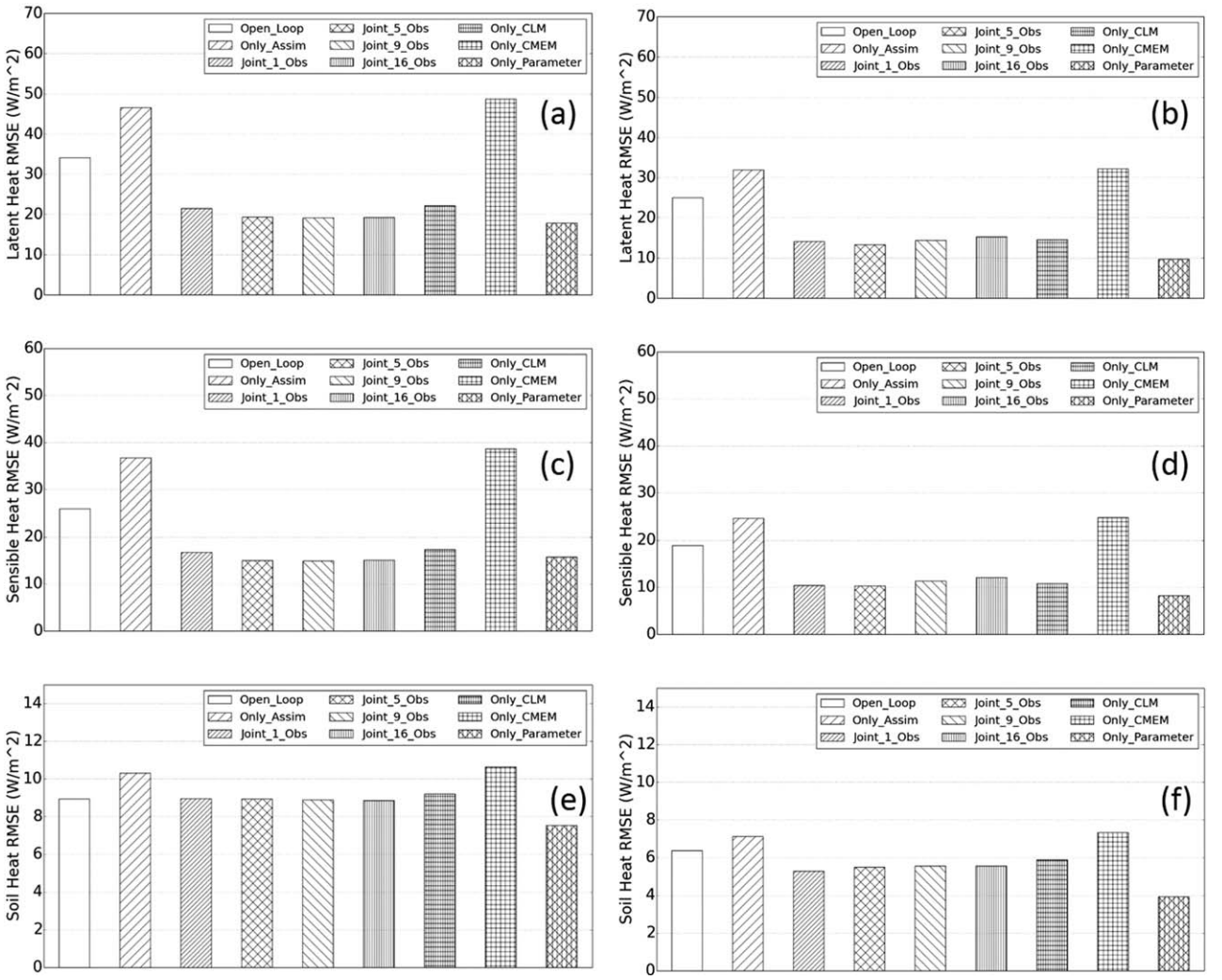

Figure 12. RMSE values of (top) latent heat flux, (middle) sensible heat flux, and (bottom) soil heat flux for all scenarios and averaged over the (left) assimilation period and (right) verification period for all grid cells.

and this will decrease the latent heat flux [*Schwinger et al*., 2010]. The soil moisture distribution is controlled by the hydraulic conductivity *K* (equation (1)), which is derived from the empirical parameter *B* of Clapp–Hornberger parameterizations (equation (10)). The clay fraction affects the calculation of *B* [*Oleson et al*., 2013; *Sakaguchi and Zeng*, 2009]. These two conditions can compensate each other and result in low RMSE values for the latent and sensible heat fluxes in the open loop runs.

The results illustrate that the scenario Only_CLM resulted in better estimates of surface fluxes than Only_Assim, and the difference with the joint estimation scenarios, where soil properties in both CLM and CMEM are updated, is small. But if the soil properties were only updated in CMEM (Only_CMEM), the estimation of the fluxes estimation was worse than for all other scenarios. As a summary, the best results are obtained if soil properties are both updated in CLM and CMEM. Updating soil properties in CLM alone or CMEM alone yields worse results as compared to updating in both models. The Only_Parameter scenario gave also here the best results.

The mean soil moisture content for the Rur catchment over the verification period at depths of 10, 30, and 50 cm for the different scenarios is shown in Figure 13. The spatial distribution of evapotranspiration over the verification period for the same scenarios is shown in Figure 14. These figures show the improvement of soil moisture and evapotranspiration characterization for the whole catchment obtained with the joint state and parameter estimation method. The scenario Joint_1_Obs shows a better estimation of soil

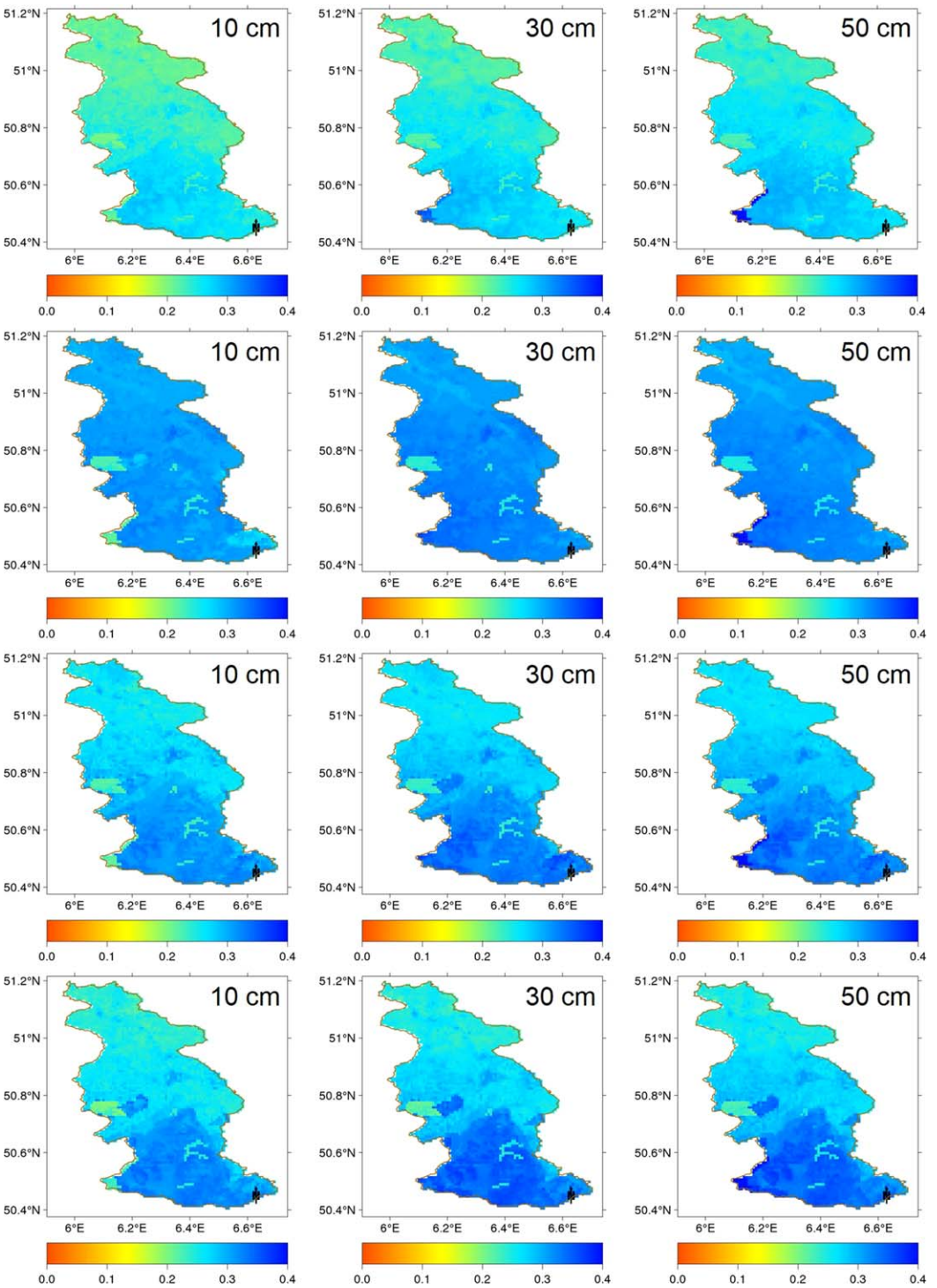

Figure 13. Mean soil moisture content (mm^3/mm^3) for the Rur catchment averaged over the verification period at depths of 10, 30, and 50 cm for the (first row) reference run, (second row) open loop run, (third row) scenario of Only_Assim, and (fourth row) scenario of Joint_1_Obs.

moisture content and evapotranspiration than obtained with the open loop run or the assimilation without parameter updating.

Finally, the soil moisture time series from 1 July to 30 September (validation period only) at 10, 30, and 50 cm depth for different scenarios are illustrated in Figure 15. The estimated soil properties from

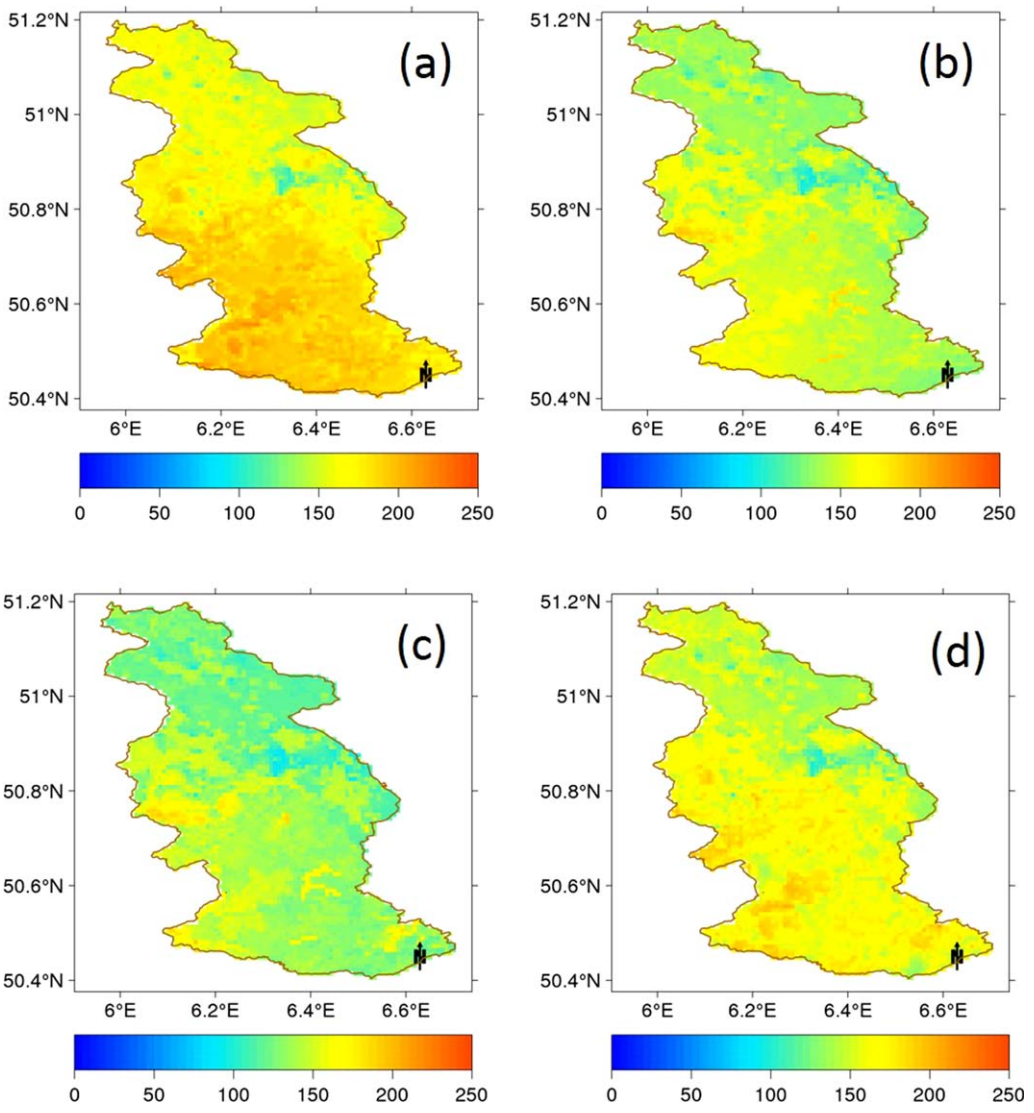

Figure 14. The overall evapotranspiration (mm) averaged over the verification period for (a) reference run, (b) open loop run, (c) the scenario of only assimilation, and (d) the scenario of Joint_1_Obs.

Joint_1_Obs and Only_Parameter were used in CLM for the simulation and no data assimilation was performed. The updated soil properties resulted in good soil moisture simulation results compared with the reference run. The soil moisture will become over time close to the open loop simulation (CLM) for the scenario Only_Assim, where no data assimilation and parameter estimation are performed. The impacts of soil moisture assimilation at 10 cm depth, 30 cm depth, and 50 cm depth on the future simulation could be prolonged for more than half month, 1 and $1\frac{1}{2}$ months, respectively.

5. Additional Discussion

The sand and clay fractions are used in CMEM for calculating soil moisture content from brightness temperature. Usually, the soil moisture product will be firstly retrieved from the passive microwave brightness temperature data and then assimilated. However, if the sand and clay fractions are not defined well (biased), the soil moisture product will be biased. The proposed schedule in this paper to jointly estimate states and parameters allows the direct use of brightness temperature data to infer the optimal sand and clay fractions in the coupled CLM-CMEM model system, i.e., in both CLM and CMEM. This would be a better procedure than the assimilation of the level 2 soil moisture product, which was calculated on the basis of biased sand

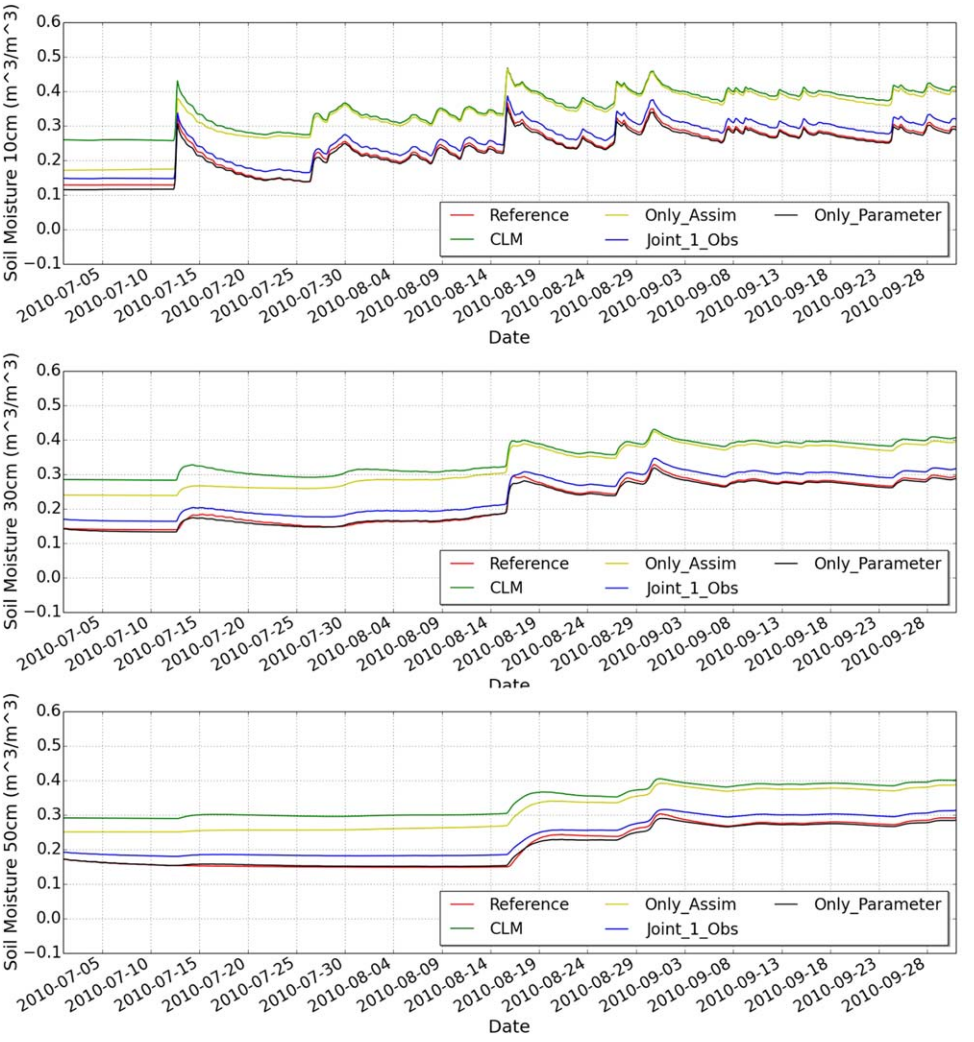

Figure 15. Soil moisture time series from 1 July to 30 September at 10, 30, and 50 cm depth for the reference run (Reference), open loop run (CLM), single assimilation scenario (Only_Assim), join estimation scenario (Joint_1_Obs), and parameter estimation scenario (Only_-Parameter) at a crop land location (6.2°E, 50.95°N).

and clay fractions. The improved soil parameters can also contribute to a better characterization of the land surface fluxes, such as the latent heat flux because the soil evaporation is sensitive to the soil properties. Therefore, the joint state and parameter estimation can be expected to be helpful especially for the brightness temperature assimilation for regions with a low vegetation density where the passive microwave is more informative for the soil moisture.

The soil properties of the complete soil profile were updated together. There are 15 soil layers in CLM, and the remote sensing data can only measure the land surface states, only the surface observation is obtained. In this study, we used a simplified vertical distribution of soil properties. Only information on two soil layers from the global soil database HWSD was available and a linear correlation among the soil properties for the layers was made to assign soil properties to 10 model layers of CLM. In the data assimilation, only soil properties for the first model layer were updated and properties for deeper layers were updated on the basis of a proportionality factor. The updated soil properties of the complete soil profile were used for the simulation of soil moisture, soil temperature, and surface fluxes after assimilation. This simplified scenario might deviate significantly from real-world conditions. If the complex vertical distribution of soil properties is introduced in the parameter estimation, it will be difficult to retrieve the deep layer soil properties using observations from the surface layer because of the limited information content of such observations for the deeper layers. On the other hand, usually we can get prior information from regional soil maps, which can

be helpful for the parameter estimation as it is expected to constrain the soil hydraulic parameter estimates.

The parameter estimation of soil properties limits the degrees of freedom of the parameter space of CLM, i.e., the number of free parameters per grid cell and enhances the stability of the approach. The hydraulic or thermal parameters of CLM are related to sand fraction, clay fraction and organic matter density through the predefined pedotransfer functions in CLM, which we assumed to be known. Parameter estimation results, however, will also be affected by the definition of these pedotransfer functions. In future work, we will focus on the direct estimation of hydraulic and thermal parameters. Moreover, for CLM used in this study, the direct update of soil hydraulic and thermal properties is not straightforward from the technical-computational point of view.

The best results were obtained for the scenario Only_Parameter, in which perturbed atmospheric forcings were used during the parameter estimation step. The results show that successful parameter estimation is possible even though the model forcings were affected by a considerable uncertainty. This is an important result as in reality model forcings (especially precipitation) are always impacted by a large uncertainty, especially for large-scale applications. The results for the scenarios where both states and parameters are updated were slightly worse than for Only_Parameter. The reason is most probably that we only imposed bias with respect to soil properties in this synthetic study and random (unbiased) uncertainty concerning the model forcings. However, in a real-world case, we expect that model uncertainties will come from different sources, including model structural error, and not from soil parameter uncertainty alone. It is expected that under such conditions the joint state-parameter estimation would outperform the updating of parameters alone.

If the updated sand fraction, clay fraction, and organic matter density were only used in CLM, the soil moisture and surface fluxes could be improved, and results were better than for the scenario Only_CMEM, but slightly worse than for the case where soil properties were both updated in CLM and CMEM. If the updated sand fraction and clay fraction were only used in CMEM, soil moisture content did not improve with assimilation. It can be concluded that the joint updating of soil properties in both CLM and CMEM is helpful.

In order to keep ensemble spread during the data assimilation cycles, a multiplicative inflation method was applied after each assimilation step. The results show that this approach allowed maintaining an adequate ensemble spread during the assimilation, although only 50 ensemble members were used. It was found that a narrow range decreased the assimilation performance. A larger initial variance could be an option for this study, but there is no common rule when to decrease/increase the inflation. It was decided to use a constant ensemble spread of soil properties.

In this study, synthetic L-Band brightness temperature data were used instead of real brightness temperature data. Otherwise, the study mimics as much as possible the real-world case with soil properties, land cover and leaf area index taken from available maps. Atmospheric forcing data from a reanalysis product were used. In a real-world case study, this information is also available, but for operational applications weather forecasts instead of reanalysis would have to be used. Therefore, information need in a real-world case study is not systematically different from this synthetic study, but input data would still be more uncertain than assumed in this study (for example, here it was assumed that LAI-data is given and error free) and probably affected by biases also neglected in this study. The leaf area index is an important variable for the surface flux simulation of CLM and brightness temperature modeling of CMEM. In this study, we did not consider the uncertainty of vegetation parameters. The joint estimation of soil and vegetation parameters is an important research topic and will be subject of study in the future. A main difference between this synthetic study and a real-world case study is the spatial resolution of real brightness temperature, which is too coarse to be assimilated at the catchment scale. For example, there are only four SMOS grid cells for our study area. The downscaling of coarse soil moisture products/brightness temperature is still an active research topic at present, and has achieved rapid progress (e.g., a 1 km soil moisture product is being distributed for regions covered by SMOS). We anticipate therefore a worse performance of our assimilation methodology under real-world conditions compared to this synthetic study. However, these results also point to the importance of considering uncertainty of soil properties and estimating soil properties. We expect that results for the assimilation of brightness temperature data could be improved taking this into account.

At the forest regions, the synthetic brightness temperature data were assumed to be unavailable because of the strong vegetation effects on the microwave signal. The local analysis method of LETKF was used to update soil moisture and soil properties for these nonobserved grid cells using the surrounding correlated observations from less vegetated regions like grassland and farmland. In real cases, the assimilation for the forested regions will however become more complex related to sudden, sharp transitions in soil moisture content and soil properties at the border of different land use types. Moreover, the local analysis method strongly depends on the spatial correlation characteristics. The improvements for the forested area can be observed from Figures 8, 13, and 14.

Many studies have proven that there is an optimal correlation length for data assimilation. Beyond this length scale, the analysis error will increase related to spurious covariances between observations and the estimation point. If many local observations are used, the system might be dominated by spurious observation increments that prevent it from converging to the truth [*Greybush et al*., 2011; *Lien et al*., 2013; *Miyoshi and Yamane*, 2007]. The optimal correlation length is case dependent; in this case, results indicate it may be very short.

6. Conclusions

A joint state and parameter estimation method based on an augmented state vector approach was implemented for the land surface model CLM version 4.5. This approach was tested in a synthetic study that mimics the Rur catchment in Western Germany. Synthetic brightness temperature data were assimilated by LETKF in the coupled CLM-CMEM model and used to update soil moisture and temperature as well as sand and clay fractions and soil organic matter density. The estimated sand and clay fractions and organic matter density converge after a series of updates (in this synthetic study about 10 updates) almost to the true values in case the ensemble spread of the ensemble is maintained large enough with help of covariance inflation. Updates of soil hydraulic and thermal parameters (such as the *B* parameter and saturated hydraulic conductivity) were obtained by updating soil textural properties and soil organic matter density in CLM. A key finding is that accurate estimates of soil hydraulic and thermal parameters are needed to obtain accurate estimates of soil moisture and surface energy fluxes. The single assimilation of brightness temperature without parameter updates resulted in a worse characterization of surface fluxes (compared to open loop simulation) if soil properties were biased/inaccurate. The update of soil properties in both CLM and CMEM resulted in a better characterization of soil moisture, temperature, and land surface fluxes than the update of soil properties in CLM (or CMEM) alone.

The soil properties in the forest region, where no reliable brightness temperature measurements were available, could also be improved with help of correlation to neighboring nonforest grid cells and surrounding L-band brightness temperature data of nonforest grid cells.

Acknowledgments
This work was supported by the Transregional Collaborative Research Centre 32, Agadapt (Climate-KIC) and the NSFC project (grant: 41271357, 91125001). The support of the supercomputing facilities of Forschungszentrum Julich (JUROPA) is gratefully acknowledged. The data used in this paper can be ordered from the authors.

References

Bateni, S. M., and D. Entekhabi (2012), Surface heat flux estimation with the ensemble Kalman smoother: Joint estimation of state and parameters, *Water Resour. Res.*, *48*, W08521, doi:10.1029/2011WR011542.

Beven, K., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*(3), 279–298.

Bogena, H. R., M. Herbst, J. A. Huisman, U. Rosenbaum, A. Weuthen, and H. Vereecken (2010), Potential of wireless sensor networks for measuring soil water content variability, *Vadose Zone J.*, *9*(4), 1002–1013.

Brocca, L., F. Melone, T. Moramarco, W. Wagner, V. Naeimi, Z. Bartalis, and S. Hasenauer (2010), Improving runoff prediction through the assimilation of the ASCAT soil moisture product, *Hydrol. Earth Syst. Sci.*, *14*(10), 1881–1893.

Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic properties, *Water Resour. Res.*, *14*(4), 601–604.

Cosby, B. J., G. M. Hornberger, R. B. Clapp, and T. R. Ginn (1984), A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils, *Water Resour. Res.*, *20*(6), 682–690.

Dai, Y., W. Shangguan, Q. Duan, B. Liu, S. Fu, and G. Niu (2013), Development of a China dataset of soil hydraulic parameters using pedotransfer functions for land surface modeling, *J. Hydrometeorol.*, *14*(3), 869–887.

de Rosnay, P., M. Drusch, A. Boone, G. Balsamo, B. Decharme, P. Harris, Y. Kerr, T. Pellarin, J. Polcher, and J. P. Wigneron (2009), AMMA land surface model intercomparison experiment coupled to the community microwave emission model: ALMIP-MEM, *J. Geophys. Res.*, *114*, D05108, doi:10.1029/2008JD010724.

De Lannoy, G. J. M., R. H. Reichle, K. R. Arsenault, P. R. Houser, S. Kumar, N. E. C. Verhoest, and V. R. N. Pauwels (2012), Multiscale assimilation of Advanced Microwave Scanning Radiometer-EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado, *Water Resour. Res.*, *48*, W01522, doi:10.1029/2011WR010588.

Dobson, M. C., F. T. Ulaby, M. T. Hallikainen, and M. A. Elrayes (1985), Microwave dielectric behavior of wet soil, 2. Dielectric mixing models, *IEEE Trans. Geosci. Remote Sens.*, *23*(1), 35–46.

Drusch, M. (2007), Initializing numerical weather prediction models with satellite-derived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set, *J. Geophys. Res.*, *112*, D03102, doi:10.1029/2006JD007478.
Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*(4), 1015–1031.
Duan, Q., et al. (2006), Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, *J. Hydrol.*, *320*(1-2), 3–17.
Entekhabi, D., et al. (2010), The Soil Moisture Active Passive (SMAP) mission, *Proc. IEEE*, *98*(5), 704–716.
FAO, IIASA, ISRIC, ISSCAS, and JRC (2010), Harmonized World Soil Database, Version 1.1, Food and Agric. Organ., Rome.
Franssen, H. J. H., and W. Kinzelbach (2008), Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem, *Water Resour. Res.*, *44*, W09408, doi:10.1029/2007WR006505.
Goovaerts, P. (1997), *Geostatistics for Natural Resources Evaluation*, Oxford Univ. Press, Oxford, U. K.
Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt (2011), Balance and ensemble Kalman filter localization techniques, *Mon. Weather Rev.*, *139*(2), 511–522.
Gutmann, E. D., and E. E. Small (2007), A comparison of land surface model soil hydraulic properties estimated by inverse modeling and pedotransfer functions, *Water Resour. Res.*, *43*, W05418, doi:10.1029/2006WR005135.
Han, X., X. Li, H. J. Hendricks Franssen, H. Vereecken, and C. Montzka (2012), Spatial horizontal correlation characteristics in the land data assimilation of soil moisture, *Hydrol. Earth Syst. Sci.*, *16*(5), 1349–1363.
Holmes, T. R. H., M. Drusch, J. P. Wigneron, and R. A. M. de Jeu (2008), A global simulation of microwave emission: Error structures based on output from ECMWF's operational integrated forecast system, *IEEE Trans. Geosci. Remote Sens.*, *46*(3), 846–856.
Huang, M., and X. Liang (2006), On the assessment of the impact of reducing parameters and identification of parameter uncertainties for a hydrologic model with applications to ungauged basins, *J. Hydrol.*, *320*(1-2), 37–61.
Hunt, B. R., E. J. Kostelich, and I. Szunyogh (2007), Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Physica D*, *230*(1-2), 112–126.
Jazwinski, A. H. (1970), *Stochastic Processes and Filtering Theory*, Academic Press, San Diego, Calif.
Kerr, Y. H., et al. (2010), The SMOS mission: New tool for monitoring key elements of the global water cycle, *Proc. IEEE*, *98*(5), 666–687.
Kumar, S. V., R. H. Reichle, R. D. Koster, W. T. Crow, and C. D. Peters-Lidard (2009), Role of subsurface physics in the assimilation of surface soil moisture observations, *J. Hydrol.*, *10*(6), 1534–1547.
Kurtz, W., H.-J. Hendricks Franssen, and H. Vereecken (2012), Identification of time-variant river bed properties with the ensemble Kalman filter, *Water Resour. Res.*, *48*, W10534, doi:10.1029/2011WR011743.
Lawrence, D. M., and A. G. Slater (2007), Incorporating organic soil into a global climate model, *Clim. Dyn.*, *30*(2-3), 145–160.
Li, C., and L. Ren (2011), Estimation of unsaturated soil hydraulic parameters using the ensemble Kalman filter, *Vadose Zone J.*, *10*(4), 1205–1227.
Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res.*, *99*(D7), 14,415–14,428.
Lien, G. Y., E. Kalnay, and T. Miyoshi (2013), Effective assimilation of global precipitation: simulation experiments, *Tellus A*, *65*(0), 19915, doi:10.3402/tellusa.v65i0.19915.
Liu, Y., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour. Res.*, *43*, W07401, doi:10.1029/2006WR005756.
Liu, Y., et al. (2012), Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities, *Hydrol. Earth Syst. Sci.*, *16*(10), 3863–3887.
Liu, Y. Y., R. M. Parinussa, W. A. Dorigo, R. A. M. De Jeu, W. Wagner, A. I. J. M. van Dijk, M. F. McCabe, and J. P. Evans (2011), Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, *Hydrol. Earth Syst. Sci.*, *15*(2), 425–436.
Livneh, B., and D. P. Lettenmaier (2012), Multi-criteria parameter estimation for the Unified Land Model, *Hydrol. Earth Syst. Sci.*, *16*(8), 3029–3048.
Mascaro, G., E. R. Vivoni, and R. Deidda (2010), Downscaling soil moisture in the southern Great Plains through a calibrated multifractal model for land surface modeling applications, *Water Resour. Res.*, *46*, W08546, doi:10.1029/2009WR008855.
Merlin, O., M. J. Escorihuela, M. A. Mayoral, O. Hagolle, A. Al Bitar, and Y. Kerr (2013), Self-calibrated evaporation-based disaggregation of SMOS soil moisture: An evaluation study at 3 km and 100 m resolution in Catalunya, Spain, *Remote Sens. Environ.*, *130*, 25–38.
Miyoshi, T., and S. Yamane (2007), Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution, *Mon. Weather Rev.*, *135*(11), 3841–3861.
Montzka, C., M. Canty, R. Kunkel, G. Menz, H. Vereecken, and F. Wendland (2008), Modelling the water balance of a mesoscale catchment basin using remotely sensed land cover data, *J. Hydrol.*, *353*(3-4), 322–334.
Montzka, C., H. Moradkhani, L. Weihermüller, H.-J. H. Franssen, M. Canty, and H. Vereecken (2011), Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, *J. Hydrol.*, *399*(3-4), 410–421.
Montzka, C., V. R. Pauwels, H. J. Franssen, X. Han, and H. Vereecken (2012), Multivariate and multiscale data assimilation in terrestrial systems: A review, *Sensors*, *12*(12), 16291–16333.
Montzka, C., J. P. Grant, H. Moradkhani, H. J. H. Franssen, L. Weihermuller, M. Drusch, and H. Vereecken (2013a), Estimation of radiative transfer parameters from L-band passive microwave brightness temperatures using advanced data assimilation, *Vadose Zone J.*, *12*(3), doi:10.2136/vzj2012.0040.
Montzka, C., et al. (2013b), Brightness temperature and soil moisture validation at different scales during the SMOS validation campaign in the Rur and Erft Catchments, Germany, *IEEE Trans. Geosci. Remote Sens.*, *51*(3), 1728–1743.
Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. R. Houser (2005a), Dual state–parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.*, *28*(2), 135–147.
Moradkhani, H., K.-L. Hsu, H. Gupta, and S. Sorooshian (2005b), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, *41*, W05012, doi:10.1029/2004WR003604.
Nie, S., J. Zhu, and Y. Luo (2011), Simultaneous estimation of land surface scheme states and parameters using the ensemble Kalman filter: Identical twin experiments, *Hydrol. Earth Syst. Sci.*, *15*(8), 2437–2457.
Niu, G.-Y., et al. (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.*, *116*, D12109, doi:10.1029/2010JD015139.
Oleson, K., et al. (2010), Technical description of version 4.0 of the Community Land Model (CLM), *NCAR Tech. Note NCAR/TN-478+STR*, 257 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.

Oleson, K., et al. (2013), Technical description of version 4.5 of the Community Land Model (CLM), *NCAR Tech.l Note NCAR/TN-503+STR*, 422 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.
Oliva, R., E. Daganzo-Eusebio, Y. H. Kerr, S. Mecklenburg, S. Nieto, P. Richaume, and C. Gruhier (2012), SMOS radio frequency interference scenario: Status and actions taken to improve the RFI environment in the 1400–1427-MHz passive band, *IEEE Trans. Geosci. Remote Sens.*, *50*(5), 1427–1439.
Pan, M., and E. F. Wood (2010), Impact of accuracy, spatial availability, and revisit time of satellite-derived surface soil moisture in a multi-scale ensemble data assimilation system, *IEEE J. STARS*, *3*(1), 49–56, doi:10.1109/JSTARS.2010.2040585.
Park, S. K., and L. Xu (2009), *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, Springer, Berlin/Heidelberg.
Parrens, M., J. C. Calvet, P. de Rosnay, and B. Decharme (2014), Benchmarking of L-band soil microwave emission models, *Remote Sens. Environ.*, *140*, 407–419.
Piles, M., A. Camps, M. Vall-llossera, I. Corbella, R. Panciera, C. Rudiger, Y. H. Kerr, and J. Walker (2011), Downscaling SMOS-derived soil moisture using MODIS visible/infrared data, *IEEE Trans. Geosci. Remote Sens.*, *49*(9), 3156–3166.
Reichle, R. H., R. D. Koster, P. Liu, S. P. P. Mahanama, E. G. Njoku, and M. Owe (2007), Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), *J. Geophys. Res.*, *112*, D09108, doi:10.1029/2006JD008033.
Reichle, R. H., S. V. Kumar, S. P. P. Mahanama, R. D. Koster, and Q. Liu (2010), Assimilation of satellite-derived skin temperature observations into land surface models, *J. Hydrometeorol.*, *11*(5), 1103–1122.
Rigon, R., G. Bertoldi, and T. M. Over (2006), GEOtop: A distributed hydrological model with coupled water and energy budgets, *J. Hydrometeorol.*, *7*(3), 371–388.
Rodell, M., et al. (2004), The global land data assimilation system, *Bull. Am. Meteorol. Soc.*, *85*(3), 381–394.
Sabater, J. M., A. Fouilloux, and P. de Rosnay (2012), Technical implementation of SMOS data in the ECMWF integrated forecasting system, *IEEE Geosci. Remote Sens. Lett.*, *9*(2), 252–256.
Sakaguchi, K., and X. Zeng (2009), Effects of soil wetness, plant litter, and under-canopy atmospheric stability on ground evaporation in the Community Land Model (CLM3.5), *J. Geophys. Res.*, *114*, D01107, doi:10.1029/2008JD010834.
Scheerlinck, K., V. R. N. Pauwels, H. Vernieuwe, and B. De Baets (2009), Calibration of a water and energy balance model: Recursive parameter estimation versus particle swarm optimization, *Water Resour. Res.*, *45*, W10422, doi:10.1029/2009WR008051.
Schöniger, A., W. Nowak, and H. J. Hendricks Franssen (2012), Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography, *Water Resour. Res.*, *48*, W04502, doi:10.1029/2011WR010462.
Schwinger, J., S. J. Kollet, C. M. Hoppe, and H. Elbern (2010), Sensitivity of latent heat fluxes to initial values and parameters of a land-surface model, *Vadose Zone J.*, *9*(4), 984–1001.
Sellers, P. J., D. A. Randall, G. J. Collatz, J. A. Berry, C. B. Field, D. A. Dazlich, C. Zhang, G. D. Collelo, and L. Bounoua (1996), A revised land surface parameterization (SiB2) for atmospheric GCMS, Part I: Model formulation, *J. Clim.*, *9*(4), 676–705.
Šimůnek, J., M. T. van Genuchten, and M. Šejna (2008), Development and applications of the HYDRUS and STANMOD software packages and related codes, *Vadose Zone J.*, *7*(2), 587–600.
Song, C., L. Jia, and M. Menenti (2014), Retrieving high-resolution surface soil moisture by downscaling AMSR-E brightness temperature using MODIS LST and NDVI data, *IEEE J. STARS*, *7*(3), 935–942.
Sun, W. X., S. L. Liang, G. Xu, H. L. Fang, and R. Dickinson (2008), Mapping plant functional types from MODIS data using multisource evidential reasoning, *Remote Sens. Environ.*, *112*(3), 1010–1024.
Swenson, S. C., D. M. Lawrence, and H. Lee (2012), Improved simulation of the terrestrial hydrological cycle in permafrost regions by the Community Land Model, *J. Adv. Model. Earth Syst.*, *4*, M08002, doi:10.1029/2012MS000165.
Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, *37*(10), 2521–2535.
Vereecken, H., M. Weynants, M. Javaux, Y. Pachepsky, M. G. Schaap, and M. T. V. Genuchten (2010), Using pedotransfer functions to estimate the van Genuchten–Mualem soil hydraulic properties: A Review, *Vadose Zone J.*, *9*(4), 795.
Vrugt, J. A., and C. J. F. Ter Braak (2011), DREAM(D): An adaptive Markov Chain Monte Carlo simulation algorithm to solve discrete, noncontinuous, and combinatorial posterior parameter estimation problems, *Hydrol. Earth Syst. Sci.*, *15*(12), 3701–3713.
Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.
Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, *41*, W01017, doi:10.1029/2004WR003059.
Whitaker, J. S., and T. M. Hamill (2012), Evaluating methods to account for system errors in ensemble data assimilation, *Mon. Weather Rev.*, *140*(9), 3078–3089.
Wigneron, J. P., et al. (2007), L-band Microwave Emission of the Biosphere (L-MEB) model: Description and calibration against experimental data sets over crop fields, *Remote Sens. Environ.*, *107*(4), 639–655.
Wilheit, T. T. (1978), Radiative-transfer in a plane stratified dielectric, *IEEE Trans. Geosci. Remote Sens.*, *16*(2), 138–143.
Yang, X. S., and T. Delsole (2009), Using the ensemble Kalman filter to estimate multiplicative model parameters, *Tellus, Ser. A*, *61*(5), 601–609.
Zacharias, S., et al. (2011), A network of terrestrial environmental observatories in Germany, *Vadose Zone J.*, *10*(3), 955–973.